MODELING AND EQUALIZATION TECHNIQUES FOR DISPERSIVE RAYLEIGH FADING CHANNELS

by

Guanghua Deng

Master of Science
Simon Fraser University, 1993

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF APPLIED SCIENCE
in the School of Engineering Science

SIMON FRASER UNIVERSITY
April 1996

ISBN 0-612-16857-3

APPROVAL

Name: Guanghua Deng

Degree: Master of Applied Science

Title of thesis: Modeling and Equalization Techniques for Dispersive Rayleigh Fading Channels

Examining Committee:

Chairman: *Dr. S. Stapleton*

Dr. J.K. Cavers
Senior Supervisor

Dr. P. Ho
Senior Supervisor

Dr. J. Vaisey
Supervisor

Dr. J. Bird
Examiner

Date Approved: April 15, 1996

PARTIAL COPYRIGHT LICENSE

Title of Thesis/Project/Extended Essay

"Modeling and Equalization Techniques For Dispersive Rayleigh Fading Channels"

Author:

_____
(signature)

G. Deng
(name)

April 17, 1996
(date)

Abstract

The mobile and indoor wireless communications environment is usually modelled as a frequency selective Rayleigh fading channel. Frequency selectivity implies that the received signal exhibits intersymbol interference (ISI). When transmitting with high data rates over these channels, the ISI becomes severe. One common techniques to combat the ISI is the adaptive equalization.

In this thesis, we propose a simplified channel model and develop two maximum likelihood sequence estimators for the Rayleigh fading channel with ISI.

Instead of using the whole channel impulse response tap gains to describe the Rayleigh fading channel as usual, the proposed reduced dimensionality channel model uses only a few random gain parameters to represent the whole discrete channel impulse response. The analysis of the modeling error shows that for channels having small delay spread the modeled channel do not deviate much from the original channel. This model is utilized in the development of an adaptive receiver using Viterbi decoder. The receiver uses a known inserted pilot training sequence to estimate the multipath time varying channel. The error performance of the receiver is analyzed and it is demonstrated that the receiver has simple structure and good bit error rate performance. We also design in the thesis a maximum likelihood sequence estimation - Viterbi algorithm (MLSE-VA) receiver that does not need the training sequence in the channel estimator. Techniques, such as the signal spatial diversity, receiver multiple sampling, and VA state reduction, are investigated to improve the error performance and reduce the complexity of the adaptive receiver. The performance of this receiver is evaluated through simulation. The numerical results show that the pilot assisted receiver has better performance and less computational complexity than the receiver that does not use pilot sequence.

Acknowledgements

I am deeply grateful to my senior supervisor, Dr. Jim Cavers, for suggesting the topic of this project, his invaluable guidance and advice, the financial support for my graduate studies and the time he spent with me throughout this thesis.

I wish to express sincere thanks to my senior supervisor, Dr. Paul Ho, for constant guidance and suggestions for my studies. I appreciate very much his time and effort.

I would like to thank Dr. Shawn Stapleton, Dr. Jacques Vaisey and Dr. John Bird for their advice and comments.

I especially thank the School of Engineering Science, Simon Fraser University for providing the opportunity for my graduate studies. My special thanks are expressed to Mrs. Brigitte Dorner, Mr. Edward Lo and others for their great assistance.

I acknowledge in all sincerity the encouragement and help from my fellow students and friends.

Contents

List of Figures

List of Abbreviations

AWGN - additive white Gaussian noise

BEP - bit error probability

BER - bit error rate

CIR - channel impulse response

CSI - channel state information

DFE - decision feedback equalizer

ISI - intersymbol interference

LMS - least mean square

MLSE - maximum likelihood sequence estimation

MSD - mean square difference

RDCM - reduced dimensionality channel model

RLS - recursive least square

RMS - root mean square

SNR - signal to noise ratio

VA - Viterbi algorithm

WSSUS - wide sense stationary uncorrelated scattering

Chapter 1

Introduction

Technical problems of digital wireless communications can be categorized into various aspects, such as speech coding, coded and uncoded channel modulation, networking protocols, cryptography, adaptive receivers and so on and so forth. This thesis deals with adaptive receivers, in particular the maximum likelihood sequence (MLSE) receiver, for frequency selective Rayleigh fading channels. In this introductory chapter we give an general overview of the popular techniques employed in adaptive receivers and a brief thesis outline. Since wireless receiver technology depends heavily on the signal propagation environment, we begin our introduction with a brief review of digital mobile channel characteristics.

1.1 Characterization of the multipath fading propagation channel

The wireless communications environment is characterised by random time-varying impulse response. Many mobile radio and indoor wireless communication channels exhibit frequency selective Rayleigh fading, which produces randomly time varying intersymbol interference (ISI) in digital transmission [1]. The Telecommunications

Industry Association has characterized the digital cellular channel as a frequency-selective Rayleigh fading channel.

Basically, the frequency-selective Rayleigh fading channel is characterized by two distinct phenomena, i.e. **multipath** and **Doppler effects** [2].

Multipath channel is where energy arrives via several paths, usually as a result of reflections, or of inhomogeneities in the physical medium that produce ray-splitting or scattering effects. Thus the received signal (apart from any additive noise) can be viewed as a weighted sum of delayed versions of the transmitted waveform. As the terminal is moved from one location to another or as people move around close to the transmitter or the receiver, the multipath affects the expected average received signal power and causes the received power to fluctuate. In most cases, the equivalent amplitude and phase fluctuations of the sum of the received signals have the same statistical characteristics as those of narrow-band additive white Gaussian noise (AWGN), and the corresponding envelope is Rayleigh-distributed, so the transmission channel is a Rayleigh fading channel. The distortion caused by the multipath is called frequency selectivity, which means the different frequency components of the transmitted signals are subject to different fading effects and thus the signal is severely distorted by the channel. The frequency selectivity depends quantitively on the multipath spread, or equivalently the coherence bandwidth of the channel, relative to the transmitted signal bandwidth.

Another important phenomenon in the mobile wireless transmission is the Doppler effect. Whenever a transmitter and a receiver are in relative motion, the received carrier frequency differs from the transmitted carrier frequency. This shift of frequency is refered to as the Doppler Shift and can be calculated according to the velocity of the mobile and the carrier frequency [1]. In a realistic environment, the received signal arrives along multipath and the velocity of movement in the direction of each arriving path is in general different from that of another path. Thus a transmitted

sinusoid, subjected not only to a single Doppler shift, is received as a spectrum which is refered to as the Doppler spectrum. This effect of spreading of the transmitted signal frequency is called in a general way as the Doppler spread of the channel, and the resulting signal attenuation is referred to as Doppler fading. Doppler spread also occurs when the propagation path characteristics is time varying. Doppler fading is grossly described by the coherence time or, equivalently, by the Doppler spread.

A detailed mathematical description of the Rayleigh fading multipath channel as well as its parameters and discrete model can be found in [2] and [3] and is also discussed in Chapter 2.

Due to the multipath, the maximum digital symbol transmission rate for a wireless communication system is bounded by the delay spread. As the symbol transmission rate increases, the duration of a transmitted symbol becomes comparable to or even smaller than the delay spread. As a result, the pulses arriving along multiple paths associated with one symbol interval will interfere with the multiple pulses associated with the neighboring symbol intervals. This intersymbol interference (ISI) is manifested as an irreducible error rate observed at high signal to noise ratio (SNR). As the symbol transmission rate increases, the worsening intersymbol interference created by the multipath increases the irreducible error rate of the receiver. Various techniques in the receiver can be used to increase the data rate for ISI channel, e.g. diversity reception, RAKE matched filter, adaptive equalization techniques, as well as employing modulation and coding techniques [2]. We will concentrate this introduction on the development of the equalization techniques which lead to optimum adaptive receiver.

1.2 Adaptive Equalization Techniques

For ISI channels, the communication receiver often adopts adaptive equalization techniques to achieve better performance [2]. For example, the European and North America digital wireless communications standards GSM and IS-54 both have the equalizer at the receiver [4].

An adaptive equalizer is an adaptive filter at the receiver whose frequency response adapts approximately to the inverse of the frequency response of the channel. Equalization techniques can be subdivided into two general types - linear and nonlinear equalization [5].

The linear equalizer is usually implemented as a FIR or IIR filter with appropriate coefficients adaption algorithms [2], e.g. least mean square (LMS) algorithm and recursive least square (RLS) algorithm [6]. In an attempt to compensate for the channel distortion, the linear equalizer places a large gain in the vicinity of the spectral null and as a consequence significantly enhances the additive noise present in the received signal. Thus it does not perform well on channels with spectral nulls in their frequency response characteristics [2]. Therefore in channels with severe frequency-selective fading, which often result in spectral nulls in received signal, the widely used techniques are the adaptive nonlinear equalizers, which utilize decisions to either cancel the interference or enhance the signal.

Two types of very effective adaptive nonlinear equalizer have been studied by many researchers over the past three decades, i.e. the adaptive decision feedback equalizer (DFE) [5] [7] [8] [9], and the adaptive version of the maximum-likelihood sequence estimation (MLSE) [10] [11] [12] [13] [7]. A summary of the nonlinear equalization techniques, DFE and MLSE, is presented in [15] . A detailed description of the DFE and MLSE structures and the associated class of algorithms that are employed to adaptively adjust the equalizer parameters is given in [5] [2].

The basic idea in the DFE is that once an information symbol has been detected, the ISI that it causes on future symbols may be estimated and subtracted out prior to symbol detection. To accomplish this, DFE adopts two sections of filters, a feedforward section and a feed back section. The received signal sequence with ISI is sent into the feedforward section, while the feedback section has as its input the sequence of decisions on previous detected symbols. The function of the feedback section is to remove the part of the ISI from the present estimate caused by previously detected symbols [2]. In an adaptive mode, the coefficients of the feedforward and feedback filters are adjusted recursively to follow time variations in the channel response. LMS algorithm, RLS algorithm and their derivatives are used for this purpose. With these rapidly adapting algorithm DFE can track the fast fading channel.

Compared with the linear equalization, the DFE yields a significant improvement in performance, mainly due to the inclusion of the decision-feedback section which eliminates the ISI from previously detected symbols. However simulation results show that there is still a significant degradation in performance of the DFE due to the residual intersymbol interference, especially on channels with severe distortion [2]. Also there is a significant performance loss due to incorrect decisions being feed back, i.e. error propagation, in DFE, caused by the sensitivity of the LMS and RLS algorithms to error propagation [14].

Among different equalization techniques, maximum likelihood sequence estimation, which is efficiently implemented by means of the Viterbi algorithm (MLSE-VA) is optimal in the sense that it minimizes the probability of the error in the receiver [2] [10] [11] [15] [16].

The MLSE is based on the computation of a-posteriori probabilities, with a known received signal sample sequence. Hypotheses are made for all the possible transmitted information sequences. The detector chooses, among these hypothetic sequences, the one that has the largest posteriori probability. The VA is used for efficiently searching

among these possible sequences.

Unlike the linear and decision-feedback equalization techniques, the MLSE needs to know the probability density function of the received signal. In linear equalization, the estimate of the information symbol in any particular signaling interval is some linear combination of the received signal samples from the output of the matched filter. In DFE, there is the additional linear combination of previously detected symbols. The tap weights in these linear combinations are determined by minimizing either the mean square error or the peak distortion. Thus the adjustment of the weights does not require knowledge of the statistics of the signal and the channel. In MLSE, in order to calculate the a-posteriori probability and hence the decision metric in VA, the statistics of the transmission environment must be known a-priori.

To eliminate the ISI, the MLSE-VA algorithms also require explicit or implicit identification of the channel characteristics in order to compute the metrics for making decisions. To accommodate a channel that is unknown, a channel estimator may be included in parallel with the detection algorithm. The channel estimator is usually an FIR transversal filter with adjustable coefficients which is computed recursively to minimize the mean square error (MSE) between the actual received sequence and the output of the estimator. These estimated coefficients are fed to the MLSE-based VA for use in the metric computations. Considerable research has been performed on the channel estimation and the corresponding digital signal structure [15].

These studies have resulted in two main channel estimation techniques. One suggests the use of a known training sequence prior to each information sequence to help the channel estimation. This technique is commonly known as the pilot assisted detection techniques [17] - [20]. With this approach the signaling format of the system is designed such that the information data sequence is organized in blocks with each block preceded by a known training pilot sequence. The training sequence at the begining of each block is used to estimate the channel response within the training

sequence duration. The channel response at other times is tracked either through the fast tracking algorithms like LMS and RLS algorithms or through a channel interpolator [20]. Generally, the whole channel estimation process is performed in two stages, a periodic channel estimation stage and a channel tracking or interpolation stage [21]-[23]. The main concerns with the channel estimation using the pilot training are the processing delay, the signalling efficiency, and the buffer space required at the receiver for interpolation.

In order to accomplish the fast fading channel identification in a short time, the other proposed technique estimates jointly the channel response and the transmitted sequence, despite the dilemma of the so-called Blind Problem : " Sequence detection is never successful without the knowledge of the channel response. But the channel response cannot be estimated without referring to the transmitted sequence." Joint channel response estimation and signal detection is usually carried through a received signal whitening filter implemented as a Kalman filter. The filtered signal output is then compared with the original received signal to determined the decoding metric in MLSE-VA algorithm [24] [25] [26] [27] [28] [29].

The main considerations in these channel estimation methods are the receiver complexity, propagation lag error, and the signal detection delay. A comparision of the advantages and disadvantages of these techniques can be found in [22].

This thesis studies the optimal equalization techniques, i.e. MLSE-VA, in the frequency selective Rayleigh fading channel. The objective is to derive an efficient and yet accurate equalizer structure to combat the ISI. The channel estimation both with and without the pilot sequence assistance is investigated. The pilot sequence assisted channel estimation scheme proposed in the thesis utilizes a novel reduced dimensionality channel model developed in Chapter 3, which results in a receiver with simple structure and not much loss of performance. The MLSE-VA under study without training sequence uses a bank of Wiener optimal filters for the fading channel

prediction. Its application under fast and slow fading environments are studied. Techniques such as multi-sampling per symbol at the receiver and the antenna diversity are investigated. The thesis outline is given below.

1.3 Thesis Outline

Following the introductory chapter, the signal and channel models for frequency selective Rayleigh fading channel are developed in Chapter 2. Concepts, such as fast and slow fading, mean square delay spread are defined and discussed in detail. In order to develop a simple and precise equalizer, the multipath Rayleigh fading channel characterization is studied and a new reduced dimensionality channel model (RDCM) is proposed in Chapter 3. The channel considered is the wide sense stationary uncorrelated scattering (WSSUS) channel, which specifies that the attenuation and phase shift of the channel associated with path delay τ_1 is uncorrelated with those associated with path delay τ_2. This is the case in most radio transmission media [29] [30] . The derived simplified discrete channel model has only a few random parameters, when the delay spread is less than the 20% of the signal symbol duration. The estimation can be carried out for these parameters instead of the whole channel impulse response. Applying this novel technique, both the simplicity and accuracy of the channel estimator can be increased as found in Chapter 3.

With the RDCM derived in Chapter 3, a reduced complexity channel estimator is proposed in Chapter 4. This channel estimator utilizes the RDCM and pilot training sequence to enhance the channel estimation accuracy. The obtained estimator is thus simple in structure and in computation and yet very accurate. The performance of the estimator is evaluated both under the mean squared measure criteria and bit error probability of the corresponding VA receiver that uses this channel estimator. The result is very satisfactory.

In Chapter 5 of the thesis, an MLSE-VA equalizer using the joint channel estimation and signal detection is studied. No training sequences are used. An implicit channel estimation technique and the VA state reduction scheme is applied so that the complexity of the conventional MLSE-VA is greatly reduced without loss of accuracy in the sense of the bit error rate of the receiver.

The thesis is summarized in Chapter 6.

Chapter 2

Signal and Channel Model

In this chapter, we define the digital signal model over the frequency selective Rayleigh fading channel. Most of the material can be found in [2]. Concepts of delay power profile, fast and slow fading are discussed. The discrete signal and channel model derived for the signaling over Rayleigh channels with intersymbol interference and additive white gaussian noise will be applied throughout the thesis.

2.1 Transmitted Signal

The baseband model of a high-speed communication system over a frequency selective Rayleigh fading channel is shown in Fig. 2.1.

The transmitted data is represented by a sequence $\{c(k)\}$ of complex symbols, and the complex envelope of the transmitted signal is

$$s(t) = A \sum_{k=-\infty}^{\infty} c(k)\, q_t(t - kT) \tag{2.1}$$

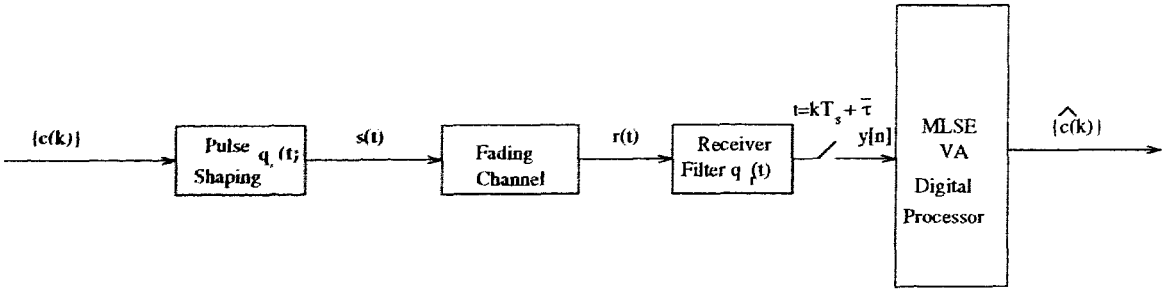

Figure 2.1: Block Diagram of the Communication System.

where A is the amplitude, and without loss of generality we set $A = 1$ in the following of the thesis. The transmitter pulse shaping function $q_t(t)$ has unit energy, and the symbol rate is $\frac{1}{T}$. The expected value $E[|c(k)|^2] = 1$.

2.2 Rayleigh Fading Dispersive Channel

The signal is transmitted over a frequency selective Rayleigh fading channel with an equivalent low-pass time-variant impulse response $u(\tau, t)$, which represents the channel response at time t due to an impulse applied at time $t - \tau$. τ is the delay in the scatters. $u(\tau, t)$ is usually modeled as a wide sense stationary uncorrelated scattering (WSSUS) process [39]. For Rayleigh fading, $u(\tau, t)$ is a Gaussian process with zero mean. It is completely characterized by its correlation function which is defined as [2]

$$\begin{aligned} \phi_u(\xi, \tau_1, \tau_2) &= \frac{1}{2} E[u(\tau_1, t) u^*(\tau_2, t + \xi)] \\ &= \phi_u(\xi, \tau_1) \delta(\tau_1 - \tau_2) \end{aligned} \tag{2.2}$$

where $*$ represents the complex conjugate. $\phi_u(\xi, \tau)$ gives the average power output as a function of the time delay τ and the difference between observation times, ξ. It reflects both the multipath spread and the fading rate of the channel. When assuming the autocorrelation function in time is the same for each delay, $\phi_u(\xi, \tau)$ is

usually represented as

$$\phi_u(\xi, \tau) = G(\tau)\phi_u(\xi) \tag{2.3}$$

where $G(\tau)$ is refered to as the delay power profile, and $\phi_u(\xi)$ is the autocorrelation function normalized to unit power.

The total fading power is normalized to unity, so that

$$\int G(\tau)d\tau = 1 \tag{2.4}$$

For convenience, we will take the first moment of $G(\tau)$, the average delay, to be zero, as given by

$$\int \tau G(\tau)d\tau = 0 \tag{2.5}$$

The second central moment of $G(\tau)$ is the mean square delay spread, with square root equal to the root mean square (rms) delay spread, given by

$$\tau_{rms}^2 = \int \tau^2 G(\tau)d\tau \tag{2.6}$$

The channel additive noise, $n_w(t)$ is a zero mean complex white Gaussian process with power density N_0. The baseband equivalent of the received signal, $r(t)$, is

$$\begin{aligned} r(t) &= u(\tau, t) \otimes s(t) + n_w(t) \\ &= \int u(\tau, t)s(t-\tau)d\tau + n_w(t) \end{aligned} \tag{2.7}$$

$\otimes$ represents the convolution.

2.3 A General Discrete Model for Rayleigh Fading Channel

The receiver front end consists of a filter followed by a sampler. Denote the impulse response of the receiver filter by the unit energy $q_r(t)$, Its output is

$$\begin{aligned} y(t) &= [r(t) + n_w(t)] \otimes q_r(t) \\ &= \sum_k c(k) \int \int u(\tau, t+\alpha) q_t(t+\alpha-kT-\tau) q_r(-\alpha) d\tau d\alpha + n(t) \end{aligned} \quad (2.8)$$

where $n(t)$ is the filtered white Gaussian noise with autocorrelation function

$$\phi_n(t) = \frac{N_0}{2} \int q_r(x) q_r(x-t) dx \quad (2.9)$$

To simplify the general expression (2.8), we need to specify the pulse shaping functions $q_t(t)$ and $q_r(t)$ as well as the channel impulse reponse $u(\tau, t)$, which all rely on the fading channel characteristics. For different fading characteristics the pulse shaping filters can be chosen to optimize the receiver.

In slow fading, it is usual to match the receiver filter to the transmitted pulse, which usually has a square root raise cosine shape, in order to maximize the detected signal to noise ratio. In this case, $q_r(t) = q_t^*(-t)$. Although the integration in (2.8) with respect to α covers the whole α axis, most of energy of the integrand centered around peaks of $q_r(-\alpha)$ and $q_t(t+\alpha-kT-\tau)$, so the practical integration region lies in a small interval from about $\alpha = 0$ to about $\alpha = kT + \tau - t$, the peak locations of the two pulses $q_r(-\alpha)$ and $q_t(t+\alpha-kT-\tau)$. Therefore for slow fading, the channel impulse response $u(\tau, t+\alpha)$ in (2.8) is considered approximately constant over the integration with respect to α. By arbitrarily identifying this value with the value at $\alpha = 0$, we can move $u(\tau, t)$ outside the α integration. Thus we have

$$y(t) = \sum_k c(k) \int u(\tau, t) p(t - kT - \tau) d\tau + n(t) \tag{2.10}$$

where $p(t) = q_t(t) \otimes q_r^*(-t)$ is the composite pulse shape.

In very fast fading, pulse distortion from the multiplicative fading can result in ISI and loss of SNR, resulting in a significant error floor [32]. One of several remedies is to make the receiver filter a rectangular lowpass with bandwidth large enough to accommodate the Doppler-spread signal, and transmit full Nyquist pulses. Therefore with $q_t(t)$ a full Nyquist pulse, usually a raised cosine pulse, and $q_r(t)$ a lowpass filter, (2.8) can still be simplified to (2.10), with $p(t) = q_t(t)$.

We see in both slow and fast fading, the received signal can be at least approximately represented as in (2.10) with $p(t)$ being a full Nyquist pulse. In this thesis, we take $p(t)$ as a raised cosine pulse.

The filter output is sampled before further processing as shown in Fig. 2.1. The sampling instant is $nT_s + \bar{\tau}$ and for convenience in this study, we assume that the timing recovery algorithm locks to the mean delay, so that $\bar{\tau}$ equals zero. Generally to combat the fast fading in the channel, multiple samples are made in per symbol duration. Let the number of samples per symbol be an integer

$$N_s = \frac{T}{T_s} \tag{2.11}$$

where $\frac{1}{T_s}$ is the sampling rate at the receiver filter output.

To simplify the notation we can define an equivalent transmission sequence as

$$b(k) = \begin{cases} c(k/N_s) & \text{if } k/N_s \text{ is integer} \\ 0 & \text{otherwise} \end{cases} \tag{2.12}$$

consisting of the information data interleaved with $(N_s - 1)$ zeros. The information sequence $\{c(k)\}$ of rate $\frac{1}{T}$ is then extended to an equivalent sequence $\{b(k)\}$ of rate $\frac{1}{T_s}$. Note that for sampling at transmission baud rate, i.e. $N_s = 1$, we have $T = T_s$ and $b(k) = c(k)$.

The sampled filter output at time nT_s is

$$y(nT_s) = \sum_k b(k) \int u(\tau, nT_s) p[(n-k)T_s - \tau] d\tau + e(nT_s) \tag{2.13}$$

where $e(nT_s)$ is a filtered white noise sample with correlation given in (2.9). For receiver matched filter and one sample per baud case, i.e. $N_s = 1$, the different $e(nT_s)'s$ are independent identically distributed (iid) complex Gaussian random variables, each having a zero mean and a variance equal to N_0.

Substitute (2.11) into (2.13) and use square braces to indicate a sample of a continuous waveform, e.g., $y[n] = y(nT_s)$, we have

$$y[n] = \sum_k b(k) \int u(\tau, nT_s) p[(n-k)T_s - \tau)] d\tau + e[n] \tag{2.14}$$

This is a general discrete signal model. In the following we will derive the discrete signal model for the discrete multipath Rayleigh fading channel.

2.4 A Signal Model over the Discrete Multipath Rayleigh Fading Channel

Section 2.2 outlined a model for the continuous fading dispersive channel. For discrete multipath Rayleigh fading channel, the impulse response of the time varying dispersive

fading channel is given by

$$u(\tau, t) = \sum_{i=1}^{M} \beta_i(t)\delta(\tau - \tau_i). \tag{2.15}$$

We adopt a WSSUS model, so that the paths $(\beta_i(t), \tau_i)$, $i = 1, 2, \ldots, M$ are statistically independent of each other, giving

$$\frac{1}{2}E[\beta_i^*(t)\beta_j(t+\xi)] = \begin{cases} \phi_i(\xi) & \text{for } i = j \\ 0 & \text{otherwise.} \end{cases} \tag{2.16}$$

We further assume that the channel scattering is separable, so that each delay has the same normalized autocorrelation function, resulting in

$$\phi_i(\xi) = \sigma_i^2 \rho(\xi) \tag{2.17}$$

where σ_i^2 is the power of the i-th path and $\rho(\xi)$ is the autocorrelation function normalized to unit power. In mobile communications, $\rho(\xi)$ is frequently taken to be $J_0(2\pi f_D \xi)$, where f_D is the maximum Doppler frequency [1]. The delay power profile is

$$G(\tau) = \sum_{i=1}^{M} \sigma_i^2 \delta(\tau - \tau_i) \tag{2.18}$$

For convenience, we take the mean delay to be zero

$$\sum_{i=1}^{M} \tau_i \sigma_i^2 = 0 \tag{2.19}$$

so that the root mean squared delay spread, τ_{rms}, is given by

$$\tau_{rms}^2 = \sum_{i=0}^{M} \tau_i^2 \sigma_i^2 \tag{2.20}$$

Substituting (2.15) into (2.10) and taking samples at nT_s, we obtain the discrete signal model over discrete multipath fading channel

$$y[n] = \sum_{k} \sum_{i=1}^{M} b(k)\beta_i[n]p[(n-k)T_s - \tau_i] + e[n] \tag{2.21}$$

where the information sequence $\{c(k)\}$ is replaced by the equivalent sequence $\{b(k)\}$.

In the remainder of the thesis, for simplicity, we will use the signal model for discrete multipath fading channel for all the derivations. The derivations can be extended to the continuous fading channels, i.e. according to (2.14), without much difficulty.

In the next chapter we will derive a reduced dimensionality channel model based on the signal model and notations developed in this chapter.

Chapter 3

A Reduced Dimensionality Propagation Model for Frequency Selective Rayleigh Fading Channels

As discussed in the last chapter, the mobile communication channels exhibit frequency selective Rayleigh fading, which produces randomly time varying intersymbol interference (ISI) in digital transmission. For ISI channels, the optimum detector uses maximum likelihood sequence estimation (MLSE), normally with the Viterbi algorithm. However, the Viterbi receiver requires an estimate of the channel impulse response (CIR) and, depending on the pulse shaping filter used and the delay spread in the channel, the CIR may span many symbols, resulting in a large number of parameters to estimate. The estimation can be facilitated by embedding known training sequences in the data [21]. According to [20], the number of training symbols required

is approximately twice the duration of the CIR. Thus, treating the CIR as a long unconstrained sequence leads to questions of accuracy in the estimation and efficiency in the transmission format.

We present in this chapter a simple, yet accurate, alternative model to describe frequency selective channel fading. It expresses the CIR in terms of only a few parameters, rather than explicitly in terms of the full set of channel taps as is usually done [2]. There are several advantages with such a reduced dimensionality model (RDM). First, since there are fewer parameters involved, they can be estimated more accurately; second, significantly fewer training symbols are needed when compared to a system that estimates explicitly the full CIR; and third, simulations of such channels are simplified through reduction of the number of complex gain generators required. We show that a RDM for the frequency selective fading channel with a small delay spread can be obtained by truncating a Taylor series representation of the transmitted signal. In a related work [31], a different mathematical transformation of CIR, i.e. channel orthogonalization, is shown to lead to a simplied channel simulator, though it has not been applied to detection in receivers.

The chapter is organized as follows. We show in Section 3.1 how the original channel model can be described in terms of a reduced dimensionality model. The accuracy of the RDM, measured in terms of the mean squared error between the original and modeled channel impulses responses, as well as the bit error rate difference of the Viterbi receiver, is given in Section 3.2. Finally the summary of the chapter is given in Section 3.3.

3.1 Reduced Dimensionality Channel Model

In this section we propose the reduced dimensionality channel model for the slow fading channel discussed in chapter 2. The block diagram of our communication system is shown in Fig. 3.1.

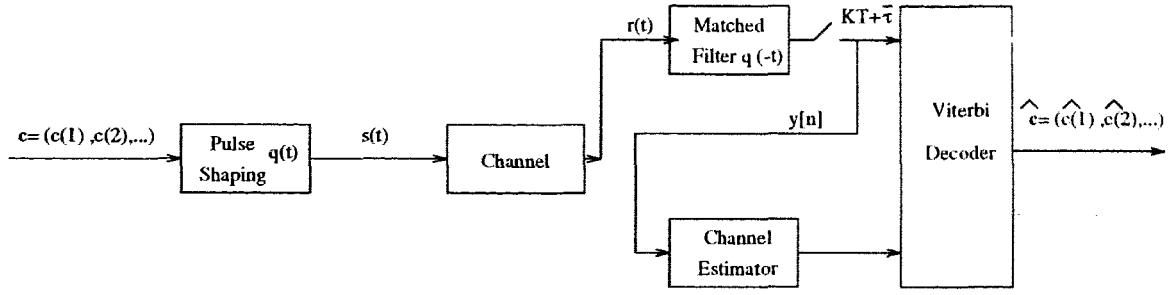

Figure 3.1: Block diagram of the system.

Recall from (2.10) that

$$y(t) = \sum_k c(k) \int u(\tau, t) p(t - kT - \tau) d\tau + n(t) \tag{3.1}$$

where $p(t) = q_t(t) \otimes q_t^*(-t)$ is the composite pulse shape.

As shown in Fig. 3.1, the receiver samples $y(t)$ at the symbol rate so that in expressions (2.11) and (2.12) we have $T = T_s$, $N_s = 1$ and $b(k) = c(k)$. According to (2.14), the received signal sample is

$$y[n] = \sum_k c(k) \int u(\tau, nT) p[(n - k)T - \tau] d\tau + e[n] \tag{3.2}$$

where $e[n]$ are iid complex Gaussian random variables.

It is clear from (3.2) that the ISI, and therefore the discrete impulse response, can span many symbol times, even for relatively modest delay spreads. Use of the

impulse response to characterize the channel, therefore, leads to high dimensionality in both channel estimation and channel simulation. We now show that an alternative description gives an effective dimensionality that is much lower for small to moderate delay spreads.

Since the transmitted signal is bandlimited, we can expand it in a Taylor series and expect that only the first few terms are required for accurate representation over a significant fraction of a symbol. Accordingly, we expand the pulse $p((n-k)T-\tau_i)$ in (3.2) as

$$p((n-k)T-\tau_i) = \sum_{j=0}^{\infty} \frac{p^{(j)}((n-k)T)(-\tau_i)^j}{j!}. \tag{3.3}$$

where $p^{(j)}(t)$ is the j-th derivative of the pulse $p(t)$. By substituting (3.3) into (3.2) and changing the time index by letting $n-k=\ell$ and hence $k=n-\ell$, we can rewrite $y[n]$ as

$$y[n] = \sum_{j=0}^{\infty} g_j[n] \sum_{\ell} c(n-\ell) p_j[\ell] + e[n] \tag{3.4}$$

where

$$g_j[n] = \int u(\tau, nT)(-\frac{\tau}{T})^j d\tau$$

$$p_j[\ell] = p^{(j)}(\ell T)\frac{T^j}{j!}$$

We can interpret (3.4) as a sum of fixed parallel subchannels, in which the j^{th} impulse response is $p_j[\ell]$, each weighted by a complex gain $g_j[n]$, see Fig. 3.2.

The $g_j[n]$, which are Gaussian random variables, are successive moments (in delay τ) of complex channel gain. They can also be interpreted in frequency domain as derivatives at $f=0$ of the Fourier transform

$$U(f,t) = \int u(\tau,t) e^{-j2\pi f\tau} d\tau \tag{3.5}$$

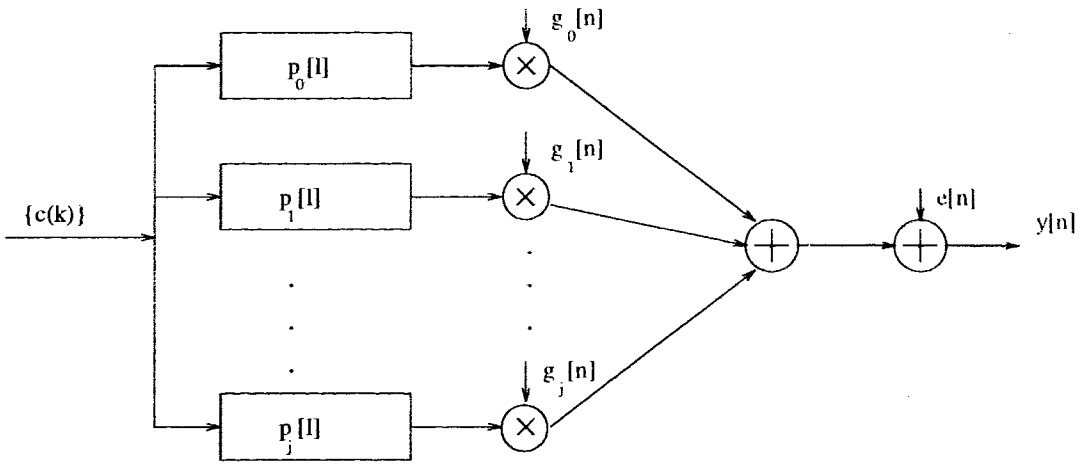

Figure 3.2: The parallel subchannel model.

$$g_j[n] = (\frac{1}{j2\pi})^j \frac{dU(f,nT)}{df}|_{f=0} \tag{3.6}$$

i.e. tilt, curvature, etc of $U(f,t)$.

The correlation of $g_j[l]$ and $g_k[l-n]$ can be obtained using (2.2) and (2.3) for the WSSUS channel

$$\begin{aligned}
\lambda_{jk}[n] &= \frac{1}{2}E\{g_j[\ell]g_k^*[l-n]\} \\
&= \frac{1}{2}\int\int u(\tau,lT)(-\frac{\tau}{T})^j u^*(\sigma,(l-n)T)(-\frac{\sigma}{T})^k d\tau d\sigma \\
&= \int\int \phi_u(nT,\tau,\sigma)(-\frac{\tau}{T})^j(-\frac{\sigma}{T})^k d\tau d\sigma \\
&= \int G(\tau)\phi_u(nT)(-\frac{\tau}{T})^{j+k} d\tau \\
&= \phi_u(nT)\int G(\tau)(-\frac{\tau}{T})^{j+k} d\tau
\end{aligned} \tag{3.7}$$

Note the integration in (3.7) is the moments of the channel power delay profile defined in section 2.2.

In the following, we take the WSSUS discrete channel as an example to derive our reduced dimensionality channel model. By substituting (2.15) - (2.17) as appropriate, we can obtain $y[n]$ as

$$y[n] = \sum_{\ell} c(n-\ell)h_{\ell}[n] + e[n] \tag{3.8}$$

where

$$h_{\ell}[n] = \sum_{j=0}^{\infty} g_j[n]p_j[\ell], \tag{3.9}$$

$$g_j[n] = \sum_{i=1}^{M} \beta_i[n] \left(\frac{-\tau_i}{T}\right)^j, \tag{3.10}$$

and

$$p_j[\ell] = p^{(j)}(\ell T)\frac{T^j}{j!} \tag{3.11}$$

The set of coefficients $\{\ldots, h_{-1}[n], h_0[n], h_1[n], \ldots\}$ is the discrete time channel impulse response (CIR) at time nT. These are complex Gaussian random variables constructed as set of fixed, known impulse responses $p_j[\ell]$ of increasing order j, each weighted by a time-varying complex gain $g_j[n]$. Fig. 3.3 depicts the discrete time channel model in (3.9), where altogether $2L+1$ channel tap coefficients are shown.

The reduction in dimensionality stems from the observation that for a small delay spread, the infinite sum in (3.9) can be approximated by the first $J+1$ terms, as shown by the covariance

$$\lambda_{jk}[n] = \frac{1}{2}E\{g_j[l]g_k^*[l-n]\} = \sum_i \phi_i[n] \left(\frac{\tau_i}{T}\right)^{j+k} \tag{3.12}$$

If all the τ_i's are much less than the symbol duration T, then the covariance decreases rapidly with increasing order $j+k$, and we can truncate the sum in (3.9) after the Jth derivative. This leads to the reduced dimensionality channel model

$$\bar{h}_{\ell}[n] = \sum_{j=0}^{J} g_j[n]p_j[\ell] \tag{3.13}$$

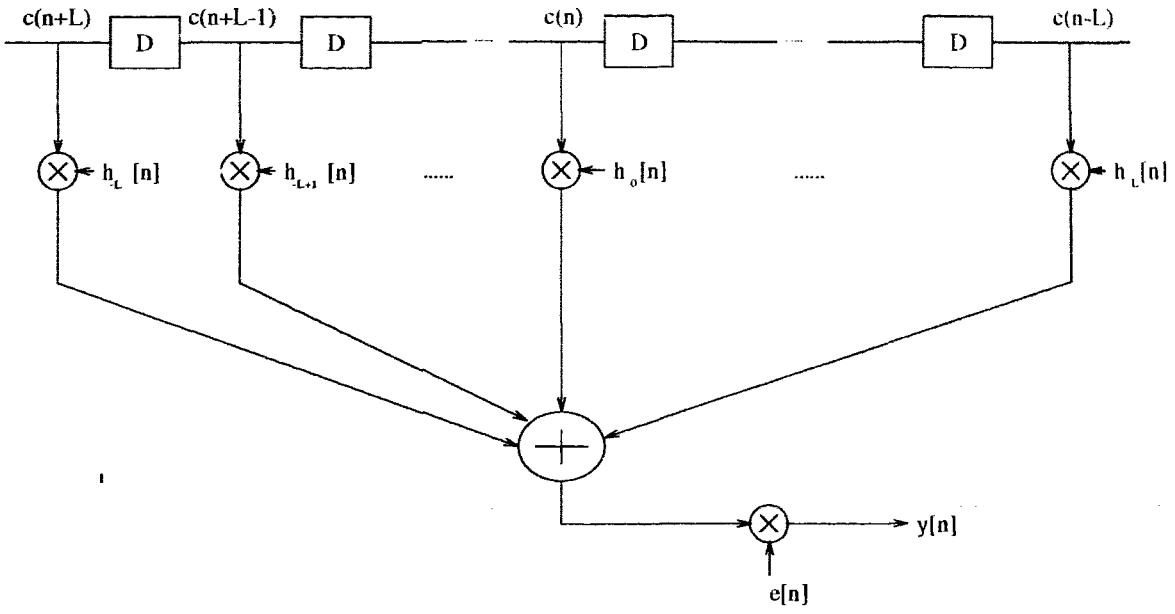

Figure 3.3: The discrete time channel model.

where J is the order of the channel model. With a proper choice of J, we see from (3.13) that the entire channel response, which strictly speaking consists of an infinite number of taps, is well approximated by only $J+1$ random processes. We will show later on that for delay spreads up to 20% of a symbol, only $J = 2$ derivatives are required.

According to (3.7), the covariance (3.12) can also be expressed as

$$\begin{aligned} \lambda_{jk}[n] &= \rho[n] \sum_i \sigma_i^2[n] \left(\frac{\tau_i}{T}\right)^{j+k} \\ &= \rho[n] M_G^{(j+k)} \end{aligned} \tag{3.14}$$

where $M_G^{(k)}$ is the kth central moment of the delay-power profile, the second such moment being the mean square delay spread. Thus we have an expansion in the moments of the delay spread, which decrease quickly with order. In particular, we have assumed that timing recovery locks to the mean delay, so that the first moment is zero and the processes $g_0[n]$ and $g_1[n]$ are independent. If a different value is returned

from timing recovery, the moments are non-central, and the first two processes are correlated.

The reduced dimensionality model has several advantages. Since there are fewer parameters, the receiver is likely to estimate them more accurately than in the case of estimating the full impulse response $h_\ell(n)$ (The point is illustrated by examples in Appendix A). On the other hand, the result is still an approximation of the true CIR. Similarly, a simulator based on the reduced dimensionality model can generate the effects of a fading dispersive channel with fewer random number generators.

3.2 Accuracy of the Model

Having established the reduced dimensionality channel model, we now evaluate its accuracy. Two criteria are used for this purpose. One is the mean squared difference between the CIRs of the RDM and the original model. The other is the bit error probability difference of the respective Viterbi receivers. Numerical results are given in both cases.

3.2.1 Mean Squared Modeling Error

One measure for the accuracy of the RDM is the mean squared difference (MSD) between the original and the modeled responses of the channel, i.e.,

$$\begin{aligned} E_h &= \frac{1}{2}\sum_{\ell} E\left\{\left|\bar{h}_\ell[n] - h_\ell[n]\right|^2\right\} \\ &= \sum_{\ell}\frac{1}{2}E\{(\sum_{j=0}^{J} g_j[n]p_j[\ell] - \sum_{j=0}^{\infty} g_j[n]p_j[\ell])(\sum_{j=0}^{J} g_j[n]p_j[\ell] - \sum_{j=0}^{\infty} g_j[n]p_j[\ell])^*\} \\ &= \sum_{\ell=-L}^{L}\sum_{i=J+1}^{\infty}\sum_{j=J+1}^{\infty}\sum_{k=1}^{M}\sigma_k^2\left(\frac{\tau_k}{T}\right)^{i+j} p_i[\ell]p_j^*[\ell] \end{aligned} \tag{3.15}$$

For a numerical comparison, we used for the composite pulse shape $p(t)$ a raised cosine characteristic with 50% excess bandwidth, truncated to $L = 5$; that is, we assume that most of the signal energy is contained in the center 11 taps of the discrete CIR. Fig. 3.4 shows the dependence of MSD on the normalized rms delay spread (τ_{rms}/T) for a channel with a uniform power-delay power profile. A similar plot is given in Fig. 3.5 for a two-ray profile. These two profiles are given respectively by

$$G(\tau) = \begin{cases} \frac{1}{2\tau_{max}} & -\tau_{max} \leq \tau \leq \tau_{max} \\ 0 & \text{otherwise} \end{cases} \tag{3.16}$$

with rms delay spread equal to $\tau_{max}/\sqrt{3}$ and

$$G(\tau) = \sigma_1^2\delta(\tau - \tau_1) + \sigma_2^2\delta(\tau - \tau_2), \tag{3.17}$$

with mean delay $\sigma_1^2\tau_1 + \sigma_2^2\tau_2 = 0$ and rms delay spread equal to $\sqrt{\sigma_1^2\sigma_2^2(\tau_2 - \tau_1)^2}$. Since both channels have a normalized power of unity, the MSDs shown in Figs 3.4 and 3.5 actually represent normalized differences between the original and the modeled CIR's.

Evident from these figures is that for rms delay spread less than 5% of the symbol duration, a first order model, i.e. $J = 1$, is sufficient to represent the actual channel with a MSD around 10^{-5}. On the other hand, if a flat fading model is assumed, i.e $J = 0$, then the MSD is approximately 10^{-2} for a 5% delay spread. Although there may not exist a simple relationship between the MSD and the bit error rate of the receiver, it is clear from the above figures that a receiver design based on a flat fading assumption will exhibit an error floor even with a few percent of delay spread.

If the rms delay spread in the channel is up to 10% of a symbol, then a second order model, $J = 2$, is needed to keep the MSD to within 10^{-5} for the 2-ray channel. The channel with the uniform profile requires a higher model for this range of delay spread. Though not shown, we have calculated the MSD for different power split ratios in the two ray channel and find that the more evenly distributed the power is in the two rays, the smaller the MSD between the reduced dimensionality channel

model and the original channel model. This is because, for a given rms delay spread, more uneven power splits correspond to larger delay differences between the two rays, and therefore require more terms in the Taylor series for sufficient accuracy.

3.2.2 Bit Error Probability of the Viterbi Receiver

In this section, we assess the effect of the reduced dimensionality model on the bit error probability (BEP) of an MLSE receiver implemented using the Viterbi algorithm. Specifically, we compare a receiver that determines the first $J+1$ complex gains $g_j[n]$ of the RDM CIR with perfect accuracy, using them in a Viterbi detector, with a similar receiver that has perfect knowledge of the true CIR. To evaluate the BEP of the Viterbi receivers, we apply the method in [33].

If perfect CSI can be obtained for $\bar{h}_\ell[n]$ in (3.13), the optimal receiver is a Viterbi decoder that selects the sequence $\hat{\mathbf{c}} = (\hat{c}(1), \cdots, \hat{c}(K))$ that minimizes the metric [2]

$$J(\hat{\mathbf{c}}) = \sum_{n=1}^{K} \left| y[n] - \sum_{\ell=-L}^{L} \hat{c}(n-\ell)\bar{h}_\ell[n] \right|^2 \tag{3.18}$$

where $2L$ is the length of the CIR and $c(-L), \cdots, c(0)$ are known to the receiver. Given the transmitted sequence $\mathbf{c} = (c(1), \cdots, c(K))$, a decoding error occurs if for some erroneous sequence $\hat{\mathbf{c}}$, the random variable D, defined as

$$D = J(\hat{\mathbf{c}}) - J(\mathbf{c}) \tag{3.19}$$

is less than zero. The metric $J(\mathbf{c})$ is given as

$$J(\mathbf{c}) = \sum_{n=1}^{K} |y[n] - \sum_{\ell=-L}^{L} c(n-\ell)\bar{h}_\ell[n]|^2 \tag{3.20}$$

Let $P(\mathbf{c} \rightarrow \hat{\mathbf{c}})$ denote the pairwise error event probability for a particular pair of $\mathbf{c}$ and $\hat{\mathbf{c}}$, i.e the chance that D is less than zero. In addition let $a(\mathbf{c}, \hat{\mathbf{c}})$ be the Hamming

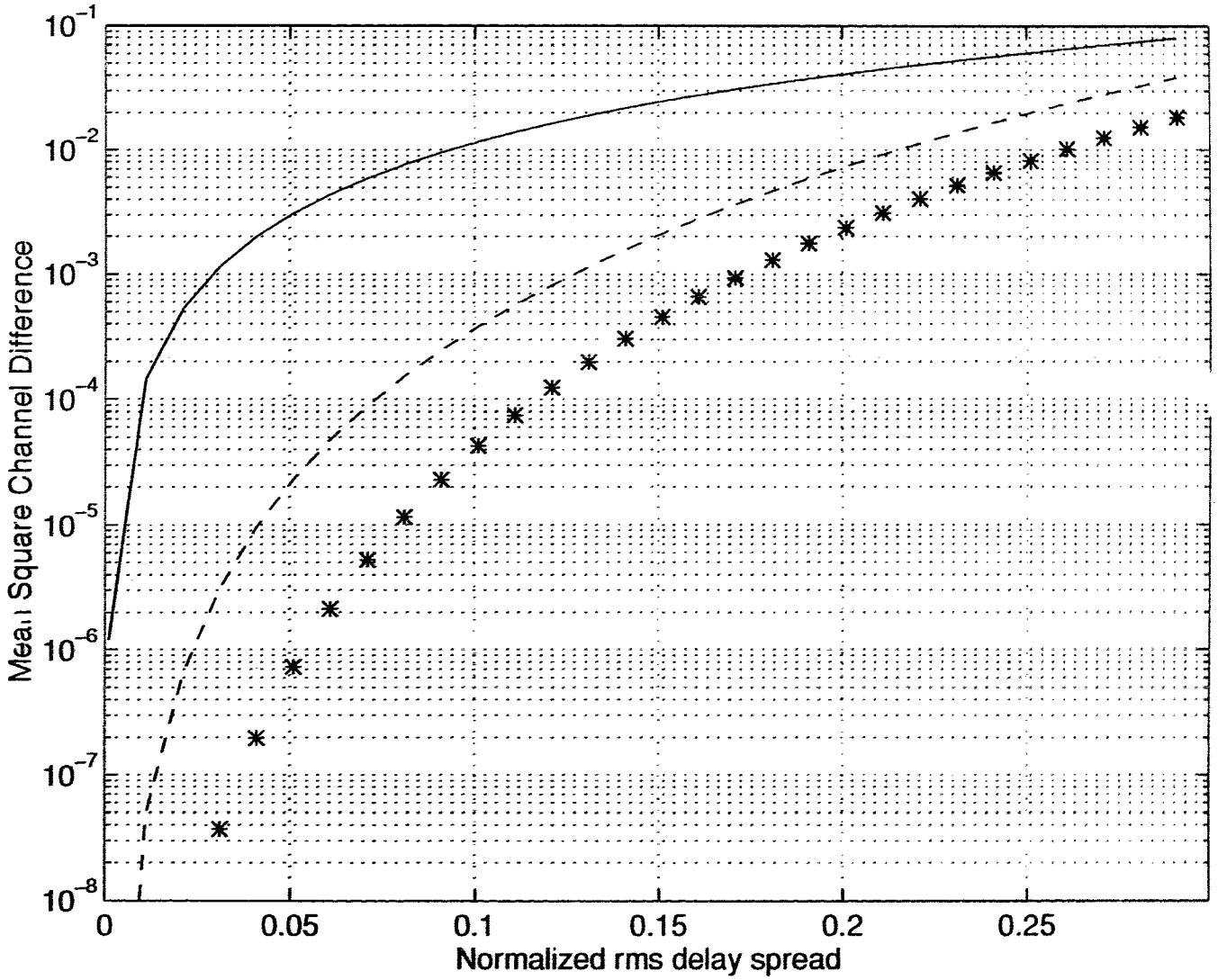

Figure 3.4: The mean squared difference vs normalized rms delay spread plot for a channel with a uniform power-delay profile; solid line : J=0, dashed line : J=1, and star : J=2.

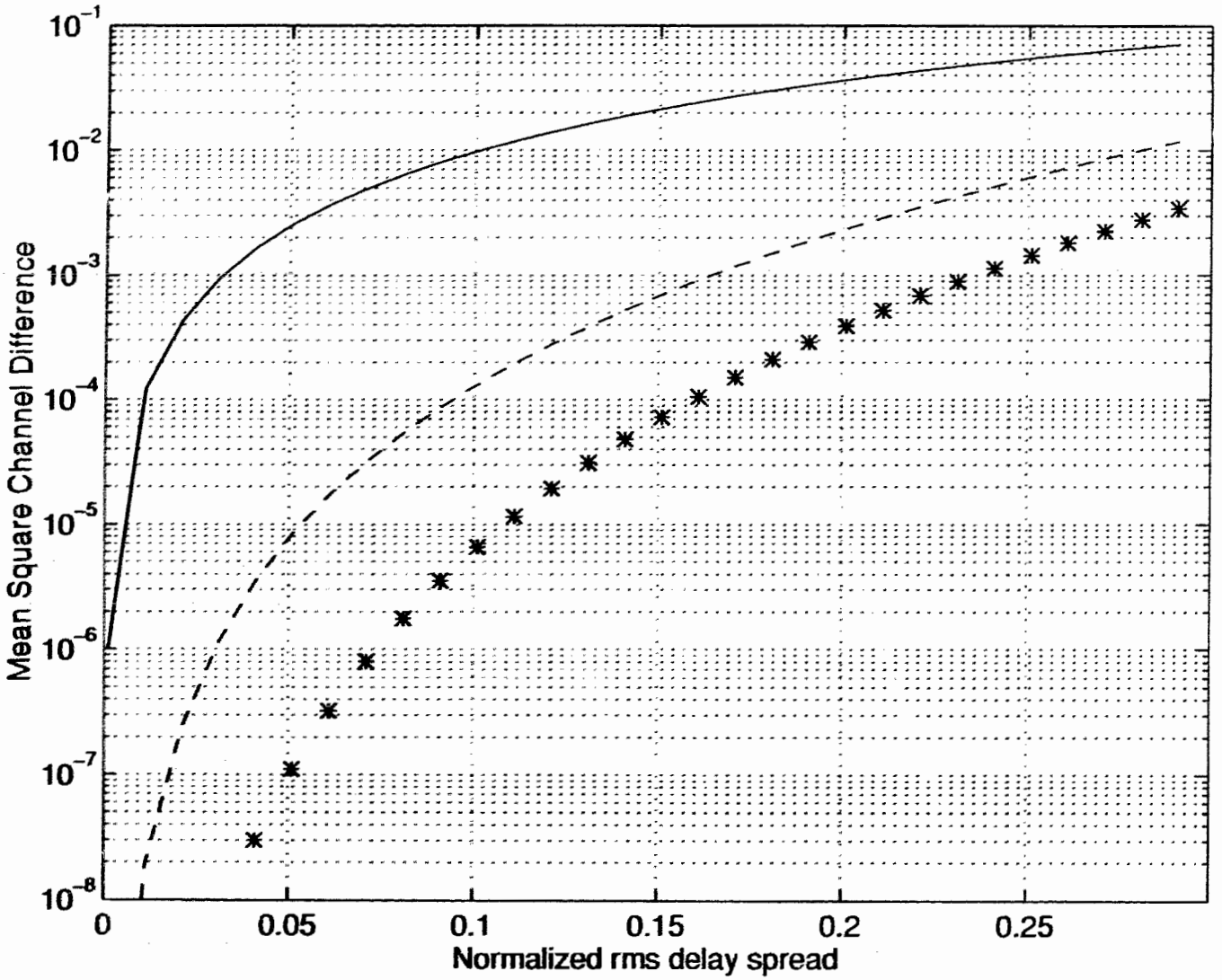

Figure 3.5: Same as the previous figure except that a 2-ray profile is considered

distance between the two corresponding information sequences. Then the overall bit error probability of the Viterbi receiver can be approximated as [33] [34]

$$P_b = \frac{1}{Z}\sum_{\mathbf{c}} \frac{a(\mathbf{c},\hat{\mathbf{c}})}{B} \sum_{\hat{\mathbf{c}}\neq\mathbf{c}} P(\mathbf{c}\rightarrow\hat{\mathbf{c}}) \tag{3.21}$$

where B is the number of the information bits per channel symbol, and Z is the number of the possible transmitted sequences. It is obvious from this equation that the key to determining the bit error performance of the receiver lies in obtaining an expression for the pairwise error probability. Using the characteristic function approach suggested in [33] we can show that the pairwise error probability is given by

$$\begin{aligned} & P(\mathbf{c}\rightarrow\hat{\mathbf{c}}) \\ = \; & p(D<0) \\ = \; & -Residue[\Phi_D(s)/s]_{Rpoles} \end{aligned} \tag{3.22}$$

where $\Phi_D(s)$ is the characteristic function of the random variable D. The structure of the characteristic function can be found in equation (4B.1) of [35].

The details in obtaining the characteristic function $\Phi_D(s)$ is shown below.

Substitute (3.8) into (3.18) and (3.20) respectively, we have

$$J(\hat{\mathbf{c}}) = \sum_{n=1}^{K} |e[n] + \sum_{\ell=-L}^{L} d(n-\ell)h_\ell[n] + \sum_{\ell=-L}^{L} \hat{c}(n-\ell)\Delta h_\ell[n]|^2 \tag{3.23}$$

and

$$J(\mathbf{c}) = \sum_{n=1}^{K} |e[n] + \sum_{\ell=-L}^{L} c(n-\ell)\Delta h_\ell[n]|^2 \tag{3.24}$$

where $d(k) = c(k) - \hat{c}(k)$, $\Delta h_\ell[n] = h_\ell[n] - \hat{h}_\ell[n] = \sum_{j=J+1}^{\infty} g_j[n]p_j[\ell]$.

Define zero mean complex Gaussian random variables a_n, b_n and t_n as follows.

$$a_n = \sum_{\ell=-L}^{L} d(n-\ell)h_\ell[n], \tag{3.25}$$

$$b_n = \sum_{\ell=-L}^{L} \hat{c}(n-\ell)\Delta h_\ell[n], \tag{3.26}$$

$$t_n = \sum_{\ell=-L}^{L} c(n-\ell)\Delta h_\ell[n]. \tag{3.27}$$

Let

$$\mathbf{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_K \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_K \end{bmatrix}, \qquad \mathbf{t} = \begin{bmatrix} t_1 \\ \vdots \\ t_K \end{bmatrix}, \qquad \mathbf{e} = \begin{bmatrix} e[1] \\ \vdots \\ e[K] \end{bmatrix}, \tag{3.28}$$

$$\mathbf{I} = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}_{K\times K}, \qquad \mathbf{0} = \begin{bmatrix} 0 & & 0 \\ & \ddots & \\ 0 & & 0 \end{bmatrix}_{K\times K}. \tag{3.29}$$

Write the complex random variable D in matrix form

$$\begin{aligned} D &= J(\hat{\mathbf{c}}) - J(\mathbf{c}) && (3.30)\\ &= \sum_{n=1}^{K}[(e[n]+a_n+b_n)(e[n]+a_n+b_n)^* - (e[n]+t_n)(e[n]+t_n)^*] \\ &= \mathbf{z}^\dagger\mathbf{F}\mathbf{z} && (3.31) \end{aligned}$$

where

$$\mathbf{z} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{t} \\ \mathbf{e} \end{bmatrix}_{4K\times 1}, \tag{3.32}$$

and

$$\mathbf{F} = \begin{bmatrix} \mathbf{I} & \mathbf{I} & \mathbf{0} & \mathbf{I} \\ \mathbf{I} & \mathbf{I} & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} & -\mathbf{I} & -\mathbf{I} \\ \mathbf{I} & \mathbf{I} & -\mathbf{I} & \mathbf{0} \end{bmatrix}_{4K \times 4K} . \tag{3.33}$$

The superscript † represents the Hermitian transpose.

The correlation matrix of $\mathbf{z}$ is obtained formally without much difficulty.

$$\begin{aligned} \mathbf{R_z} &= \frac{1}{2} E \mathbf{z}^* \mathbf{z}^t \\ &= \frac{1}{2} E \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{t} \\ \mathbf{e} \end{bmatrix}^* \begin{bmatrix} \mathbf{a} & \mathbf{b} & \mathbf{t} & \mathbf{e} \end{bmatrix} \\ &= \frac{1}{2} E \begin{bmatrix} \mathbf{a}^*\mathbf{a} & \mathbf{a}^*\mathbf{b} & \mathbf{a}^*\mathbf{t} & \mathbf{a}^*\mathbf{e} \\ \mathbf{b}^*\mathbf{a} & \mathbf{b}^*\mathbf{b} & \mathbf{b}^*\mathbf{t} & \mathbf{b}^*\mathbf{e} \\ \mathbf{t}^*\mathbf{a} & \mathbf{t}^*\mathbf{b} & \mathbf{t}^*\mathbf{t} & \mathbf{t}^*\mathbf{e} \\ \mathbf{e}^*\mathbf{a} & \mathbf{e}^*\mathbf{b} & \mathbf{e}^*\mathbf{t} & \mathbf{e}^*\mathbf{e} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{R_{aa}} & \mathbf{R_{ab}} & \mathbf{R_{at}} & \mathbf{0}_{K \times K} \\ \mathbf{R_{ba}} & \mathbf{R_{bb}} & \mathbf{R_{bt}} & \mathbf{0}_{K \times K} \\ \mathbf{R_{ta}} & \mathbf{R_{tb}} & \mathbf{R_{tt}} & \mathbf{0}_{K \times K} \\ \mathbf{0}_{K \times K} & \mathbf{0}_{K \times K} & \mathbf{0}_{K \times K} & \mathbf{R_{ee}} \end{bmatrix} \end{aligned} \tag{3.34}$$

Note that $\mathbf{e}$ is uncorrelated with $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{t}$, and $\mathbf{R_z}$ is a Hermitian matrix. Determination of the submatrices of $\mathbf{R_z}$ is straightforward, but detailed. To simplify the expression, define

$$
\begin{aligned}
& \mathcal{F}(\mathbf{w}, \mathbf{v}, n_1, n_2, j_{1l}, j_{1h}, j_{2l}, j_{2h}) \\
= & \sum_{\ell=-L}^{L} \sum_{m=-L}^{L} w(n_1 - l) v^*(n_2 - m) \sum_{j_1=j_{1l}}^{j_{1h}} \sum_{j_2=j_{2l}}^{j_{2h}} p_{j_1}[\ell] p^*_{j_2}[m] \\
& [\sum_{i=1}^{M} \phi_i(n1 - n2) \tau^{j_1 + j_2}].
\end{aligned} \tag{3.35}
$$

where $\mathbf{w}$ and $\mathbf{v}$ are vectors with elements $\{w(i), i = n1 - L, \cdots, n1 + L\}$ and $\{v(i), i = n1 - L, \cdots, n1 + L\}$ respectively.

The $(n_1, n_2)'th$ element, $n_1,\ n_2 = 1, \cdots, K$, in submatrices $R_{\mathbf{aa}}$, $R_{\mathbf{ab}}$, $R_{\mathbf{at}}$, $R_{\mathbf{bb}}$, $R_{\mathbf{bt}}$, $R_{\mathbf{tt}}$, and $R_{\mathbf{ee}}$ can be obtained respectively as follows.

$$
\begin{aligned}
r_{a_{n_1} a_{n_2}} &= \mathcal{F}(\mathbf{d}, \mathbf{d}, n_1, n_2, 0, \infty, 0, \infty), \\
r_{a_{n_1} b_{n_2}} &= \mathcal{F}(\mathbf{d}, \hat{\mathbf{c}}, n_1, n_2, 0, \infty, J+1, \infty), \\
r_{a_{n_1} t_{n_2}} &= \mathcal{F}(\mathbf{d}, \mathbf{c}, n_1, n_2, 0, \infty, J+1, \infty), \\
r_{b_{n_1} b_{n_2}} &= \mathcal{F}(\hat{\mathbf{c}}, \hat{\mathbf{c}}, n_1, n_2, J+1, \infty, J+1, \infty), \\
r_{b_{n_1} t_{n_2}} &= \mathcal{F}(\hat{\mathbf{c}}, \mathbf{c}, n_1, n_2, J+1, \infty, J+1, \infty), \\
r_{t_{n_1} t_{n_2}} &= \mathcal{F}(\mathbf{c}, \mathbf{c}, n_1, n_2, J+1, \infty, J+1, \infty), \\
r_{e_{n_1} e_{n_2}} &= N_0 \delta(n_1 - n_2).
\end{aligned} \tag{3.36}
$$

Where $\mathbf{d} = \mathbf{c} - \hat{\mathbf{c}}$ and all the symbol meanings are consistent throughout the thesis.

With the matrix $\mathbf{F}$ and correlation matrix $\mathbf{R_z}$, the characteristic function of D is obtained as

$$
\begin{aligned}
\Phi_D(s) &= \frac{1}{det(I + 2sR^*_{\mathbf{z}}\mathbf{F})} \\
&= \frac{1}{\Pi_{i=1}^{4N}(1 + 2s\lambda_i)}
\end{aligned} \tag{3.37}
$$

where $\{\lambda_i, i = 1, \cdots, 4K\}$ are $4K$ real eigenvalues of $R^*_{\mathbf{z}}\mathbf{F}$.

The pairwise error probability in (3.22) can be calculated from (3.37).

Using the technique described above, we calculated the bit error probability for a BPSK signal transmitted over the two ray channel defined in section 3.2.1. A raised cosine $p(t)$ with a 50% roll-off is used as the overall pulse shaping filter. Since the main purpose of the analysis is to study the effect of the order of the RDM on the receiver performance, we truncate the overall CIR to 3 taps ($L = 1$). Although the result may not produce the exact BEP, it is nevertheless suitable for comparison of the two receivers, because, for small delay spreads, most of the energy of the CIR, and hence most of the energy of the representation error of the RDM, will be concentrated in the first 3 taps. The motivation, of course, is that use of such a short CIR length greatly simplifies the bit error probability calculation, since there are fewer erroneous sequences to consider. Fig. 3.6 shows the trellis diagram of BPSK for a 3-tap channel, with the circles being the channel states and the numbers associated with the transitions being the BPSK symbols. To simplify our error probability calculation further, we assume a transmitted sequence of $(-1, -1, -1, -1, -1, -1, \ldots)$ and consider only those erroneous paths that merge with it in no more than 4 steps.

Fig. 3.7 shows the computed BEP versus the signal to noise ratio (SNR) for a normalized rms delay spread of 0.1, with the order of the reduced dimensionality channel model as a parameter. Also shown in this figure is the BEP of the optimal receiver, i.e. the one based on the actual CIR (i.e., the infinite order model). A similar plot is shown in Fig. 3.8 for a normalized rms delay spread of 0.2. In both figures, the two rays in the channel are assumed to have equal power. It can be observed from these two figures that a receiver based on the flat fading assumption ($J = 0$) leads to an irreducible error floor. The magnitude of the error floor increases with the actual delay spread in the channel. With a first order model, the error floor can be suppressed significantly, and, with a second order model, it completely

disappears in the range of SNR shown. As a matter of fact, the performance of the receiver based on the second order model is almost indistinguishable from the optimal receiver. Although not shown, we also studied the effect of the power split ratio σ_1^2/σ_2^2 in the 2-ray model on the bit error probability. We found that even power split gives the best performance, and this result is consistent with those found in [36].

Finally, we plot in Fig. 3.9 the BEP vs normalized τ_{rms} curves for different reduced dimensionality channel models. Also shown is the curve for the the optimal receiver. A channel SNR of 40 dB is considered in this figure. It is observed that up to a delay spread of 1/5 of a symbol, the performance of the second order receiver tracks perfectly that of the optimal receiver, whereas the first order receiver tracks perfectly up to about 5% of delay spread. These numbers are consistent with those found in the mean squared channel difference analysis in Figs 3.4 and 3.5. As in [36], it is also observed here that as the delay spread increases, the BEP actually improves. This stems from the diversity effect provided by multipath fading. It is necessary to note that in the North America digital mobile standard, the mobile channel delay spread is 5 10 μs, and the symbol duration is 45μs. So the second order RDCM model is applicable to the standard.

States	Information Bits
0	(-1,-1)
1	(-1 , 1)
2	(1 , -1)
3	(1 , 1)

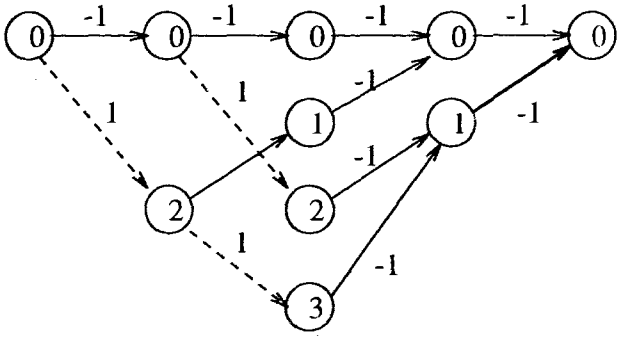

Figure 3.6: Trellis diagram for a BPSK system with a channel memory of 2 symbols

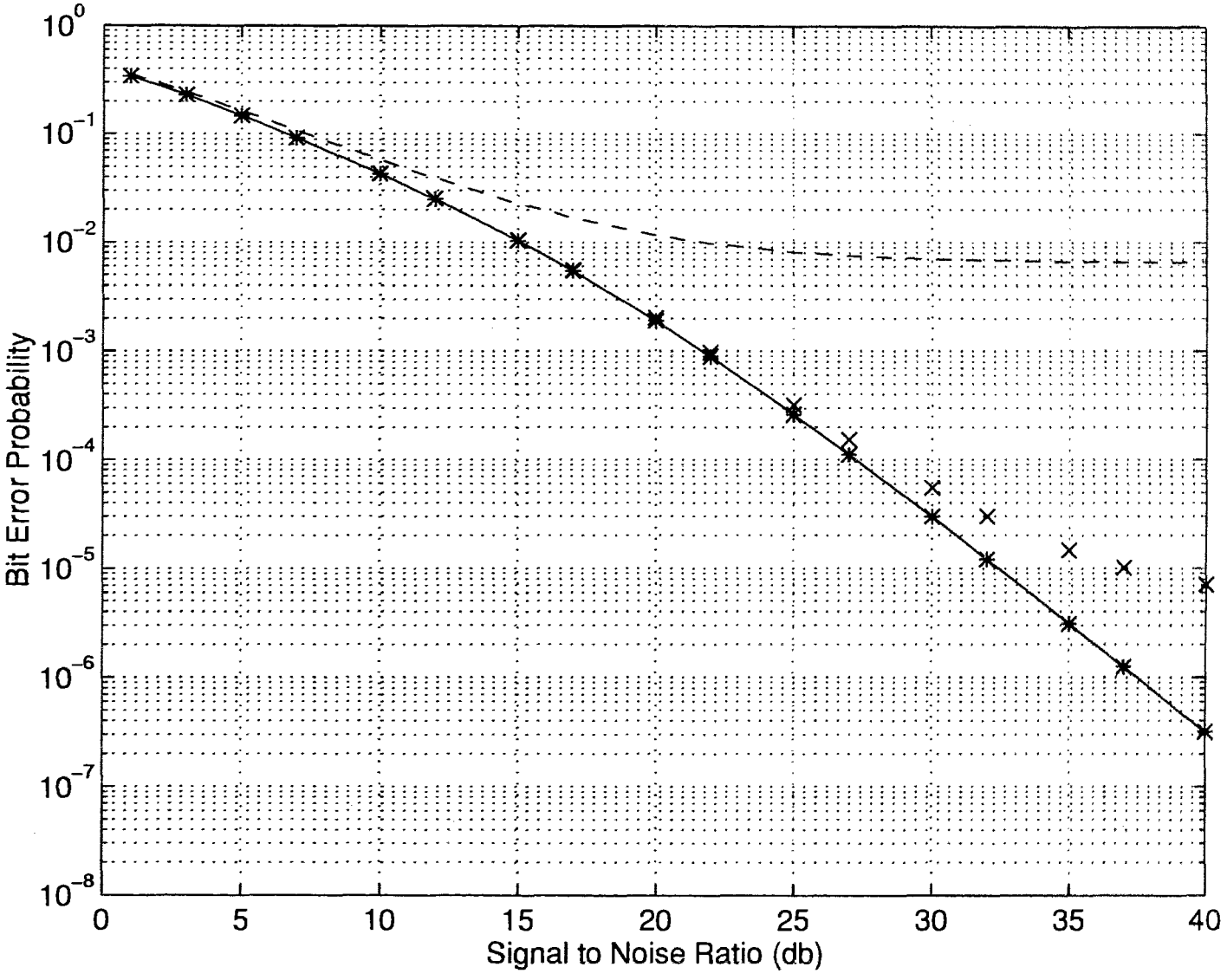

Figure 3.7: Bit error probability vs SNR plot for a 2-ray power-delay profile with a rms delay spread of 0.1 symbol; solid line : full order receiver, dashed line : zero order receiver, crosses : first order receiver, stars : second order receiver.

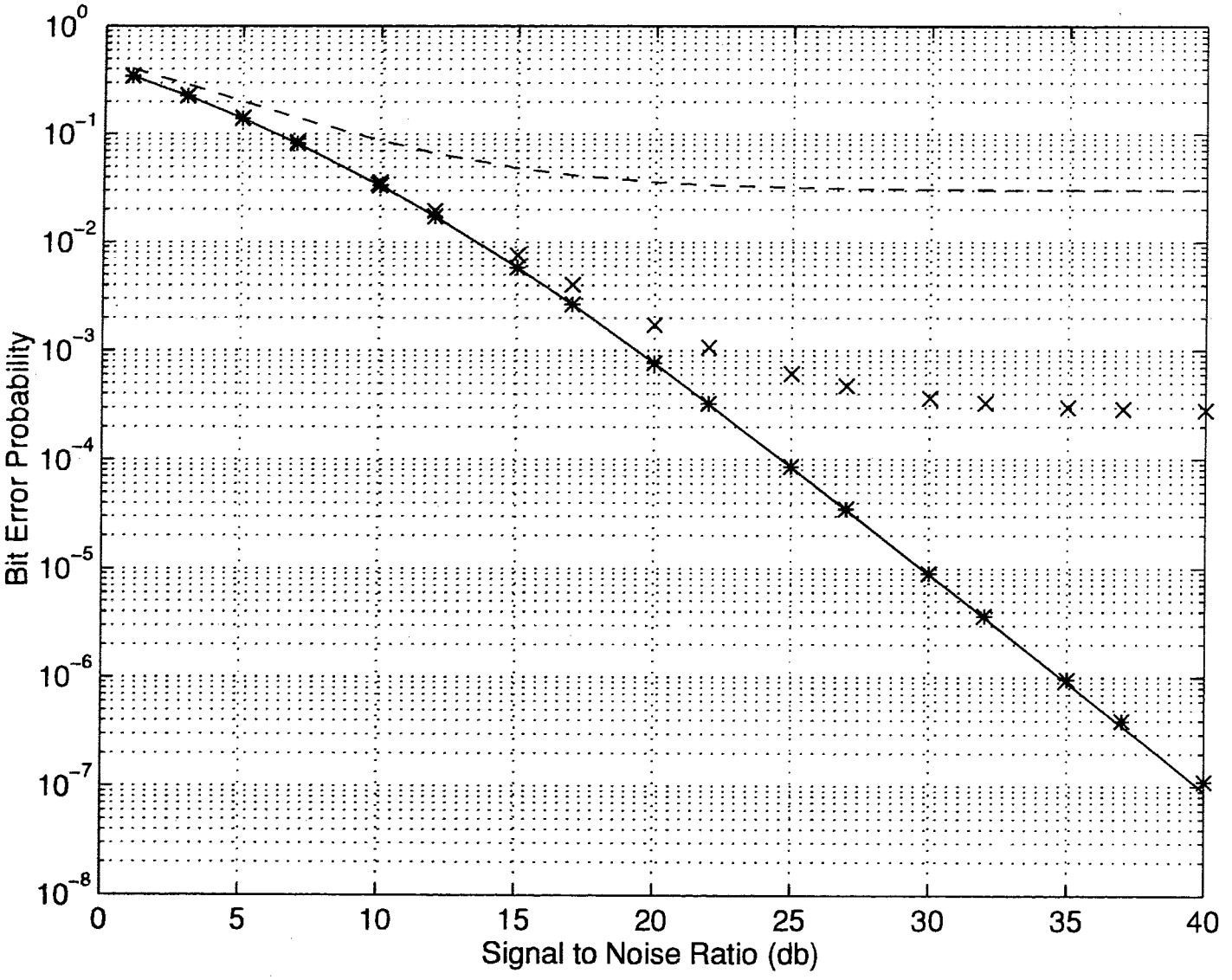

Figure 3.8: Same as the previous figure except that the normalized rms delay spread is increased to 0.2.

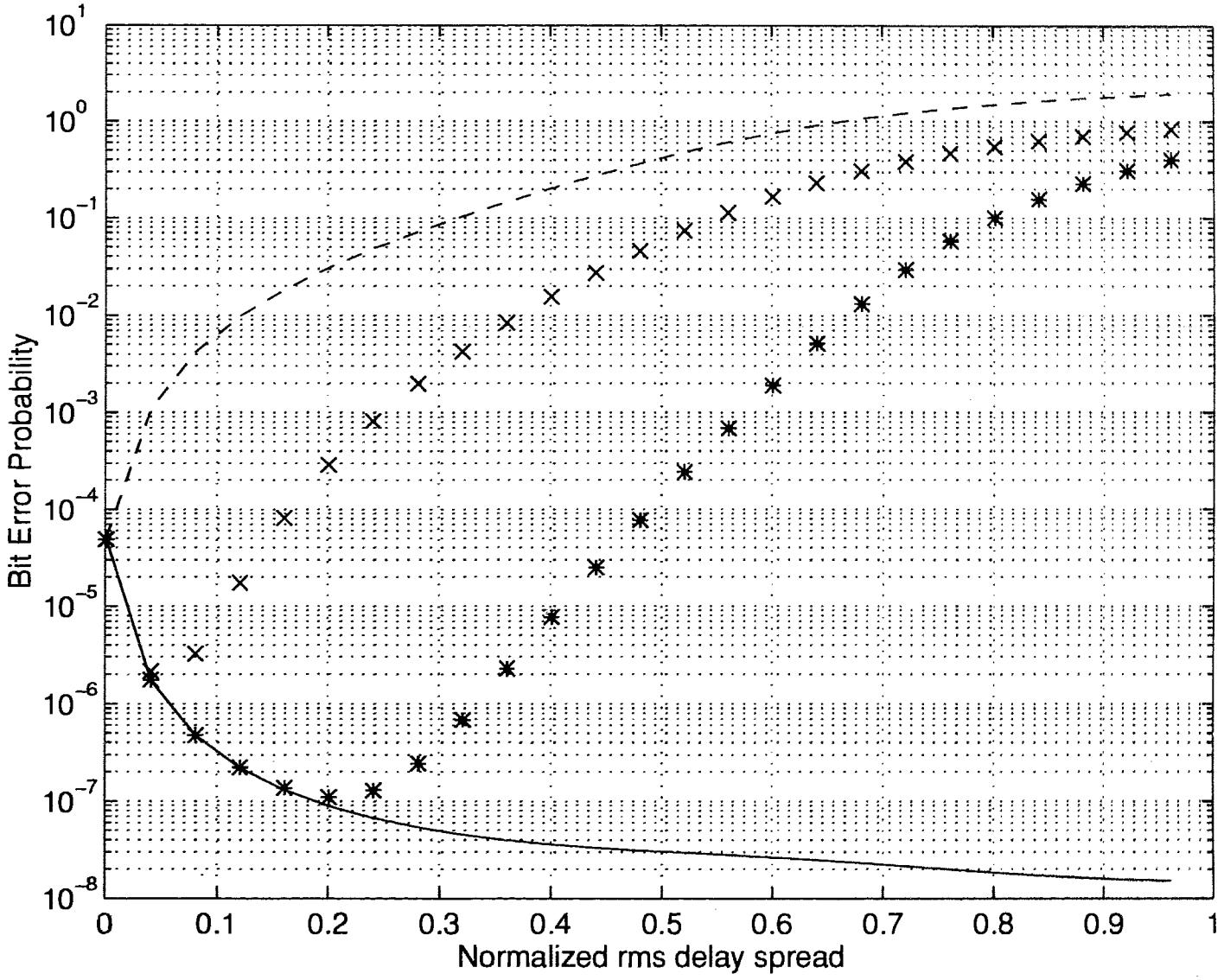

Figure 3.9: Bit error probability vs delay spread plot for a 2-ray power-delay profile; solid line : full order receiver, dashed line : zero order receiver, crosses : first order receiver, stars : second order receiver.

3.3 Conclusions

We have presented a reduced dimensionality channel model for digital transmission over frequency selective Rayleigh fading channels. The model is general, in the sense that, by selecting a proper value for the model order J, we can vary our assumption about the channel from flat fading ($J = 0$) to the original frequency selective fading model ($J = \infty$). For a given delay spread, the correct choice of the model order is determined through a mean squared difference analysis of the original and modeled channel impulse responses. For a delay spread less than 20% of a symbol, we can safely use a second order model to design the receiver without sacrificing any performance. This is confirmed through analysis of the bit error rates of a Viterbi receiver using the RDM CIR and one using the full CIR, both with perfect knowledge of their respective channel parameters. In a real system, estimation of the two or three parameters of the RDM is likely to be more accurate, as well as simpler, than the corresponding estimation of the $2L + 1$ parameters of the full CIR. We can therefore speculate that the an MLSE receiver based on estimating and using the RDM in its metric will have better performance than a similar one based on the full CIR. The channel estimation based on the reduced dimensionality model is the subject of the next chapter.

Chapter 4

A Reduced Complexity Channel Estimator for Linear Modulations Operating in Fading Dispersive Channels

The design and performance of communication systems primarily depend on the channel characteristics and modeling. Using the digital signaling model proposed in chapter 1, we design in this chapter a reduced complexity, pilot symbol assisted channel estimator integrated with the VA receiver for linear modulations operating in the frequency selective Rayleigh fading channel.

As discussed in chapter 1, the optimal adaptive receiver over time varying, multipath radio channels requires identification of the transmission environment in order to eliminate the ISI. One approach for the dispersive fading channel estimation is to use a training sequence; see for examples [17]-[20]. With this approach the signaling format of the system is designed such that the data sequence is divided into frames, with each

frame further divided into a training (or pilot symbol) block and a data block; see Fig. 4.1. The estimation process is carried out in two stages, a periodic channel estimation stage and a channel tracking or interpolation stage [21]-[23]. The advantages of pilot symbol assisted channel estimation over other approaches are summarized in [22].

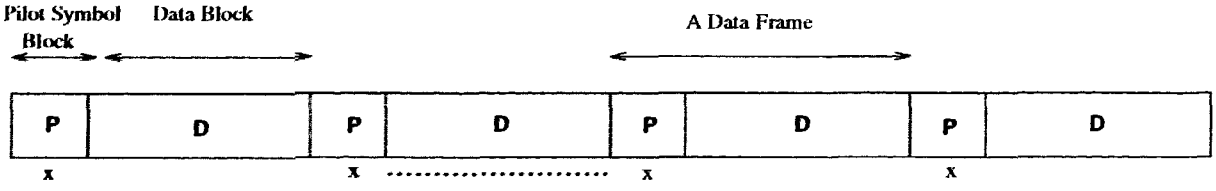

Figure 4.1: Structure of the transmitted frames and the associated channel estimator. A periodic channel estimator is used to estimate the channel impulse responses at those locations marked "X". An interpolator is then used to generate the estimates at those locations marked "...."

In most of the proposed pilot symbol assisted channel estimation techniques (see for examples [20]-[22]), the 'snapshot' assumption is used to simplify the design of the channel estimator. This means the channel impulse response (CIR) is assumed either constant over the entire frame or at least assumed constant over the training block. Though this assumption is probably valid for an indoor microcellular type of operating environment, it imposes a serious limitation on the underlying channel estimation techniques in applications where the delay and Doppler spreads are relatively large. Note that for a channel with an impulse response that spans ℓ symbol intervals, the length of the training block required for accurate channel estimation is approximately $2\ell + 1$ [20] - [22]. For a fast fading channel, such as the one encountered in an outdoor cellular environment, it becomes unreasonable to assume the channel remains constant over the training interval.

We introduce in this chapter a two stage, reduced complexity channel estimation strategy for linear modulations transmitted over a multipath fast fading environment.

The reduction in complexity is a result of using the reduced dimensionality model (RDM) in the last chapter to model the dispersive fading channel. This means only a few channel related parameters, rather than the entire channel impulse response, needed to be estimated. In addition, it also permits us to use shorter pilot symbol blocks, since the length of these blocks now depends on the number of parameters that needed to be estimated, rather than on the length of the CIR. Another feature in our estimator design is that in the periodic channel estimation stage we take into account the fact that the channel parameters can vary significantly over the training block.

The chapter is organized as follows. After a brief review of the RDM for the channel, the optimal filter required for periodic channel parameters estimation is derived in Section 1. Also presented in Section 1 is the optimal interpolation filter. The performance of the channel estimator, measured in terms of the bit error probability of the corresponding Viterbi equalizer, is presented in Section 2. Finally, a summary of the findings in the present study is given in Section 3.

4.1 Channel Estimation Using Pilot Symbols

Use the notations in the last chapter, it was derived that the multipath Rayleigh fading channel, specially the WSSU channel, is modeled as the weighted infinite sum of the random gain processes in (3.9).

It was demonstrated that for small delay spreads, this infinite sum can be approximated by the first $J+1$ terms as follows :

$$\bar{h}_\ell[n] = \sum_{j=0}^{J} g_j[n]\, p_j[\ell]. \tag{4.1}$$

The above represents a reduced dimensionality channel model with J being the order

of the model. It was shown that for a rms delay spread up to 20% of a symbol, the second order model is sufficient to accurately represent the channel impulse response. Consequently, only three fading gain terms, $g_0[k]$, $g_1[k]$, and $g_2[k]$, need to be estimated. The channel estimator is significantly simpler than the one that need to estimate the entire CIR which may contain up to 7 to 10 coefficients.

As shown in Fig. 4.2, the received signal samples $\{y[n]\}$ are passed to the channel estimator. The channel estimator first generates the estimate $\hat{g}_j[k]$ for the derivative fading process $g_j[k]$. Then these estimates are applied to (4.1) to reconstruct an estimate of the CIR. These CIR estimates, denoted by $\{\hat{h}_\ell[n]\}$ are then passed, along with the received samples, to a maximum likelihood sequence estimator (MLSE) implemented as a Viterbi decoder. It should be pointed out that the number of coefficients used in the CIR reconstruction depends strongly on the computational complexity allowed in the Viterbi algorithm; however this potential limitation is not unique to our RDM approach and is common to all Viterbi equalization.

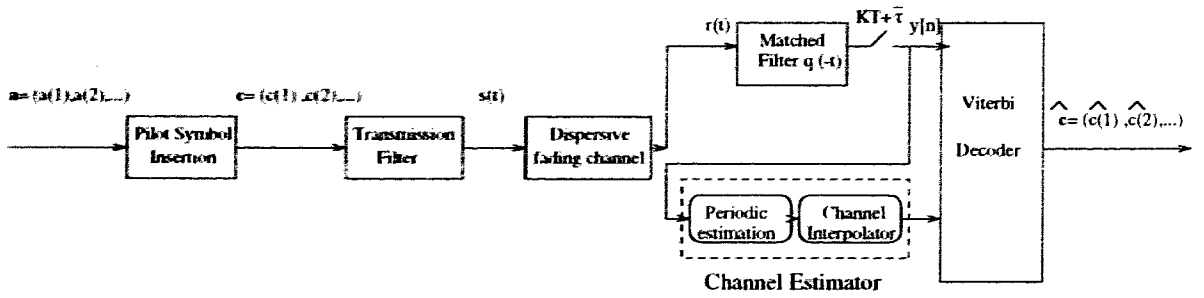

Figure 4.2: System diagram

Since the whole channel estimation is performed in two stages : (1) periodic channel estimation, and (2) interpolation, we will first consider the subject of periodic channel estimation based on the reduced dimensionality channel model in (4.1).

4.1.1 Periodic Channel Estimation

Assuming the channel is accurately represented by the $J+1$ derivative processes $g_0[k], g_1[k], \ldots, g_J[k]$. Then the received sample $y[n]$ can be written as

$$y[n] = \sum_{\ell=-L}^{L} \sum_{j=0}^{J} g_j[n] c(n-l) p_j[\ell] + e[n] \tag{4.2}$$

Without loss of generality, let's assume we want to estimate the derivative processes

$$\mathbf{g}[k] = (g_0[k], g_1[k], \ldots, g_J[k])^t, \tag{4.3}$$

at time $k = 0$, from the received samples

$$\mathbf{Y} = (y[-N], y[-N+1], \ldots, y[N])^t, \tag{4.4}$$

where $(\bullet)^t$ represents a matrix transpose. It can be shown that the optimal estimator, in the minimum mean squared error sense, is given by [37] [38]

$$\tilde{\mathbf{g}} = \left(\mathbf{\Phi}_{gY} \mathbf{\Phi}_{YY}^{-1}\right) \mathbf{Y} \tag{4.5}$$

where $\mathbf{g}$ is the short hand notation for $\mathbf{g}[0]$, $\tilde{\mathbf{g}}$ the corresponding estimate, $\mathbf{\Phi}_{gY}$ is the correlation of $\mathbf{g}$ and $\mathbf{Y}$, and $\mathbf{\Phi}_{YY}$ is the covariance matrix of $\mathbf{Y}$.

To determine $\mathbf{\Phi}_{gY}$ and $\mathbf{\Phi}_{YY}$, and consequently the optimal estimation filter

$$\mathbf{H} = \mathbf{\Phi}_{gY} \mathbf{\Phi}_{YY}^{-1}, \tag{4.6}$$

we can make use of the fact that the received vector $\mathbf{Y}$ can be written as

$$\mathbf{Y} = \mathbf{F}\mathbf{G} \tag{4.7}$$

where

$$\mathbf{G} = \begin{bmatrix} \mathbf{g}[-N-L] \\ \vdots \\ \mathbf{g}[-N] \\ \mathbf{g}[-N+1] \\ \vdots \\ \mathbf{g}[N+L] \end{bmatrix} \tag{4.8}$$

contains $2(N+L)+1$ $\mathbf{g}$ vectors and $\mathbf{F}$ is partitioned into $2N+1$ by $2(N+L)+1$ submatrices, with the (i,k)-th submatrix , $i = 1, 2, \ldots, 2N+1$ and $k = 1, 2, \ldots, 2(N+L)+1$, given by

$$\mathbf{f}_{i,k} = c(k-N-L-1)\,[p_0[i-k+L], p_1[i-k+L], \ldots, p_J[i-k+L]] \tag{4.9}$$

Note that by definition, $p_j[k] = 0$ for $|k| > L$. Consequently, the matrix $\mathbf{F}$ is uniquely determined if

$$\mathbf{P} = [c(-N-L), c(1-N-L), \ldots, c(N+L)] \tag{4.10}$$

are the known pilot symbols. Given the above structure for the received vector $\mathbf{Y}$, it is straightforward to determine the matrices $\mathbf{\Phi}_{gY}$ and $\mathbf{\Phi}_{YY}$.

$$\begin{aligned} \mathbf{\Phi}_{gY} &= \frac{1}{2}\overline{(g\,\mathbf{Y}^\dagger)} \\ &= \frac{1}{2}\overline{[g\,(\mathbf{FG})^\dagger)]} \\ &= \mathbf{\Phi}_{gG}\mathbf{F}^\dagger \end{aligned} \tag{4.11}$$

$$\begin{aligned} \mathbf{\Phi}_{YY} &= \frac{1}{2}\overline{(\mathbf{Y}\,\mathbf{Y}^\dagger)} \\ &= \frac{1}{2}\overline{[(\mathbf{FG})\,(\mathbf{FG})^\dagger)]} \\ &= \mathbf{F}\mathbf{\Phi}_{GG}\mathbf{F}^\dagger \end{aligned} \tag{4.12}$$

where $\dagger$ represents Hermitian transpose. $\mathbf{\Phi}_{gG}$ is the correlation matrix of $\mathbf{g}[k]$ and $\mathbf{G}$. $\mathbf{\Phi}_{GG}$ is the correlation matrix of $\mathbf{G}$ and $\mathbf{G}$. All work left in obtaining these matrices is the correlation element of $g_i[k_1]g_j[k_2]$ at times k_1 and k_2. This correlation expression is given in (3.12).

The mean squared estimation error for the estimator in (4.5) equals

$$\begin{aligned}\epsilon_1^2 &= \frac{1}{2}\overline{|\mathbf{g}-\tilde{\mathbf{g}}|^2} \\ &= \frac{1}{2}\overline{\left[\mathbf{g}-\left(\mathbf{\Phi}_{gY}\mathbf{\Phi}_{YY}^{-1}\right)\mathbf{Y}\right]\left[\mathbf{g}-\left(\mathbf{\Phi}_{gY}\mathbf{\Phi}_{YY}^{-1}\right)\mathbf{Y}\right]^{\dagger}} \\ &= trace\left(\mathbf{\Phi}_{gg}-\mathbf{\Phi}_{gY}\mathbf{\Phi}_{YY}^{-1}\mathbf{\Phi}_{Yg}\right), \end{aligned} \tag{4.13}$$

where $\mathbf{\Phi}_{gg}$ is the covariance matrix of $\mathbf{g}[k]$, and $\mathbf{\Phi}_{Yg}$ is the Hermitian transpose of the correlation matrix $\mathbf{\Phi}_{gY}$. Note that the mean squared error is a function of the pilot symbol sequence $\mathbf{P}$ in (4.10). In this study, we use a computer search to select the optimal pilot symbol sequence that minimizes the normalized mean squared error :

$$NMSE = \frac{\epsilon^2}{trace\left(\mathbf{\Phi}_{gg}\right)} \tag{4.14}$$

For BPSK with $N = 1$, $L = 1$, $fdT = 0.003$, $SNR = 30\,db$, the optimal pilot symbol sequence is $(1, -1, -1, -1, 1)$, which is different from the one reported in [20] with the same length.

4.1.2 Interpolation of the Derivative Processes

The periodic estimates $(\ldots, \tilde{\mathbf{g}}[-K], \tilde{\mathbf{g}}[0], \tilde{\mathbf{g}}[K], \ldots)$ obtained from the estimation filter in (4.5) are passed to an interpolator to generate the fading gain estimates at other instants. Without loss of generality, let's consider the estimation of $\mathbf{g}[k]$ for $k = 1, 2, \ldots, K-1$, where K is the frame size.

Assume the interpolator $\mathbf{A}$ has an order of $Q = 2I + 1$ and let

$$\tilde{\mathbf{G}} = \begin{bmatrix} \tilde{\mathbf{g}}[-IK + K] \\ \tilde{\mathbf{g}}[-IK + 2K] \\ \vdots \\ \tilde{\mathbf{g}}[IK] \end{bmatrix}. \tag{4.15}$$

be the collection of the Q relevant periodic estimates. Using a minimum mean squared error criterion, the optimal estimate of $\mathbf{g}[k]$, $k = 1, 2, \ldots, K-1$, is [37] [38]

$$\hat{\mathbf{g}}[k] = \mathbf{A}\tilde{\mathbf{G}} \tag{4.16}$$

where

$$\mathbf{A} = \mathbf{\Phi}_{g\tilde{G}}\mathbf{\Phi}_{\tilde{G}\tilde{G}}^{-1}, \tag{4.17}$$

with $\mathbf{\Phi}_{g\tilde{G}}$ being the correlation of $\mathbf{g}[k]$ and $\tilde{\mathbf{G}}$, and $\mathbf{\Phi}_{\tilde{G}\tilde{G}}$ the covariance of $\tilde{\mathbf{G}}$. It is understood that the optimal interpolator $\mathbf{A}$ depends on the time index k. As in the last subsection, the determination of $\mathbf{\Phi}_{g\tilde{G}}$ and $\mathbf{\Phi}_{\tilde{G}\tilde{G}}$, and consequently the optimal interpolator $\mathbf{A}$ is quite straight forward, especially when the signal structure reported in Section 3.1.1 is being exploited. The correlation submatrix involved in the computation is

$$\begin{aligned} \mathbf{\Phi}_{g[k_1]\tilde{g}[k_2]} &= \frac{1}{2}\overline{g[k_1]\tilde{g}^{\dagger}[k_2]} \\ &= \frac{1}{2}\overline{g[k_1]\left[\left(\mathbf{\Phi}_{gY}\mathbf{\Phi}_{YY}^{-1}\right)\mathbf{Y}\right]} \\ &= \mathbf{\Phi}_{gY}(\mathbf{\Phi}_{YY}^{-1})^{\dagger}\mathbf{\Phi}_{Yg} \end{aligned} \tag{4.18}$$

which can be evaluated as in the last subsection.

The mean squared error of the interpolator is

$$\begin{aligned} \epsilon_2^2 &= \frac{1}{2}\overline{|\mathbf{g} - \hat{\mathbf{g}}|^2} \\ &= \frac{1}{2}\overline{\left[\mathbf{g} - \left(\mathbf{\Phi}_{g\tilde{G}}\mathbf{\Phi}_{\tilde{G}\tilde{G}}^{-1}\right)\tilde{\mathbf{G}}\right]\left[\mathbf{g} - \left(\mathbf{\Phi}_{g\tilde{G}}\mathbf{\Phi}_{\tilde{G}\tilde{G}}^{-1}\right)\tilde{\mathbf{G}}\right]^{\dagger}} \\ &= \textit{trace}\left(\mathbf{\Phi}_{gg} - \mathbf{\Phi}_{g\tilde{G}}\mathbf{\Phi}_{\tilde{G}\tilde{G}}^{-1}\mathbf{\Phi}_{\tilde{G}g}\right), \end{aligned} \tag{4.19}$$

Once the estimates of the derivative processes at a given time instant are obtained, they will be converted into a channel impulse response estimate according to (4.1). The estimated channel response is denoted by $\hat{\mathbf{h}}[n] = \left(\hat{h}_{-L}[n], \hat{h}_{-L+1}[n], \ldots, \hat{h}_{L}[n]\right)$ with the ℓ-th coefficient given by

$$\hat{h}_{\ell}[n] = \sum_{j=0}^{J} \hat{g}_j[n]\, p_j[\ell] \tag{4.20}$$

4.1.3 Numerical Results

The performance of the channel estimator presented in subsections 3.1.1 and 3.1.2 is evaluated according the normalized mean squared error (NMSE) between $\mathbf{g}[k]$ and $\hat{\mathbf{g}}[k]$ at the interpolator output. The definition of the NMSE was given earlier in (4.14) for the periodic estimator. A similar expression can be easily derived for the interpolator. Although the optimal interpolator varies with the symbol position within a data frame, we found that the NMSE is rather insensitive to the data symbol position. For the two examples in Figs 4.3 and 4.4, the NMSE being reported are those that would be encountered in the center of a frame. It should be pointed out that pulse shape $p(t)$ being considered in these figures is a raised cosine pulse with a roll-off factor of 0.5 and that the order of the reduced dimensionality channel model, J, is equal to 2.

Fig. 4.3 illustrates the effect of the interpolator order Q on the accuracy of the estimates. Here a uniform delay-power profile with a rms delay spread of 0.2 symbol and a normalized fade rate of $f_dT = 0.003$ is considered. The frame size K is set to 100 symbols and the channel and the periodic channel estimation filter have parameters $L = 1$ and $N = 1$ respectively . It should be pointed out that the signal-to-noise ratio, SNR, is defined as the area under the delay-power profile, divided by the power spectral density of the channel noise, N_0. It is observed from Fig. 4.3 that the 3 curves run almost in parallel. The difference between the $Q = 2$ and the $Q =$

10 curve is about 4 dB in channel SNR. Although using a 10-th order interpolator can provide a significant improvement in terms of the SNR, the price to pay is the additional computational complexity and the extra delay introduced by using a longer interpolator.

Fig 4.4 illustrates the effect of fade rate on the estimator accuracy. Once again, a uniform delay power profile is assumed. The interpolator order is fixed at $Q = 2$ while the fade rate varies. It is observed that up to a channel SNR of 35 db, the pilot-symbol assisted channel estimator can track the channel accurately even at a fade rate as high as 3 % the symbol rate. Note that in the context of the North American digital cellular standard IS-54, a $f_d T = 0.03$ corresponds to a vehicle speed of approximately 600 km/hr !

The NMSE curves in Figs 4.3 and 4.4 decay inversely with the SNR. However, the corresponding bit error rate in the Viterbi equalizer will decrease at a much faster rate, due to the diversity effect provided by the multipath channel. In the next section, we provide an analysis of the bit error rate of the Viterbi equalizer.

4.2 Bit Error Probability of the Viterbi Equalizer

The sequence of received samples $\{y[n]\}$ and the corresponding channel estimates $\left\{\hat{\mathbf{h}}[n]\right\}$ are the input to a Viterbi receiver. The Viterbi receiver selects the sequence $\hat{\mathbf{c}} = (\hat{c}(1), \cdots, \hat{c}(K))$ that minimizes the metric

$$J(\hat{\mathbf{c}}) = \sum_{n=1}^{K} \left| y[n] - \sum_{l=-L}^{L} \hat{c}(n-\ell)\hat{h}_\ell[n] \right|^2 \tag{4.21}$$

where K is the frame size. Note that both the initial and the terminating states in the Viterbi receiver are determined by the pilot symbols in (4.10).

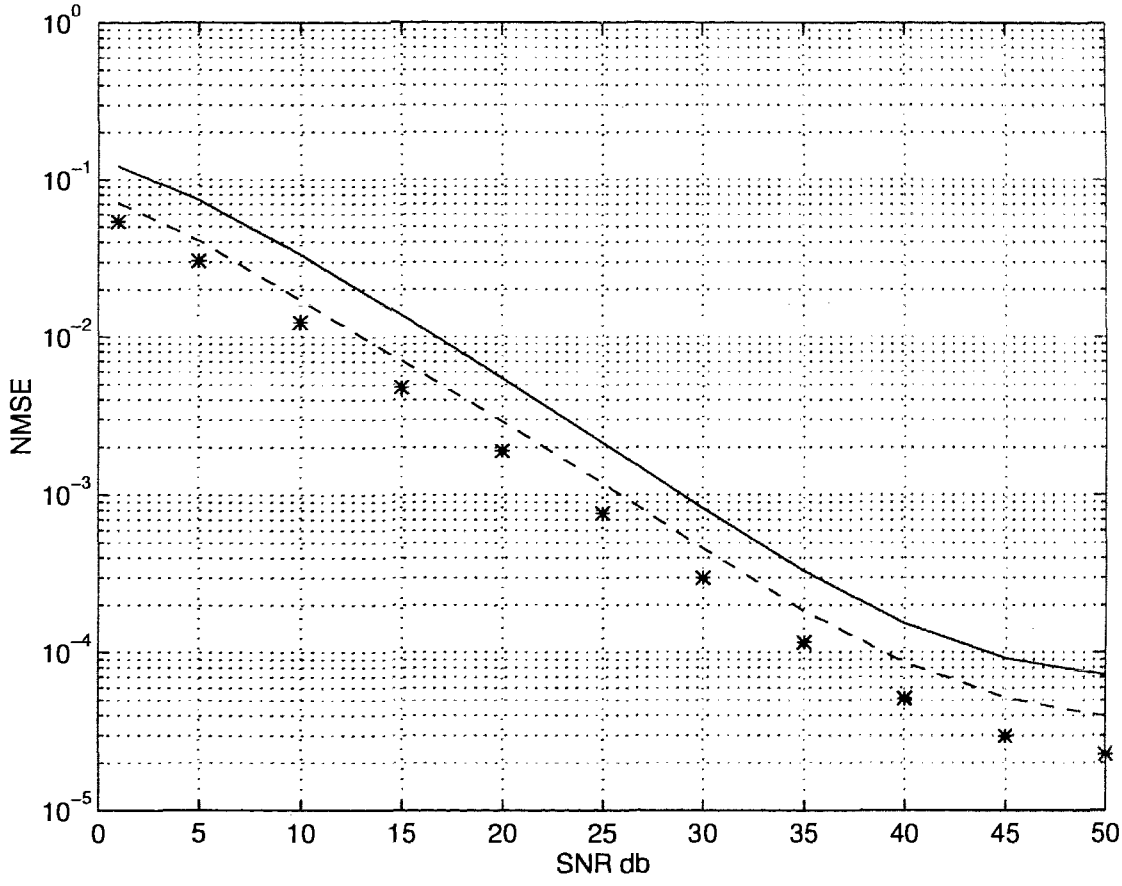

Figure 4.3: The normalized mean squared error in the reduced complexity channel estimator. The fade rate is $f_dT = 0.003$ and a uniform power delay profile with $\tau_{rms} = 0.2$ is used. The different curves correspond to different interpolator orders. Solid line: Q=2, dashed line: Q=4, star: Q=10.

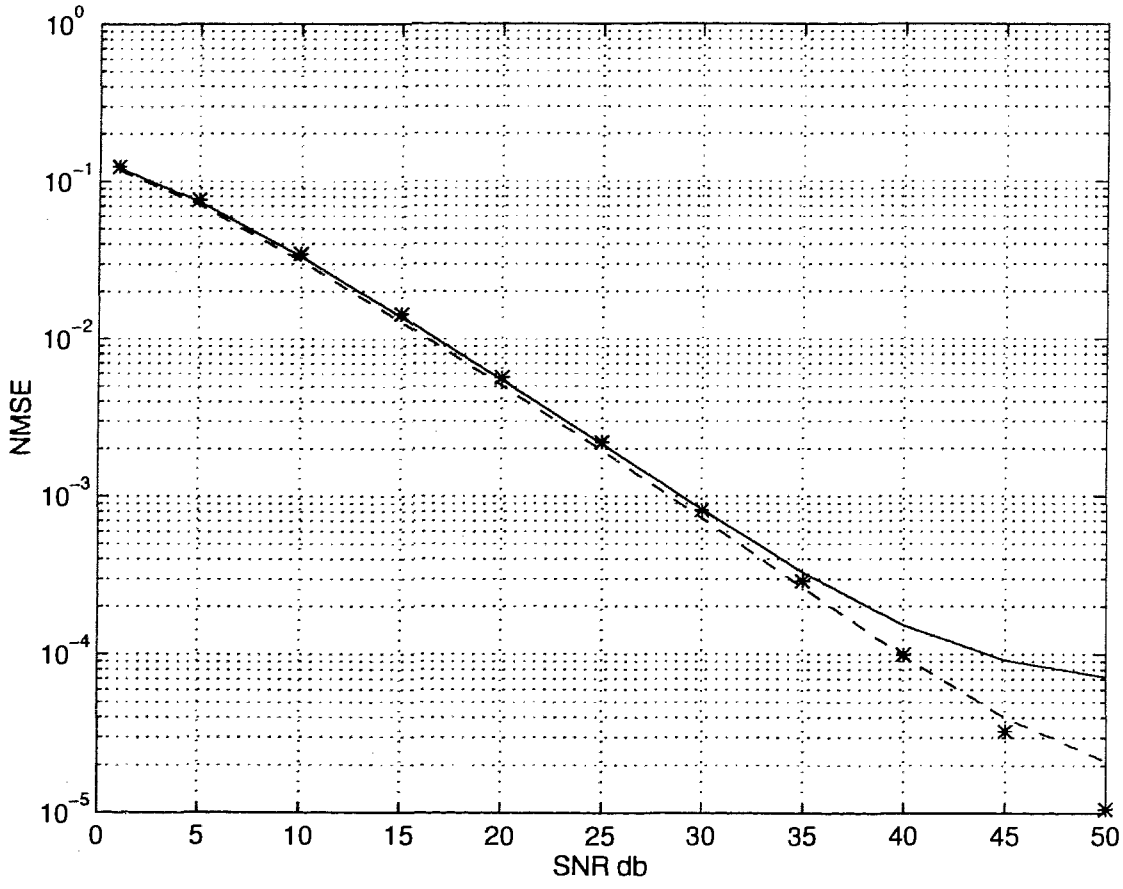

Figure 4.4: The normalized mean squared error of the channel estimator as a function of the fade rate. A uniform delay power profile with $\tau_{rms} = 0.2$ and an interpolator with $Q = 2$ was used. Solid line : $f_dT = 0.03$, dashed line: $f_dT = 0.003$, star: $f_dT = 0.0003$. The frame sizes are respectively 15, 100, and 1000 symbols.

Given the transmitted sequence $\mathbf{c} = (c(1), \cdots, c(K))$, a decoding error occurs if for some erroneous sequence $\hat{\mathbf{c}}$, the random variable D, defined as

$$D = J(\hat{\mathbf{C}}) - J(\mathbf{C}), \tag{4.22}$$

is less than zero. Let $P(\mathbf{c} \rightarrow \hat{\mathbf{c}})$ denotes the pairwise error event probability for a particular pair of $\mathbf{c}$ and $\hat{\mathbf{c}}$, i.e the chance that D is less than zero. In addition, let $a(\mathbf{c}, \hat{\mathbf{c}})$ be the Hamming distance between the two corresponding information sequences. Then the overall bit error probability of the Viterbi receiver can be approximated as [33] [34]

$$P_b = \frac{1}{Z} \sum_{\mathbf{c}} \frac{a(\mathbf{c}, \hat{\mathbf{c}})}{B} \sum_{\hat{\mathbf{c}} \neq \mathbf{c}} P(\mathbf{c} \rightarrow \hat{\mathbf{c}}) \tag{4.23}$$

where B is the number of the information bits per channel symbol, and Z is the number of possible transmitted sequences. Using the characteristic function approach suggested in [33], the pairwise error probability is given by

$$P(\mathbf{c} \rightarrow \hat{\mathbf{c}}) = p(D < 0) = -\sum \text{Residue} \left[\frac{\Phi_D(s)}{s} \right]_{RPpoles} \tag{4.24}$$

where $\Phi_D(s)$ is the characteristic function of the random variable D and the sum is over all the residues calculated at the right plane poles. The structure of the characteristic function can be found in equation (4B.1) of [35].

The calculation of the characteristic function $\Phi_D(s)$ is very similar to that in section 3.2.2. The detailed derivation is given in Appendix B.

4.2.1 Numerical Results

The bit error rate of BPSK, using a Viterbi equalizer in conjunction with a second order ($J = 2$) reduced dimensionality channel estimator, was determined analytically

and the results are shown in Figs 4.5 and 4.6. The fade rate in both figures is $f_dT = 0.003$ and the underlying raised cosine pulse shape, $p(t)$ has a 50% spectral roll-off. In addition, a uniform delay-power profile with a rms delay spread of 0.2 symbol is used. The two figures differ in the number of taps used in the channel impulse response estimates, and consequently in the number of states in the Viterbi receiver. Note that we use only the shortest and the next shortest error events in the error analysis.

It is observed from these figures that a good choice for the interpolator order is $Q = 8$. There is practically no gain in the bit error performance in going from $Q = 8$ to $Q = 10$. However, when Q is too small, say $Q = 4$, there exists an irreducible error floor in both figures.

In comparing the two figures, we can clearly see the effect of the length of the CIR, relative to number of symbols in the Viterbi algorithm state, on the bit error performance. For example, at a bit error rate of 10^{-5}, the $Q = 8$ receiver in Fig. 4.6 is about 3 dB better than the corresponding one in Fig. 4.5. In addition, the $Q = 8$ error curve in Fig. 4.6 is able to track the ideal performance curve, while the one in Fig. 4.5 begins to diverges from the ideal performance curve. Of course, the price to pay for this performance improvement is the computational complexity. Note that the receiver in Fig. 4.5 has 4 states, while the one in Fig. 4.6 has 16 states.

4.3 Conclusions

We present in this chapter a reduced complexity, pilot symbol assisted channel estimation technique for linear modulations operating in fading dispersive channels. The reduction in complexity is made possible by adopting a reduced dimensionality model

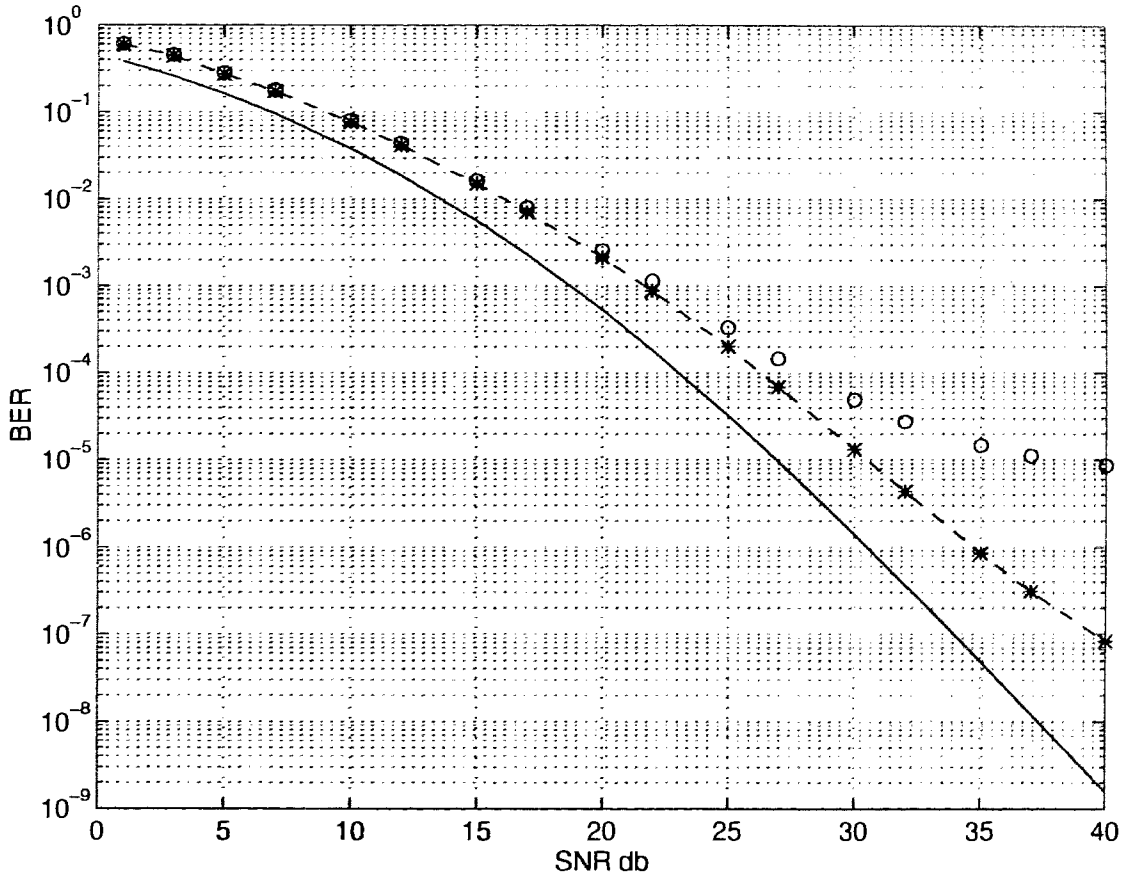

Figure 4.5: Bit error rate of a 4 states Viterbi equalizer for BPSK. A uniform delay power profile with $\tau_{rms} = 0.2$, and a fade rate of $f_d = 0.003$ was used. Solid line: perfect channel information, dashed line: Q=10, star: Q=8, circle: Q=4.

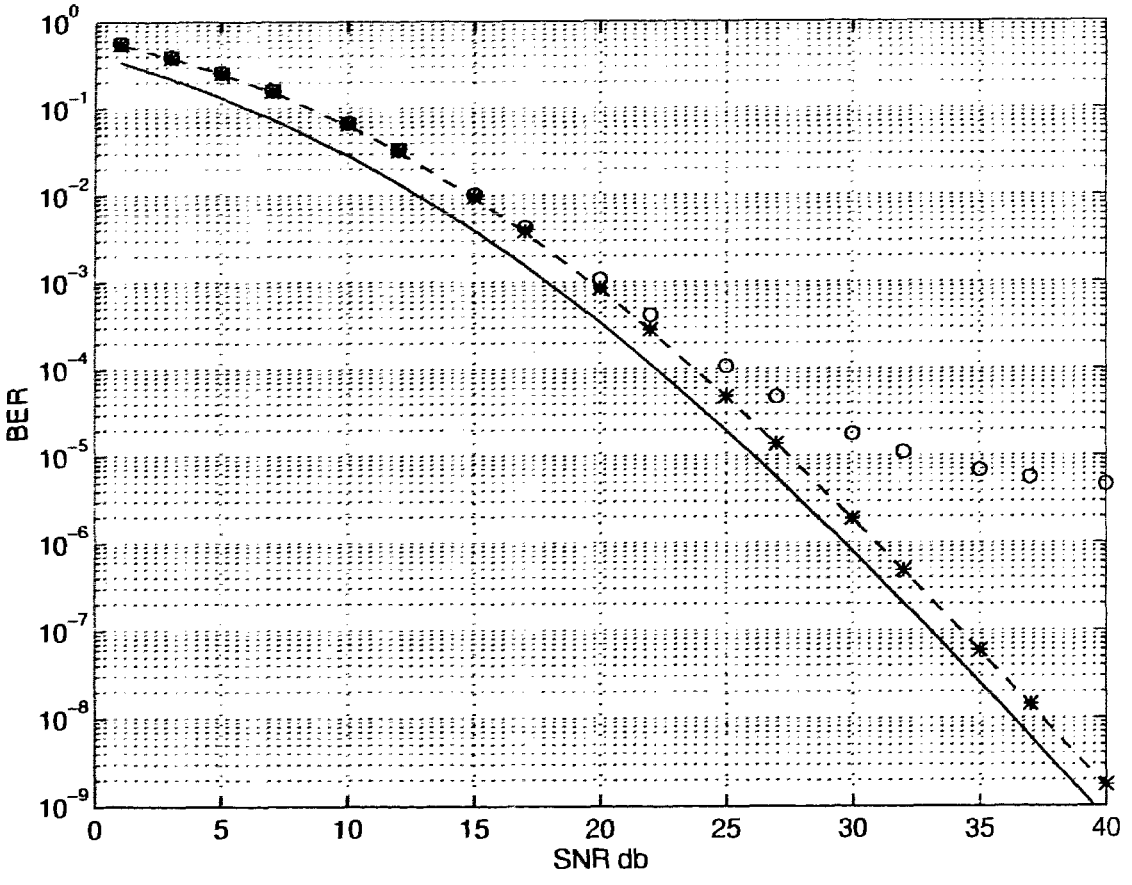

Figure 4.6: Same as the previous Figure, except that the Viterbi equalizer now has 16 states. Solid line : perfect channel information, dashed line: Q=10, star: Q=8, circle: Q=4.

for this type of channel. According to the model, instead of unnecessarily estimating the entire channel response, which may contain many taps, it is sufficient just to estimate 2 to 3 derivative channel processes.

We illustrate in the chapter how to perform channel estimation based on the reduced dimensionality channel model. The mean squared error of the estimator is presented, along with the bit error performance of the companion Viterbi equalizer. It was found that for a delay spread of up to 20% of a symbol, a Viterbi equalizer, in conjunction with a reduced complexity channel estimator, can provide good diversity effect for BPSK operating in a dispersive fading channel. The larger the number of states in Viterbi equalizer, the better the diversity effect will be achieved. It was also found that for all practical situations of interest, the performance of the Viterbi equalizer is relatively insensitive to the fade rate.

Chapter 5

A Semi-Blind Maximum Likelihood Sequence Estimation Receiver for Fading Dispersive Channels with Multiple Antennas

It is well known that the MLSE coupled with VA is an optimal receiver for the signal detection in the fading multipath channel [2] [11]. Due to the two kinds of channel impairments in wireless environment, i.e. the rapidly time varying and occasional null spectrum in the transmission band, the receiver in mobile communication should estimate jointly the channel response and the transmitted sequence in a short time [28]. In this chapter we proposed a 'semi-blind' MLSE-VA receiver that meets these requirements. The term 'semi-blind' comes from the fact that the receiver requires approximate knowledge of the fading spectrum.

The MLSE-VA receiver structures without forced pilot training sequence for explicit channel estimation have been studied for a long time since D.J. Forney's pioneer

paper [11]. Recent work on this topic centered on utilizing a bank of whitening filters, usually Kalman filters, for the implicit channel estimation purpose, and then deriving a corresponding decoding metric for the VA [25] [28] [29] .

Following the same track, we design a MLSE-VA receiver which differs from those of others, e.g. [25] [28] [29], in three aspects. Firstly, our channel model is different from the commonly used model, in that we considered in the model the intersymbol interference coming not only from the previous transmitted symbols but also from limited future symbols. This phenomenon appears when the transmission time is comparable with a symbol duration or the receiver has a delay of a couple of symbol durations in the processing of data, which is common in wireless communications. Utilizing this channel model, the Viterbi algorithm needs to incorporate future symbols in the state trellis diagram. Secondly, our channel estimator uses a Wiener filter while the Kalman filter is widely used in other papers. The Kalman filter is based on finite order model - an approximation. We use the Wiener filter, a FIR filter but truncated, which is a different approximation model. The rationale is that a Wiener filter is optimal in a complex Gaussian channel [37], and it provides the exact information we need while the Kalman filter provides extra information that is not used in the VA [25]. Thirdly, in order to reduce the computation complexity we make use in our branch metric calculation the survivor sequence in the state trellis of the VA. By so doing, we can reduce greatly the number of states in the trellis.

These three new features in our MLSE-VA, make our receiver accurate and efficient.

This chapter is organized as follows. After a brief review of the signal model (derived mainly in chapter 2) for the discrete multipath fading channels in Section 5.1, we describe our semi-blind MLSE-VA receiver in Section 5.2. The branch metric in the VA state transition is derived and the algorithm using diversity antennas is investigated. In Section 5.3, we provide some simulaiton results and the chapter is

concluded in Section 5.4.

5.1 Signal Model

This section gives a review of the communications signal model over multipath fading channels. Since our MLSE-VA is demonsttrated numerically for discrete channels, we will concentrate our detailed derivation for the discrete case only. By following through the signal model in chapter 2, it is easy to derive the specific MLSE-VA formulations for the continuous channels.

The block diagram of our system is shown in Fig. 2.1. The discrete channel model is given in (2.15). Rewrite the received sampled signal in (2.21) here

$$y[n] = \sum_k \sum_{i=1}^{M} b(k)\beta_i[n]p[(n-k)T_s - \tau_i] + e[n] \tag{5.1}$$

where $\{b(k)\}$ is the equivalent information sequence defined in (2.12). β_i is the random gain in the i^{th} path, and τ_i is the corresponding delay. $p(t)$ is the composite pulse shape. $e[n]$'s are correlated white gaussian random samples with the correlation given in (2.9). Recall that $\frac{1}{T_s}$ is the receiver sampling rate, $\frac{1}{T}$ is the transmission data rate, and the number of samples per baud defined in (2.11) is N_s. Also recall that in obtaining (5.1) we assume the timing recovery circuit produces the mean delay of the received signal (which is assumed to be zero in previous chapters) so that the sample with index iN_s is nominally associated with information symbol $c(i)$.

When using the orignal information data sequence sequence $\{c(k)\}$, we can rewrite $y[n]$ as

$$y[n] = \sum_k \sum_{i=1}^{M} c(kN_sT_s)\beta_i[n]p[(n-kN_s)T_s - \tau_i] + e[n] \tag{5.2}$$

Practically, if we assume the transmission channel has a limited memory of $2L$ signal symbols (with duration T) centered around sampling instant nT_s, by letting $n-k=\ell$, and hence $k=n-l$, (5.1) can be written as

$$y[n]=\sum_{l=-N_sL}^{N_sL} b(-l)\{\sum_i \beta_i[n]p(lT_s-\tau_i)\}+e[n] \tag{5.3}$$

Note that in (5.3) the received signal sample is related to L previous and L future transmitted data symbols, besides the current data symbol associating with time instance nT_s. This is because of the multipath as well as the transmission and processing delay in the channel and receiver.

We can rewrite (5.3) more compactly as

$$y[n]=\sum_{l=-N_sL}^{N_sL} b(n-l)h_l[n]+e[n] \tag{5.4}$$

where $h_l[n]=\sum_i \beta_i[n]p(-lT_s-\tau_i)$.

(5.4) is the discrete signaling model which we will use throughout this chapter for the Rayleigh fading channel. $h_l[n]$ is our channel impulse response tapped delay line model. The relationship between the data sequence $\{c(k)\}$ and received sampled sequence $\{y[n]\}$ can be modelled as a finite state machine shown in Fig. 5.1.

5.2 Maximum Likelihood Sequence Estimation and Viterbi Algorithm

The received signal sample $y[n]$ in (5.4) is a sufficient statistics for data signal detection [11]. Since the channel $h_l[n]$ as well as additive noise $e[n]$ are all correlated

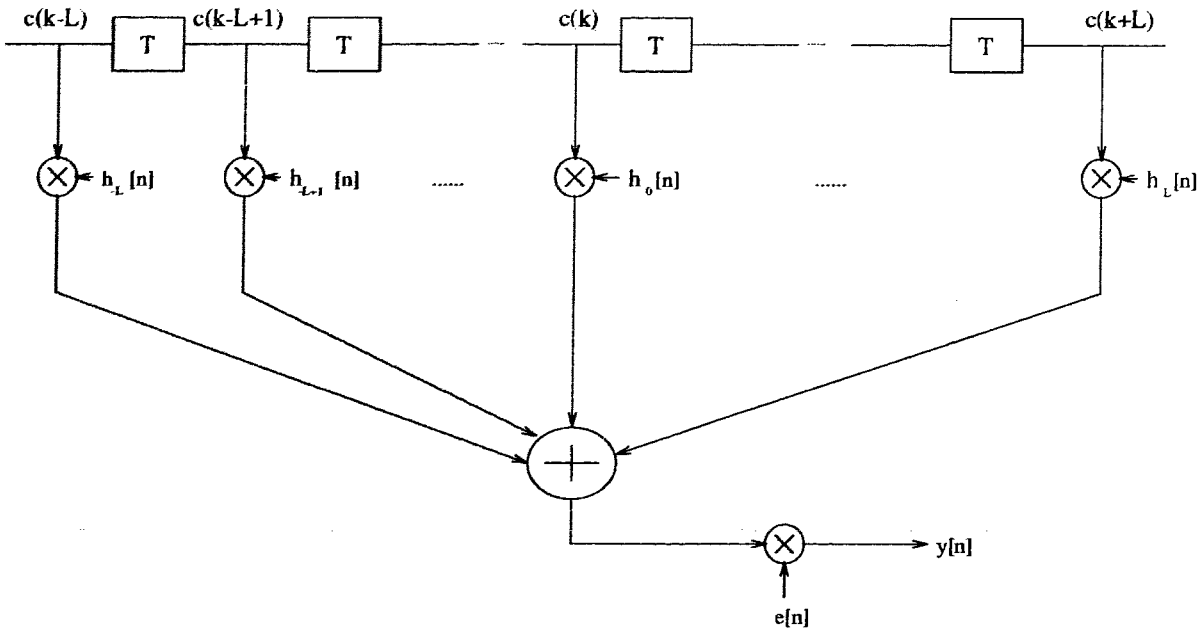

Figure 5.1: Finite state machine channel model.

complex Gaussian processes, the signal $y[n]$ is also correlated complex Gaussian random process. This process can be whitened by a whitening filter implemented as a Wiener filter. The whitened sufficient statistics are then used for the metric computation in MLSE-VA. In this section we derive the MLSE and VA that evaluates the derived metric recursively. Since the search computation complexity and the number of whitening filters in the VA grows exponentially with the number of trellis states [2], we design in section 5.2.2 a state reduction scheme for VA that makes use of the survival sequence to each trellis state. The total trellis states required to carry out VA is thus greatly reduced.

5.2.1 Maximum Likelihood Sequence Estimation with Wiener Whitening Filter

In the following, a boldface symbol represents vector. In our communications system, consider transmission of S q-ary data symbols $\mathbf{c}$, thus the equivalent information sequence $\mathbf{b}$ has a length N_sS and the received sampled sequence $\mathbf{y}$ has a length of N_sS. Denote the possible transmitted data sequences as

$$\mathbf{b}_1,\ \mathbf{b}_2,\ \cdots,\ \mathbf{b}_N$$

where $\mathbf{b}_i$, $i = 1, \cdots, N$ has a length N_sS and $N = q^S$. Let $\mathbf{y}_n = (y[1], y[2], \cdots, y[n])$ be the output sequence up to n'th symbol, $\mathbf{y}$ be the whole received information sequence of length S, and $\mathbf{y}_M[n] = (y[n - M + 1], y[n - M + 2], \cdots, y[n])$ be the length M sequence up to sample n.

The maximum likelihood sequence estimation chooses the decoded data sequence $\mathbf{b}_i$ if

$$p[\mathbf{y}|\mathbf{b}_i] > p[\mathbf{y}|\mathbf{b}_j],\ \text{where}\, j = 1, 2, \cdots, N, j \neq i. \tag{5.5}$$

It is clear that since $\mathbf{b}_i$ is equivalent to the information sequence $\mathbf{c}_i$ of length S, the sequence $\mathbf{b}_i$ is also a maximum likelihood estimate of $\mathbf{c}_i$.

By applying the conditional probability repeatedly as in [25] we have

$$p[\mathbf{y}|\mathbf{b}_i] = \prod_{n=1}^{N_sS} p[y[n]|\mathbf{y}_{n-1}, \mathbf{b}_i] \tag{5.6}$$

Taking the logarithm of both sides in (5.6) we obtain the log-likelihood function

$$log\, p[\mathbf{y}|\mathbf{b}_i] = \sum_{n=1}^{N_sS} log\, p[y[n]|\mathbf{y}_{n-1}, \mathbf{b}_i]. \tag{5.7}$$

Since each received sample $y[n]$ is complex Gaussian distributed, the conditional probability $p[y[n]|\mathbf{y}_{n-1}, \mathbf{b}_i]$ in (5.7) is also a complex Gaussian density function [38]. The mean and the variance of the conditional random process are respectively $\bar{y}_{n|n-1}(i)$ and $\sigma^2_{n|n-1}(i)$. Therefore we have the following simplification

$$\begin{aligned} log\, p[\mathbf{y}|\mathbf{b}_i] &= \sum_{n=1}^{N_sS} log\,\{\frac{1}{\sqrt{2\pi\sigma^2_{n|n-1}(i)}} exp[-\frac{|y[n]-\bar{y}_{n|n-1}(i)|^2}{2\sigma^2_{n|n-1}(i)}]\} \\ &= \sum_{n=1}^{N_sS}\{-\frac{|y[n]-\bar{y}_{n|n-1}(i)|^2}{2\sigma^2_{n|n-1}(i)} - \frac{1}{2} log[2\pi\sigma^2_{n|n-1}(i)]\} \end{aligned} \tag{5.8}$$

Now the decision rule in (5.5) becomes: choose a sequence $\mathbf{b}_i$ among $\mathbf{b}_1, \cdots, \mathbf{b}_N$ which has the minimum metric

$$\Lambda^i_S(\mathbf{y}) = \sum_{n=1}^{N_sS}\{\frac{|y[n]-\bar{y}_{n|n-1}(i)|^2}{\sigma^2_{n|n-1}(i)} + log[\sigma^2_{n|n-1}(i)]\},\ i = 1, 2, \cdots, q^S \tag{5.9}$$

In the metric calculation in (5.9), the conditional mean and variance need to be estimated. The estimation can be made through a Wiener prediction filter which is optimal for the Gaussian processes [37] [38]. Practically the prediction filter will have a limited order, say order M. Thus given the received length M sample sequence up to sample $n-1$, i.e. $\mathbf{y}_M[n-1]$, the Wiener filter will produce the sample estimate $\hat{y}_{n|n-1}$ with the variance $\hat{\sigma}^2_{n|n-1}$.

$$\hat{y}_{n|n-1} = \mathbf{\Phi}_{y[n]\mathbf{y}_M[n-1]}\, \mathbf{\Phi}^{-1}_{\mathbf{y}_M[n-1]\mathbf{y}_M[n-1]}\, \mathbf{y}_M[n-1] \tag{5.10}$$

and

$$\hat{\sigma}^2_{n|n-1} = \mathbf{\Phi}_{y[n]y[n]} - \mathbf{\Phi}_{y[n]\mathbf{y}_M[n-1]}\, \mathbf{\Phi}^{-1}_{\mathbf{y}_M[n-1]\mathbf{y}_M[n-1]}\, \mathbf{\Phi}_{\mathbf{y}_M[n-1]y[n]} \tag{5.11}$$

where function $\mathbf{\Phi}_{y[n]\mathbf{y}_M[n-1]}$ is the correlation matrix of $y[n]$ and vector $\mathbf{y}_M[n-1]$, $\mathbf{\Phi}_{\mathbf{y}_M[n-1]y[n]}$ is the Hermitian transpose of $\mathbf{\Phi}_{y[n]\mathbf{y}_M[n-1]}$. $\mathbf{\Phi}_{\mathbf{y}_M[n-1]\mathbf{y}_M[n-1]}$ is the covariance matrix of $\mathbf{y}_M[n-1]$, and $\mathbf{\Phi}_{y[n]y[n]}$ is the covariance of $y[n]$. These correlation

matrices can be computed without difficulty from the definitions as we did in the previous chapter. The $(n_1, n_2)^{th}$ correlation element involved in these matrices is obtained from (5.4) as

$$\begin{aligned}
\Phi_{y[n_1]y[n_2]} &= \frac{1}{2}E[y[n_1]y^*[n_2]] \\
&= \frac{1}{2}E\{[\sum_{l_1=-N_s(L+1)+1}^{N_sL} b_1 h_{l_1}[n_1]] + e[n_1][\sum_{l_2=-N_s(L+1)+1}^{N_sL} b_2 h_{l_2}[n_2] + e[n_2]] \\
&= \sum_{l_1=-N_s(L+1)+1}^{N_sL}\sum_{l_2=-N_s(L+1)+1}^{N_sL} b_1 b_2 \sum_i p(l_1T_s - \tau_i)p(l_2T_s - \tau_i)\phi_i[(n_1 - n_2)T_s] \\
&\quad + N_0 R_{qr}[(n_1 - n_2)T_s] \qquad (5.12)
\end{aligned}$$

where $\lceil\cdot\rceil$ is the ceiling function, $b_1 = b(-l_1 + \lceil\frac{n_1}{N_s}\rceil N_s)$, and $b_2 = b(-l_2 + \lceil\frac{n_2}{N_s}\rceil N_s)$.

Now we see the conditional mean and variance in (5.9), i.e. $\bar{y}_{n|n-1}(i)$ and $\bar{\sigma}_{n|n-1}(i)$, can be estimated by using this Wiener filter. Although (5.12) is a relatively complex expression, fortunately the receiver only has to compute it once to obtain a set of precalculated prediction filter coefficients. Practically a bank of filters will be used for all the hypothesed data sequence $\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_N$. The receiver structure using a bank of Wiener filters are given in the next subsection regarding the VA.

5.2.2 Viterbi Algorithm for the Sequence Detection

From above subsection we see the maximum likelihood sequence can be selected among other possible data sequences using a Wiener whitening filter to provide the estimate $\bar{y}_{n|n-1}$ and the associated estimation variance $\sigma^2_{n|n-1}$. Since the channel memory is finite, the relationship between the input data sequence $\mathbf{c}$ and the received sample sequence $\mathbf{y}$ can be described by a trellis. The VA can be used to carry out the sequential estimation by traversing through a state trellis and evaluating the metric $\Lambda^i_S(\mathbf{y})$

recursively in the state transition process.

To facilitate the use of the above metric in Viterbi algorithm to decode the data sequence $\mathbf{c}$, and by realizing the information sequence has length S and symbol rate $\frac{1}{T}$, the metric needs to be broken into a sum over the information symbols and a sum over the samples associated with each symbol. We have the metric to be minimized with respect to the data sequence as

$$\begin{aligned}\Lambda_S^i(\mathbf{y}) &= \sum_{n_c=1}^{S}\sum_{n_s=1}^{N_s}\{\frac{|y[N_s(n_c-1)+n_s]-\bar{y}_{N_s(n_c-1)+n_s}(i)|^2}{\sigma^2_{N_s(n_c-1)+n_s}(i)}+log\,\sigma^2_{N_s(n_c-1)+n_s}(i)\}\\ &= \sum_{n_c=1}^{S}\lambda_i(\mathbf{y}_M[n_c],\mathbf{c}(n_c)),\ i=1,2,\cdots,q^S \qquad (5.13)\end{aligned}$$

where we use a short-hand notation $\bar{y}_{N_s(n_c-1)+n_s}(i)$ and $\sigma^2_{N_s(n_c-1)+n_s}(i)$ to represent the previous notation $\bar{y}_{N_s(n_c-1)+n_s|N_s(n_c-1)+n_s-1}$ and $\sigma^2_{N_s(n_c-1)+n_s|N_s(n_c-1)+n_s-1}(i)$ respectively.

$\lambda_i(\mathbf{y}_M[n_c],\mathbf{c}(n_c))$ is the i^{th} branch metric defined as

$$\begin{aligned}&= \lambda_i(\mathbf{y}_M[n_c],\mathbf{c}(n_c))\\ &= \sum_{n_s=1}^{N_s}\{\frac{|y[N_s(n_c-1)+n_s]-\bar{y}_{N_s(n_c-1)+n_s}(i)|^2}{\sigma^2_{N_s(n_c-1)+n_s}(i)}+log\,\sigma^2_{N_s(n_c-1)+n_s}(i)\} \qquad (5.14)\end{aligned}$$

It should be pointed out that in the above branch metric $\lambda_i(\mathbf{y}_M[n_c],\mathbf{c}(n_c))$, we assumed that the number of the correlated output samples at time k is M, which corresponds to the order of the Wiener prediction filter discussed in the last subsection. In reality, M can be generally determined by the truncated length of the bandlimited transmission pulse shaping $q_t(t)$ and the multipath spread of the transmission channel [29].

There are in total $M+2L$ information symbols involved in the above branch metric calculation. As defined in 5.4, each received sample $y[n]$ is associated with L previous

and L future data symbols as well as the data symbol at time instant kT. Therefore the VA trellis in total needs q^{2L-1} states if we use the usual VA sequence estimation as in [2] [29] and [25] or the blind sequence detection as in [28]. The computation complexity can be excessive.

Realize that associated with each state in the trellis there is a survivor sequence, which is determined in the previous sequential detection iteration. We can make use of this survivor sequence in the calculation of the branch metric for all transitions emanating from that state. By doing so, we don't need to consider the 'previous' data symbols in the trellis state design. Therefore the state trellis can be defined with only L future symbols. In this way, we have the same detection accuracy and the number of states in the trellis can be reduced to q^L.

The correlation expression needed in the Wiener filter estimation of the conditional quantities in (5.13) can be obtained in terms of the information sequence as

$$\begin{aligned} & \mathbf{\Phi}(n_1T_s, n_2T_s) \\ = \; & \frac{1}{2}y(n_1T_s)y^*(n_2T_s) \\ = \; & \sum_{l_1=-L}^{L}\sum_{l_2=-L}^{L} c_1c_2 \sum_m p[-l_1N_sT_s - \tau_i]p[-l_2N_sT_s - \tau_i] \\ & \phi_i[(n_1 - n_2)T_s] + N_0R_{qr}[(n_1 - n_2)T_s] \end{aligned} \tag{5.15}$$

where $c_1 = c[(-l_1N_s + n_1)T_s]$ and $c_2 = c[(-l_2N_s + n_2)T_s]$.

As an example, state trellis in our VA with $L = 2$ is the same as in Fig. 3.6. The MLSE-VA receiver structure is shown in Fig. 5.2.

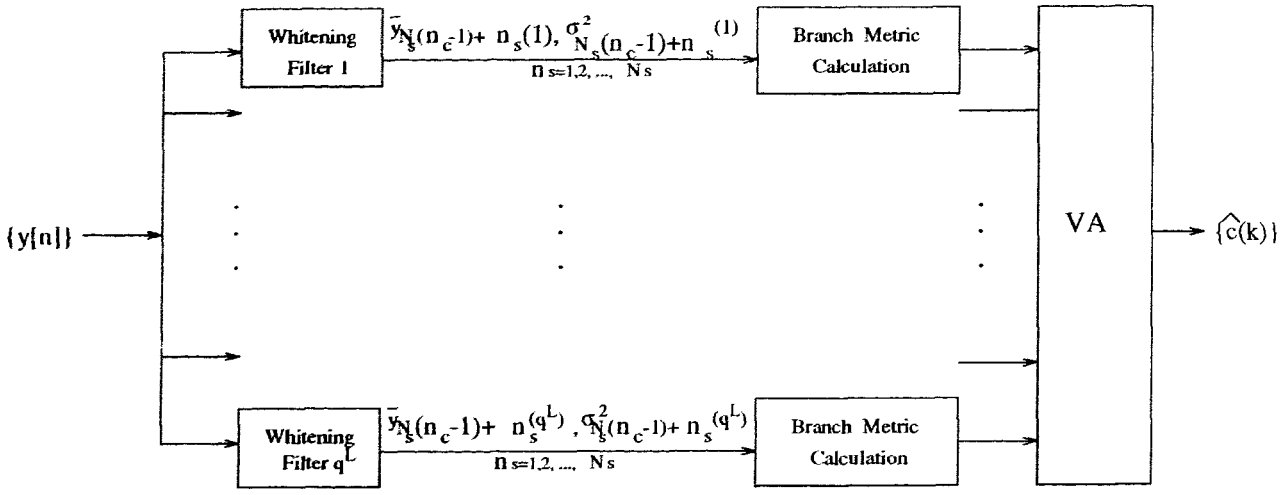

Figure 5.2: Diagram of the MLSE-VA receiver with a bank of whitening filters.

5.2.3 MLSE-VA Metric for Receiver with Diversity Channels

To this point, the development has been concerned with single channels. If measurements from D statistically independent channels are available, then performance can be improved significantly. Extending (5.6), we note that the maximum likelihood receiver maximizes

$$p[\mathbf{y}^{<1>}, \cdots, \mathbf{y}^{<D>} | \mathbf{b}_i] = \prod_{j=1}^{D} p[\mathbf{y}^{<j>} | \mathbf{b}_i] \tag{5.16}$$

where the superscripts identify the channel and the product follows from the assumed independence of the channels. The log likelihood in (5.7) and the rest of the development are similar to the single channel case and the symbol branch metric (5.13) is rewritten as the sum of symbol branch metric calculated for each channel, given by

$$\lambda_i^D(\mathbf{y}_M^{<1>}[n_c], \cdots, \mathbf{y}_M^{<D>}[n_c], \mathbf{c}(n_c)) = \sum_{j=1}^{D} \lambda_i(\mathbf{y}_M^{<j>}[n_c], \mathbf{c}(n_c)) \; i = 1, 2, \cdots, q^S \tag{5.17}$$

where

$$\lambda_i(\mathbf{y}_M^{<j>}[n_c], \mathbf{c}(n_c)) = \sum_{n_s=1}^{N_s} \{ \frac{|y^{<j>}[N_s(n_c-1)+n_s] - \bar{y}^{<j>}_{N_s(n_c-1)+n_s}(i)|^2}{\sigma^{<j>2}_{N_s(n_c-1)+n_s}(i)} + log\, \sigma^{<j>2}_{N_s(n_c-1)+n_s}(i) \} \tag{5.18}$$

This is a simple replication of the single channel metric calculation.

A diversity communication system for Rayleigh fading channels with ISI is shown in Fig. 5.3.

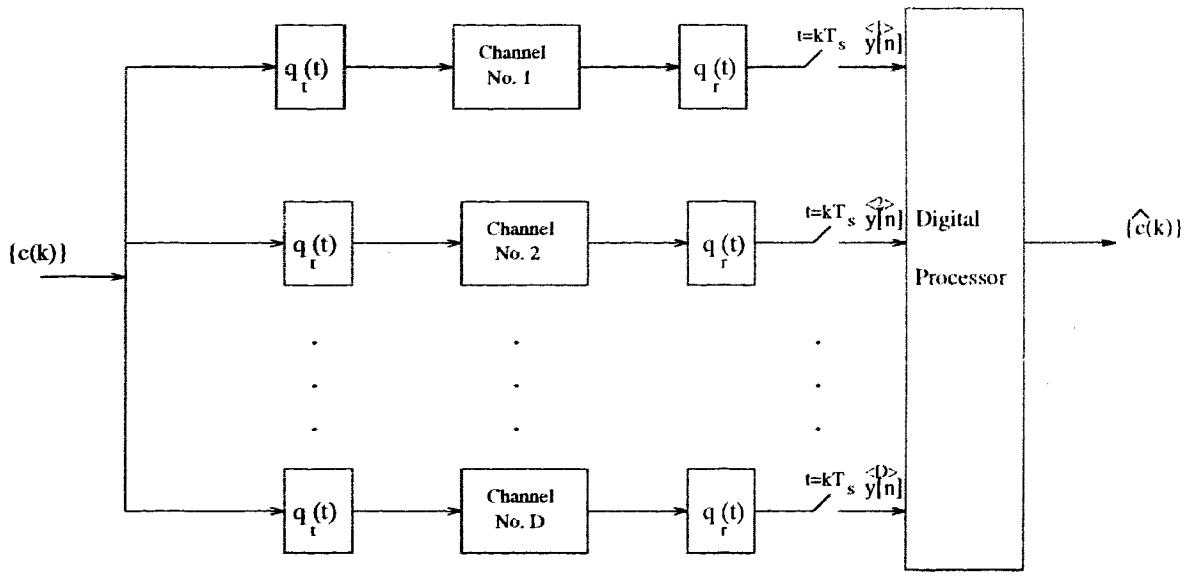

Figure 5.3: A communication system with diversity antennas.

5.3 Simulation Results

To evaluate the error performance of MLSE-VA we presented in the previous sections, computer simulations are done for two-ray dispersive channel.

In these examples, the transmission pulse shape $q_t(t)$ and the receiver filter $q_r(t)$ both are the squared raised cosine function with a 50% spectral rool-off. The transmitted data symbol rate is $\frac{1}{T} = 1$, while the receiver sampling rates are $\frac{1}{T_s} = 1$ and $\frac{1}{T_s} = 2$ respectively. The channel memory length considered is 5 data symbols, i.e. $L = 2$. Using our state reduction strategy, the VA contains 4 states. The channel estimation filter discussed in section 5.2 has a highest order of $M = 5$.

The bit error rates of BPSK in slow and fast fading channels are obtained in Fig. 5.5 and Fig. 5.4. The fading rates considered are $f_dT = 0.03$ in Fig. 5.4, and $f_dT = 0.003$ in Fig. 5.5. In both figures, the channel rms delay is $\tau_{rms} = 0.2$, and power split ratio in the two rays is 1.

Both figures show that the double sampling VA with $\frac{1}{T_s} = 2$ tracks the fading channel much better than the single sample scheme, although the BER is still high compared with the VA with perfect channel impulse respinse (CIR). It is observed that bit error rate curve corresponding to the VA with the sampling rate of $\frac{1}{T_s} = 1$ departs largely from the curve for VA with perfect CIR when the SNR becomes high. For high fade rate of $f_dT = 0.03$, the BER of the double sampling VA receiver is roughly 1.4db worse than the one with the perfect CIR VA. There is an error floor of the order 10^{-2} in one sample per baud case and an error floor in the order of 10^{-3} in the two samples per baud case. In the slow fading channel with $f_dT = 0.003$, the VA performance is a bit closer to that of the perfect CIR VA, when comparing with the fast fading case. No obvious error floor is observed within the scope of the SNR under study.

It needs to be pointed out that according to [29], an irreducible error rate exists in their type of MLSE-VA receiver.

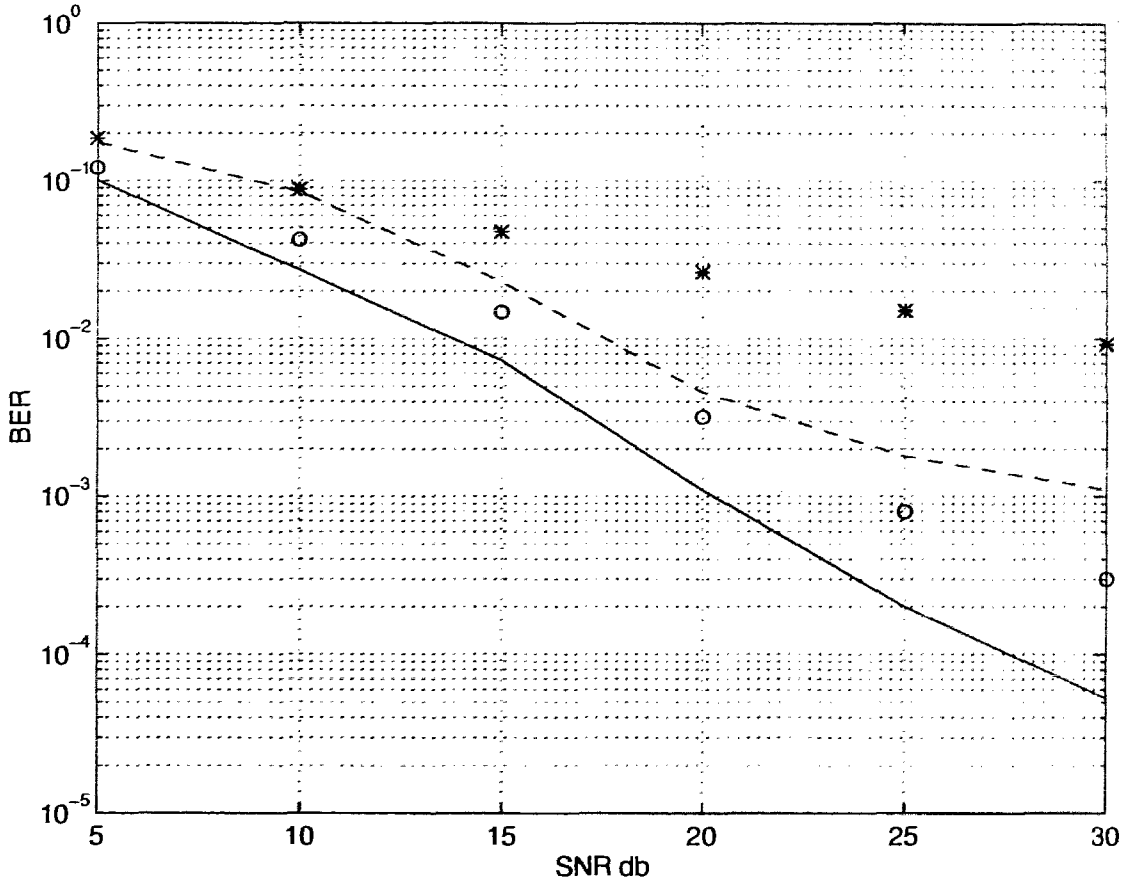

Figure 5.4: Bit error rate of MLSE-VA for BPSK. A two-ray delay power profile with $\tau_{rms} = 0.2$, and a fade rate of $f_d T = 0.03$. Solid line: perfect channel information with $T_s = 0.5$, dashed line: $T_s = 0.5$, star: $T_s = 1$, circle: perfect channel information with $T_s = 1$

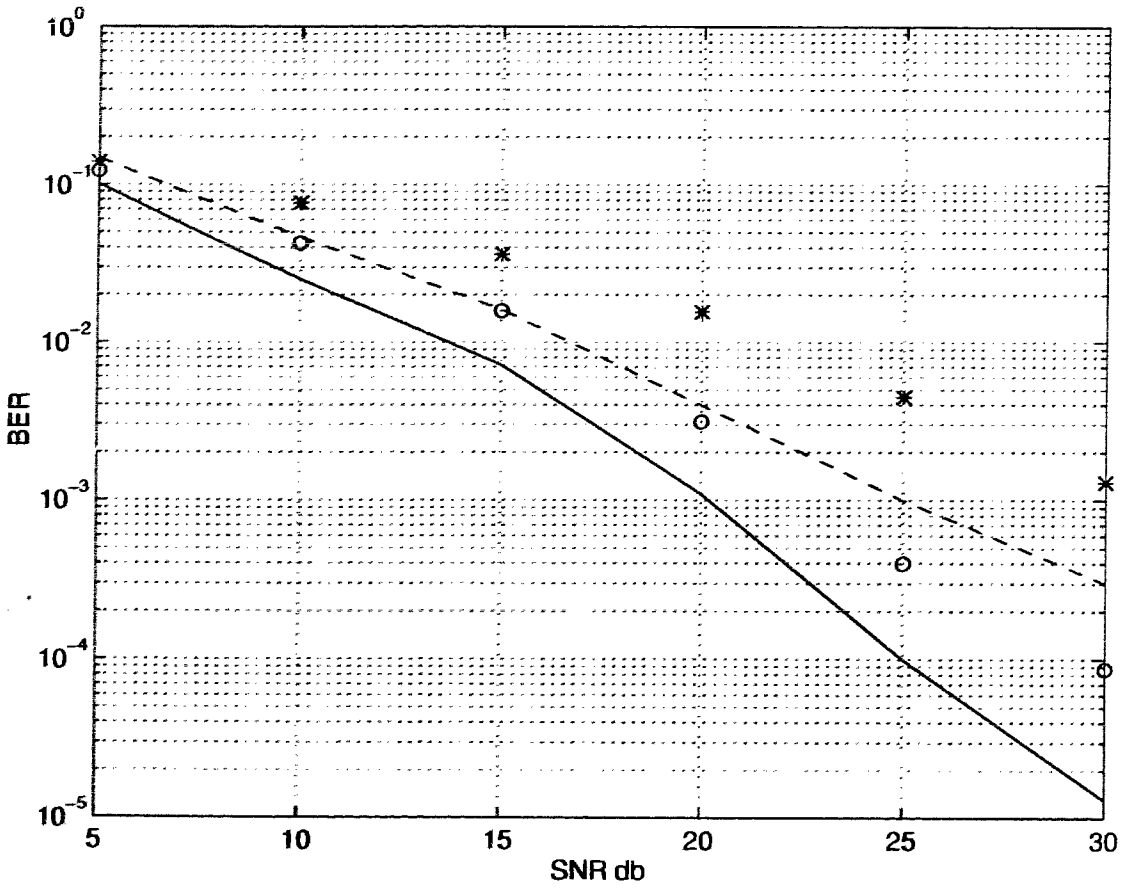

Figure 5.5: Same as the previous figure except that $f_dT = 0.003$. Solid line : perfect channel information with $T_s = 0.5$, dashed line: $T_s = 0.5$, star: $T_s = 1$. circle: perfect channel information with $T_s = 1$

In Fig. 5.6 and Fig. 5.7, we show the bit error rate versus the normalized rms delay spread (wrt symbol rate) for $SNR = 30$db and 20db respectively. The fading rate considered is $f_dT = 0.03$.

We see that the performance of the 2 samples per symbol VA is much better than that of the one sample per baud VA. Note the diversity effect in the VA when we have the perfect channel state information. For small delay spread, i.e. $\tau_{rms} < 0.3$, the 2 sample per baud VA receiver shows the diversity effect. For higher SNR, the BER performance of our VA receiver departs largely from that of the VA receiver having perfect knowledge of CSI. This result agrees with that shown in Fig. 5.4 and Fig. 5.5. The findings are the same for slow fading channel when $f_dT = 0.003$, which is not shown here.

We investigated the applications of multiple receive antennas in the MLSE-VA, which is developed in Section 5.2.3. The simulation results are shown in Fig. 5.8 and Fig. 5.9. In these examples, two independent antennas are used. The MLSE-VA performance is compared with the one using only one antenna.

Fig. 5.8 shows the BER vs. SNR for VA receiver using one and two antennas. It shows that the BER performance of VA receiver for $T_s = 1$ using two independent antennas is at least as good as that of the VA receiver with $T_s = 0.5$ using only one antenna. The BER performance of VA receiver with 2 sample per symbol sampling and 2 independent antennas is very close to and can be better (for small SNR) than that of the the 2 sample per symbol receiver with perfect CSI using only one receive antenna. We see from the figure that there is a BER performance improvement of around 10*db* when 2 independent receiver antennas are in use instead of only one antenna. In the graph, no irreducible error floor is observed for VA receiver with two

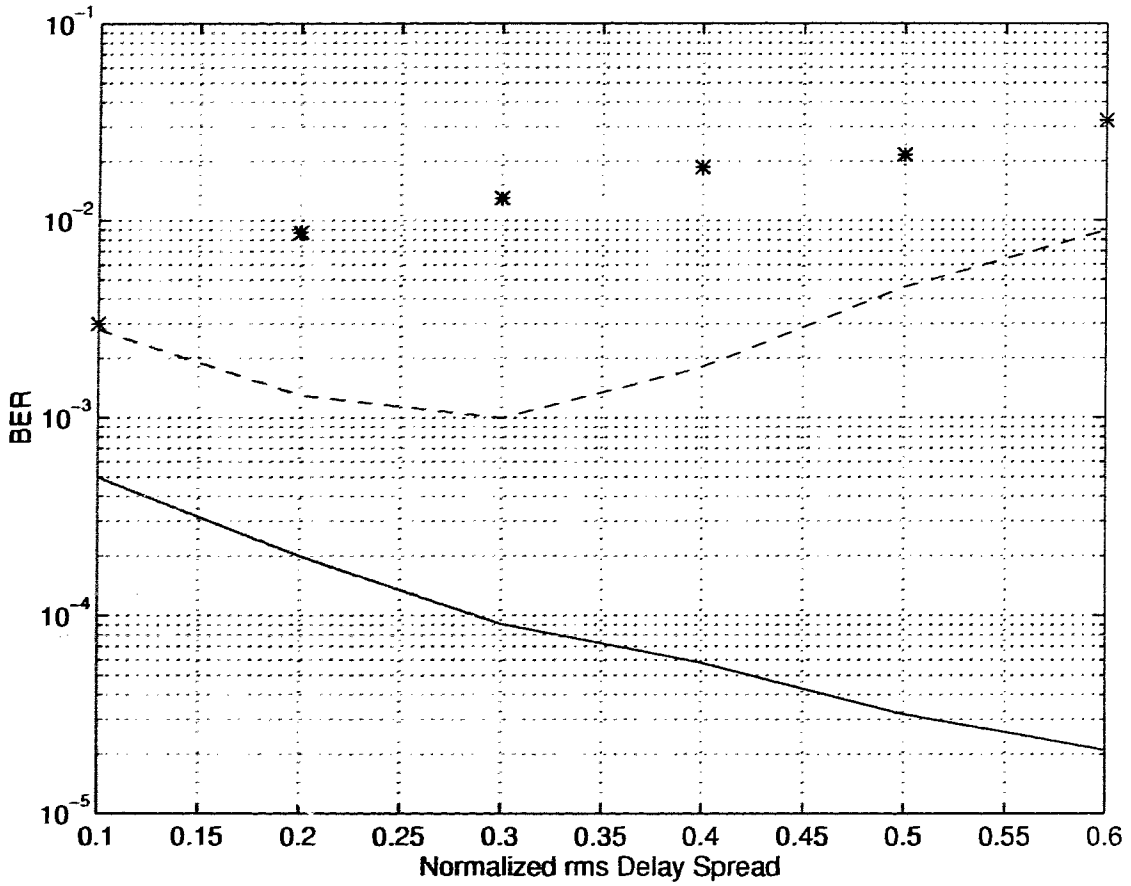

Figure 5.6: Bit error rate of MLSE-VA for BPSK. A two-ray channel with fading rate of $f_dT = 0.03$ and $SNR = 30$db. Solid line: perfect channel information with $T_s = 0.5$, dashed line: $T_s = 0.5$, star: $T_s = 1$

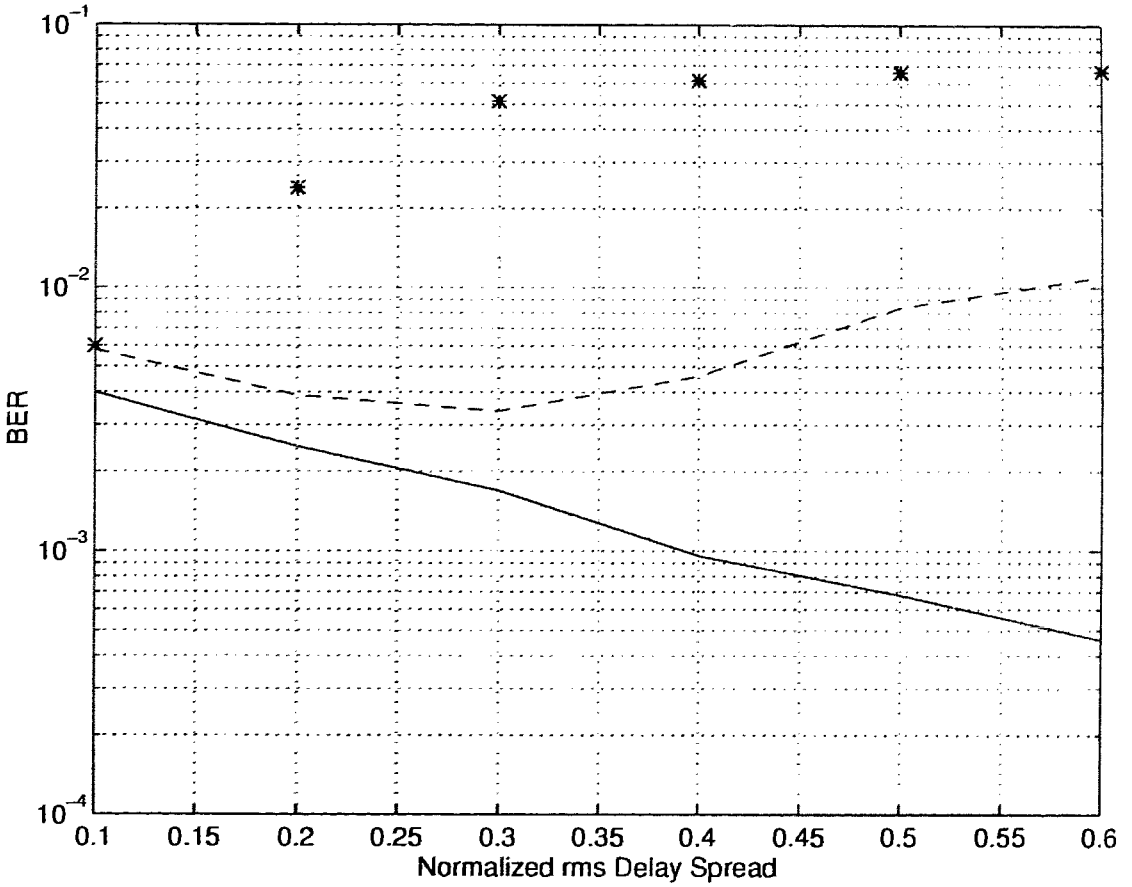

Figure 5.7: Same as the previous figure except that $SNR = 20$db. Solid line : perfect channel information with $T_s = 0.5$, dashed line: $T_s = 0.5$, star: $T_s = 1$

antennas.

Fig. 5.9 shows the BER versus normalized rms delay spread for the application of 2 independent antennas. Again we see similar findings shown in Fig. 5.8 when comparing the performance of VA receivers using one and two antennas. Similar to the result shown in Fig. 5.8, there is a large BER performance improvement with the applicaiton of two independent receiver antennas. The performance of the two-antenna VA receiver is very close to that of the receiver with perfect CSI using one receiver antenna. Note the range of the diversity effect of the VA receiver using 2 antennas is now extended to $\tau_{rms} < 0.5$, as compared with the performance of VA receiver using only one antenna which is shown in Fig. 5.6.

In our simulations, we also studied the situation when we don't know the whole channel delay power profile but only some statistics of it, as is usually the case in practice. Though not shown here, we find that, for small delay spread, the MLSE-VA receiver designed with only the knowledge of the first and second moments of the delay power profile has very close BER performance to the one designed with the whole channel delay power profile. This agrees with the results found in Chapter 3 and 4.

5.4 Conclusions

In this chapter, we present a novel semi-blind MLSE-VA receiver for fading dispersive channels that requires no training sequences, performs no CIR tracking and makes use of one or more antenna signals, as available. Both the fading channel response and the transmitted data sequence are estimated jointly by utilizing a Wiener prediction filter incorporated in the VA. A state reduction scheme which uses the survival sequence

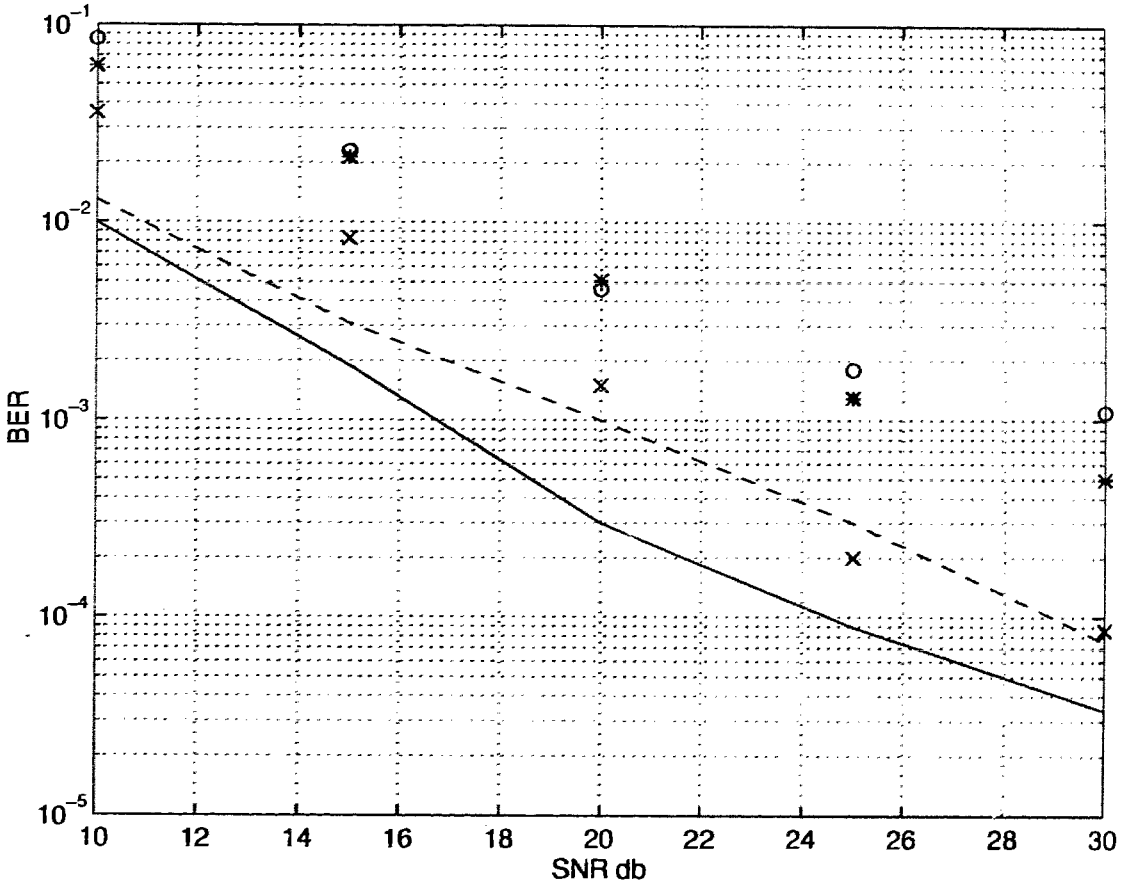

Figure 5.8: Bit error rate of MLSE-VA for BPSK. A two-ray delay power profile with $\tau_{rms} = 0.2$, and a fade rate of $f_d T = 0.03$. Solid line: perfect channel information with $T_s = 0.5$ using 2 antennas, dashed line: $T_s = 0.5$, 2 antennas cross: perfect channel information with $T_s = 0.5$ using 1 antenna star: $T_s = 1$, 2 antennas circle: $T_s = 0.5$, 1 antenna.

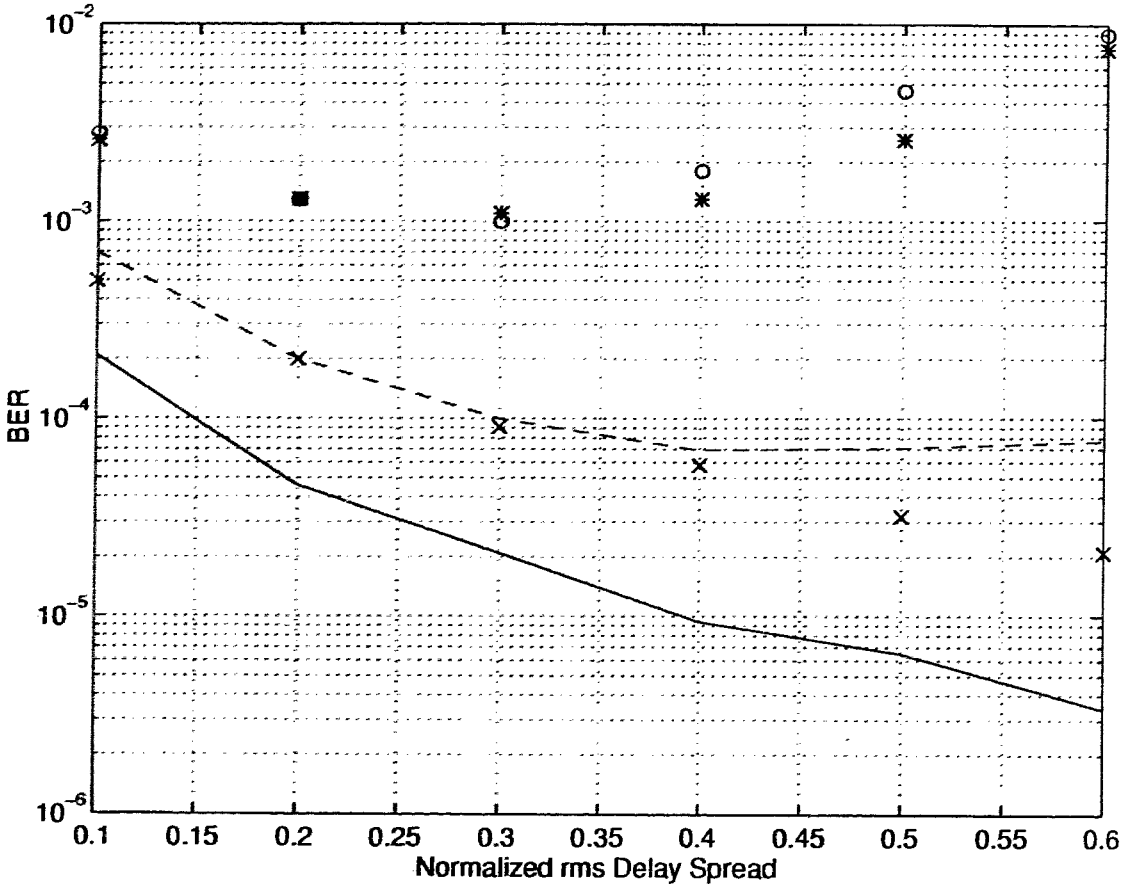

Figure 5.9: Bit error rate of MLSE-VA for BPSK. A two-ray channel with fading rate of $f_dT = 0.03$ and $SNR = 30$db. Solid line: perfect channel information with $T_s = 0.5$ using 2 antennas, dashed line: $T_s = 0.5$, 2 antennas cross: perfect channel information with $T_s = 0.5$ using 1 antenna star: $T_s = 1$, 2 antennas circle: $T_s = 0.5$, 1 antenna.

to each state in VA is proposed to reduce greatly the number of states in the VA. The multiple sampling rate scheme is studied and the result is compared with MLSE-VA using data symbol rate at the receiver filter output. The MLSE-VA is applied to a multiple antenna environment, where it enjoys both performance enhancement through diversity and reduction of complexity.

In addition to derivations of the receiver structure and metric, we present simulation results for performance, including the effect of Doppler and delay spread, state reduction methods, effect of multiple sampling and multiple antennas. It is illustrated through the simulation that the double sampling scheme at the receiver output results in a simple MLSE-VA whose performance is reasonably close to the MLSE-VA with perfect CIR, especially for the slow fading channels. The antenna diversity improves the performance. It is demonstrated through the chapter that the proposed MLSE-VA is simple in structure with a fair accuracy.

Chapter 6

Conclusions

Communication systems such as mobile radio and indoor wireless communication typically are modeled as Rayleigh fading channels. As data rates for these channels are increased, intersymbol interference (ISI) may result from both the multipath spread and the channels having a finite bandwidth.

In this thesis, we developed maximum likelihood sequence estimation with Viterbi algorithm for Rayleigh fading channels with ISI. The MLSE-VA receivers with and without training sequence assistance are presented, along with a reduced dimensionality channel model which is utilized to simplify the receiver structure.

We demonstrate that for small to moderate channel delay spread the dispersive Rayleigh fading channel can be modeled by only a few random gain parameters instead of by the whole discrete channel impulse tap gains. Especially in the scope of our study our, for rms delay spread less than 20% of the transmitted symbol duration 3 random parameters are enough to describe the multipath Rayleigh fading channel.

By utilizing this novel channel modeling technique, we designed a reduced complexity channel estimator which use the inserted pilot training sequence to estimate the time varying fading channel. The channel estimator uses a block adaptation technique in which two stages, periodic channel estimation and channel interpolation, are involved in order to track the time varying channel. It is shown that this channel estimator has a simple structure due to the contribution of the simplified channel model and can track the dispersive fading channel very well. For small delay spread channel, the Viterbi receiver using our reduced complexity channel estimator shows the same diversity effect and very close bit error rate performance as the one that utilizes the perfect channel state information.

In the thesis, we also proposed a semi-blind MLSE-VA receiver which uses a bank of Wiener filters to estimate the channel information. No pilot sequence is necessary. Techniques such as multiple sampling at the receiver, VA state reduction, antenna diversity, and ect. are investigated. Simulation results show that our MLSE-VA receiver has fairly close performance comparing with the VA receiver that has the perfect knowledge of the channel information. It is demonstrated that the MLSE-VA receiver using antenna diversity technique has a simple structure and better bit error rate performance.

When comparing the proposed VA receivers with and without the pilot training sequence, we found that, considering the complexity and the bit error performance, the receiver using pilot assisted channel estimator shows a great advantage. The cost is that the training sequence occupies certain channel bandwidth.

Appendix A

Estimation error examples

This appendix gives two trivial examples to illustrate the point that estimating fewer channel parameters from received signal is likely to be more accurate than estimating the full impulse response.

Case 1: Assume the total statistical independence among the transmitted signal samples and noise samples.

The received signal is

$$\mathbf{r} = \mathbf{s} + \mathbf{n} \tag{A.1}$$

where $\mathbf{r}$, $\mathbf{s}$ and $\mathbf{n}$ are respectively received signal vector, transmitted signal vector and noise vector, all of length N. The independent transmitted signal samples and additive noise samples are respectively denoted as $s_i,\ i = 1, \cdots, N$ and $n_i,\ i = 1, \cdots, N$. Assume s_i and n_i have variances σ_s^2 and σ_n^2 respectly for $i = 1, \cdots, N$.

The best linear estimate (minimum mean square estimate) for $\mathbf{s}$ is [38]

$$\hat{\mathbf{s}} = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_n^2}\mathbf{r} \tag{A.2}$$

and the total estimation error is

$$(\mathbf{s} - \hat{\mathbf{s}})^\dagger(\mathbf{s} - \hat{\mathbf{s}}) = N\sigma_s^2\left(1 - \frac{\sigma_s^2}{\sigma_s^2 + \sigma_n^2}\right) \tag{A.3}$$

where $\dagger$ denotes the Hermitian transpose.

From (A.3) we see the mean squared estimation error is propotional to N. The error becomes smaller if fewer elements s_i, $i = 1, \cdots$ are estimated.

Case 2: Assume transmitted signal samples are exactly the same.

Assume all the transmitted signal elements are the same x. The received signal vector is

$$\begin{aligned} \mathbf{r} &= \mathbf{s} + \mathbf{n} \\ &= \mathbf{1}x + \mathbf{n} \end{aligned} \tag{A.4}$$

where $\mathbf{r}$, $\mathbf{1}$ and $\mathbf{n}$ are respectively received signal vector, unit vector and noise vector, all of length N. The signal vector $\mathbf{s} = \mathbf{1}x$ and x is a scalar. Assume the transmitted signal elements x is statistically independent with the noise elements $n_i, i = 1, \cdots, N$, also n_i and n_j are independent when $i \neq j$. Assume x has energy σ_x^2, and n_i has the variance σ_n^2 for $i = 1, \cdots, N$.

The best linear estimate for the transmitted signal sample x is [38]

$$\hat{x} = \left(\mathbf{\Phi}_{xr}\mathbf{\Phi}_{rr}^{-1}\right)\mathbf{r} \tag{A.5}$$

where $\mathbf{\Phi}_{rr}$ is the covariance matrix of $\mathbf{r}$, and $\mathbf{\Phi}_{xr}$ is the correlation vector of x and $\mathbf{r}$.

The total estimation error is

$$(\mathbf{s}-\hat{\mathbf{s}})^{\dagger}(\mathbf{s}-\hat{\mathbf{s}}) = N\sigma_x^2\left(1-\frac{N\sigma_x^2}{N\sigma_x^2+\sigma_n^2}\right) \tag{A.6}$$

Again we see the mean squared estimation error becomes smaller if fewer elements s_i is estimated.

Appendix B

Calculation of Characteristic Function in (4.24)

The calculation of the characteristic function $\phi_D(s)$ in (4.24) is shown in this appendix following the procedure similar to that in section 2.2.

The optimal receiver selects the sequence $\hat{\mathbf{c}} = (\hat{c}(1), \cdots, \hat{c}(K))$ that minimizes the metric

$$J(\hat{\mathbf{c}}) = \sum_{n=1}^{K} \left| y[n] - \sum_{\ell=-L}^{L} \hat{c}(n-\ell)\hat{h}_\ell[n] \right|^2 \tag{B.1}$$

where $2L$ is the length of the CIR and $c(-L), \cdots, c(0)$ are known to the receiver. Given the transmitted sequence $\mathbf{c} = (c(1), \cdots, c(K))$, a decoding error occurs if for some erroneous sequence $\hat{\mathbf{c}}$, the random variable D, defined as

$$D = J(\hat{\mathbf{c}}) - J(\mathbf{c}) \tag{B.2}$$

is less than zero. The metric $J(\mathbf{c})$ is given as

$$J(\mathbf{c}) = \sum_{n=1}^{K} |y[n] - \sum_{\ell=-L}^{L} c(n-\ell)\hat{h}_\ell[n]|^2 \tag{B.3}$$

Substitute (3.8) into (B.1) and (B.3) respectively, we have

$$J(\hat{\mathbf{c}}) = \sum_{n=1}^{K} \left| e[n] + \sum_{\ell=-L}^{L} c(n-\ell)h_\ell[n] - \sum_{\ell=-L}^{L} \hat{c}(n-\ell)\hat{h}_\ell[n] \right|^2 \tag{B.4}$$

and

$$J(\mathbf{c}) = \sum_{n=1}^{K} \left| e[n] + \sum_{\ell=-L}^{L} c(n-\ell)h_\ell[n] - \sum_{\ell=-L}^{L} \hat{c}(n-\ell)\hat{h}_\ell[n] \right|^2 \tag{B.5}$$

where $h_\ell[n] = \sum_{j=0}^{\infty} g_j[n]p_j[\ell]$ and $\hat{h}_\ell[n] = \sum_{j=0}^{J} \hat{g}_j[n]p_j[\ell]$.

Define zero mean complex Gaussian random variables a_n, b_n and t_n as follows.

$$a_n = \sum_{\ell=-L}^{L} c(n-\ell)h_\ell[n], \tag{B.6}$$

$$b_n = \sum_{\ell=-L}^{L} \hat{c}(n-\ell)\hat{h}_\ell[n], \tag{B.7}$$

$$t_n = \sum_{\ell=-L}^{L} c(n-\ell)\hat{h}_\ell[n]. \tag{B.8}$$

Let

$$\mathbf{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_K \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_K \end{bmatrix}, \qquad \mathbf{t} = \begin{bmatrix} t_1 \\ \vdots \\ t_K \end{bmatrix}, \qquad \mathbf{e} = \begin{bmatrix} e[1] \\ \vdots \\ e[K] \end{bmatrix}, \tag{B.9}$$

$$\mathbf{I} = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}_{K\times K}, \qquad \mathbf{0} = \begin{bmatrix} 0 & & 0 \\ & \ddots & \\ 0 & & 0 \end{bmatrix}_{K\times K}. \tag{B.10}$$

Write the complex random variable D in matrix form

$$D = J(\hat{\mathbf{c}}) - J(\mathbf{c}) \tag{B.11}$$

$$= \sum_{n=1}^{K} [(e[n] + a_n - b_n)(e[n] + a_n - b_n)^* - (e[n] + a_n - t_n)(e[n] + a_n - t_n)^*]$$

$$= \mathbf{z}^\dagger \mathbf{F} \mathbf{z} \tag{B.12}$$

where

$$\mathbf{z} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{t} \\ \mathbf{e} \end{bmatrix}_{4K\times 1}, \tag{B.13}$$

and

$$\mathbf{F} = \begin{bmatrix} \mathbf{0} & -\mathbf{I} & \mathbf{I} & \mathbf{0} \\ -\mathbf{I} & \mathbf{I} & \mathbf{0} & -\mathbf{I} \\ \mathbf{I} & \mathbf{0} & -\mathbf{I} & -\mathbf{I} \\ \mathbf{0} & -\mathbf{I} & -\mathbf{I} & \mathbf{0} \end{bmatrix}_{4K\times 4K}. \tag{B.14}$$

The superscript † represents the Hermitian transpose.

The correlation matrix of $\mathbf{z}$ is obtained formally without much difficulty.

$$\begin{aligned} \mathbf{R_z} &= \frac{1}{2}E\mathbf{z}^*\mathbf{z}^t \\ &= \frac{1}{2}E\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{t} \\ \mathbf{e} \end{bmatrix}^* \begin{bmatrix} \mathbf{a} & \mathbf{b} & \mathbf{t} & \mathbf{e} \end{bmatrix} \\ &= \frac{1}{2}E\begin{bmatrix} \mathbf{a}^*\mathbf{a} & \mathbf{a}^*\mathbf{b} & \mathbf{a}^*\mathbf{t} & \mathbf{a}^*\mathbf{e} \\ \mathbf{b}^*\mathbf{a} & \mathbf{b}^*\mathbf{b} & \mathbf{b}^*\mathbf{t} & \mathbf{b}^*\mathbf{e} \\ \mathbf{t}^*\mathbf{a} & \mathbf{t}^*\mathbf{b} & \mathbf{t}^*\mathbf{t} & \mathbf{t}^*\mathbf{e} \\ \mathbf{e}^*\mathbf{a} & \mathbf{e}^*\mathbf{b} & \mathbf{e}^*\mathbf{t} & \mathbf{e}^*\mathbf{e} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{R_{aa}} & \mathbf{R_{ab}} & \mathbf{R_{at}} & \mathbf{0}_{K\times K} \\ \mathbf{R_{ba}} & \mathbf{R_{bb}} & \mathbf{R_{bt}} & \mathbf{0}_{K\times K} \\ \mathbf{R_{ta}} & \mathbf{R_{tb}} & \mathbf{R_{tt}} & \mathbf{0}_{K\times K} \\ \mathbf{0}_{K\times K} & \mathbf{0}_{K\times K} & \mathbf{0}_{K\times K} & \mathbf{R_{ee}} \end{bmatrix} \end{aligned} \tag{B.15}$$

Note that $\mathbf{e}$ is uncorrelated with $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{t}$, and $\mathbf{R_z}$ is a Hermitian matrix. Determination of the submatrices of $\mathbf{R_z}$ is straightforward, but detailed.

The $(n_1, n_2)'th$ element, $n_1,\ n_2 = 1, \cdots, K$, in submatrices $R_{\mathbf{aa}}$, $R_{\mathbf{ab}}$, $R_{\mathbf{at}}$, $R_{\mathbf{bb}}$, $R_{\mathbf{bt}}$, $R_{\mathbf{tt}}$, and $R_{\mathbf{ee}}$ can be obtained respectively as follows.

$$
\begin{aligned}
r_{a_{n_1}a_{n_2}} &= \textstyle\sum_{\ell_1=-L}^{L}\sum_{\ell_2=-L}^{L} c(n_1-\ell_1)c(n_2-\ell_2)\overline{h_{\ell_1}[n_1]h_{\ell_2}[n_2]},\\
r_{a_{n_1}b_{n_2}} &= \textstyle\sum_{\ell_1=-L}^{L}\sum_{\ell_2=-L}^{L} c(n_1-\ell_1)\hat{c}(n_2-\ell_2)\overline{h_{\ell_1}[n_1]\hat{h}_{\ell_2}[n_2]},\\
r_{a_{n_1}t_{n_2}} &= \textstyle\sum_{\ell_1=-L}^{L}\sum_{\ell_2=-L}^{L} c(n_1-\ell_1)c(n_2-\ell_2)\overline{h_{\ell_1}[n_1]\hat{h}_{\ell_2}[n_2]},\\
r_{b_{n_1}b_{n_2}} &= \textstyle\sum_{\ell_1=-L}^{L}\sum_{\ell_2=-L}^{L} \hat{c}(n_1-\ell_1)\hat{c}(n_2-\ell_2)\overline{\hat{h}_{\ell_1}[n_1]\hat{h}_{\ell_2}[n_2]},\\
r_{b_{n_1}t_{n_2}} &= \textstyle\sum_{\ell_1=-L}^{L}\sum_{\ell_2=-L}^{L} \hat{c}(n_1-\ell_1)c(n_2-\ell_2)\overline{\hat{h}_{\ell_1}[n_1]\hat{h}_{\ell_2}[n_2]},\\
r_{t_{n_1}t_{n_2}} &= \textstyle\sum_{\ell_1=-L}^{L}\sum_{\ell_2=-L}^{L} c(n_1-\ell_1)c(n_2-\ell_2)\overline{\hat{h}_{\ell_1}[n_1]\hat{h}_{\ell_2}[n_2]},\\
r_{e_{n_1}e_{n_2}} &= N_0\delta(n_1-n_2).
\end{aligned}
\tag{B.16}
$$

Alhough the expression for auto-correlation and cross-correlation between $h_{\ell_1}[m]$ and $h_{\ell_2}[n]$, where m and n represent different time instants, are lenthy, they can be obtained from their defination without much difficulty.

With the matrix $\mathbf{F}$ and correlation matrix $\mathbf{R_z}$, the characteristic function of D is obtained as

$$
\begin{aligned}
\Phi_D(s) &= \frac{1}{det(I+2sR_{\mathbf{z}}^{*}\mathbf{F})}\\
&= \frac{1}{\Pi_{i=1}^{4N}(1+2s\lambda_i)}
\end{aligned}
\tag{B.17}
$$

where $\{\lambda_i, i = 1, \cdots, 4K\}$ are $4K$ real eigenvalues of $R_{\mathbf{z}}^{*}\mathbf{F}$.

The pairwise error probability in (4.24) can be calculated.

Bibliography

[1] W.Y.C.Lee, "Mobile Communications Engineering". *New York: McGraw-Hill*, 1982.

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

[3] K. Pahlavan and A.H. Levesque, "Wireless Data Communications", *Proceedings of the IEEE*, Vol. 82, No.9, pp. 1398-1430, Sept. 1994.

[4] J.E. Padgett, C.G. Gunther and T. Hattori, "Overview of Wireless Personal Communications", *IEEE Communications Magazine*, pp. 28-41, Jan. 1995.

[5] J.G. Proakis, "Adaptive equalization for TDMA digital mobile radio", *IEEE Trans. Veh. Technology*, vol 40, pp. 333-341, May 1991.

[6] D.D. Falconer and L. Ljung, "Application of fast Kalman estimation to adaptive equalization", *IEEE trans. Commun.*, vol. COM-26 pp. 1439-1446, Oct. 1978.

[7] D.D. Falconer, et.al., "Comparison of DFE and MLSE receiver performance on HF channels", *IEEE Trans. Commun.*, vol. COM-33, pp. 484-486, may 1985.

[8] K.Pahlavan, S. Howard, and T. Sexton, "Adaptive Equalization of the indoor radio channel", *IEEE Trans. Commun.*, vol. 41, pp. 164-170, Jan. 1993.

[9] J.G. Proakis, "Advances in equalization for intersymbol interference", *Advances in Communication Systems*, vol 4, A.J. Viterbi, Edit. New York: Academic Press, 1975.

[10] P. Monsen, "MLSE equalization of interference on fading diversity channels", *IEEE Trans, Commun,.*, vol. COM-32, pp. 5-12, Jan. 1984.

[11] G.D.Forney, Jr, "Maximum Likelihood Sequence Estimation of Digital Sequence in the Presence of Intersymbol Interference", *IEEE Trans. on Information Theory*, vol. IT-18, pp. 363-378, may 1972.

[12] R.F. Magee and J.G. Proakis, "Adaptive Maximum Likelihood Sequence Estimation for Digital Signaling in the Presence of ISI", *IEEE Trans. Informat. Theory*, vol. IT-19, pp. 120-124, 1973.

[13] F.Xiong, A.Zerik, and E. Shwedyk, "Sequential Sequence Estimation for Channels with Intersymbol Interference of Finite of Infinite Length", *IEEE Trans. on Communications*, vol.38, N0.6, pp.795-804, June 1990.

[14] D.L. Duttweiler, J.E. Mazo, and D.G. Messerschmitt, "Error propagation in Decision-feed-back equalizers", *IEEE Trans. Information Theory*, vol. IT-20, pp. 490-497, July 1974.

[15] D.Borth, K. Baum. B.Mueller, "A comparison of Nonlinear equalization Methods for the U.S. Digital Cellular System", *Proc., IEEE supercom/ICC 92*, pp. 312.1.1-5, Chicago, June 92.

[16] R.E. Morley, Jr., and D.L.Snyder, "Maximum Likelihood Sequence Estimation for randomly Dispersive Channels", *IEEE Trans. on Communications*, vol. COM-27, No. 6, pp.833-839, June 1979.

[17] S. Aghamohammadi and H. Meyr, "A new method for phase synchronization and automatic gain control of linearly modulated signals on frequency-flat fading channels", *IEEE Trans. Communications*, Vol. 39, pp. 25-29, Jan. 1991.

[18] M.L. Moher and J.H. Lodge, "TCMP A modulation and coding strategy for Raician fading channels", *IEEE J. Sel. Areas Commun.*, Vol. 7, No. 9, pp. 1347-1355, Dec. 1989.

[19] J.K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels", *IEEE Transactions on Vehicular Technology, Vol. 40, No. 4*, 686-693, 1991.

[20] S.N. Crozier, D.D. Falconer and S.A. Mahmoud, "Least sum of square errors (LSSE) channel estimation", *Inst. Elec. Eng. Proc., pt. F*, vol. 138, pp. 371-378,

Aug. 1991.

[21] N.W.K. Lo, D.D. Falconer, and A.U.H.Sheikh, "Adaptive Equalization and Diversity Combining for Mobile Radio using Interpolated Channel Estimates", *IEEE Trans. on Vehicular Technology*, Vol. 40, No. 3, pp. 636-645, August 1991.

[22] S.A. Fechtel and Heinrich Meyr, "Optimal parametric feedforward estimation of frequency-selective fading radio channels", *IEEE Trans. on Communications*, VOL. 42, No. 2/3/4, pp. 1639-1650, 1994.

[23] P. Ho and J.H. Kim, "On pilot symbol assisted detection of MSK and GTFM in fast fading channels", *To be published.*

[24] R. Hab and H. Meyr, "A Systematic approach to Carrier Recovery and Detection of Digitally Phase-Modulated signals on fading Channels", *IEEE Trans. Commun.*, Vol. 37, No.7, pp. 748-754, July, 1989.

[25] J.H. Lodge and M.L. Moher, "Maximum Likelihood Sequence Estimation of CPM Signals Transmitted over rayleigh Flating fading Channels", *IEEE Trans Commun.*, vol. 38, No. 6, pp. 787-794, June 1990.

[26] P. Hoher, "An Adaptive Channel Estimator for Frequency Selective fading Channels", *Aachener Symposium fuer Signaltheorie*, pp. 168-173, Sept. 1990.

[27] S. McLaughlin, B. Mulgrew, and C.F.N. Cowan, "Performance Study of the Extended Kalman Algorithm as an HF Channel Estimator", *IEE 4th International*

Conference HF Radio Systems and Techniques, pp. 335-338, Apr. 1988.

[28] Y. Sato, "A Blind Sequence Detection and Its Application to Digital Mobile Communication", *IEEE Journal on Selected Areas in Communications*, vol. 13, No.1, pp. 49-58, Jan. 1995.

[29] Q. Dai, E.D. Shwedyk, "Detection of Bandlimited Signals Over Frequency Selective Rayleigh Fading Channels", *IEEE Trans. on Communications*, Vol. 42, No. 2/3/4, pp. 941-950, 1994.

[30] G.L. Turin, et.al., "A Statistical Model of Urbon Multipath Propagation", *IEEE Trans. Veh. Technology*, vol. VT-21, No. 2, pp. 1-9, Jan. 1972.

[31] S.A. Fechtel, "A Novel Approach to Modeling and Efficient Simulation of Frequency- Selective fading radio Channels", *IEEE Journal on Selected Areas in Communications*, Vol. 11, No.3, pp. 442-431, 1993.

[32] J.K. Cavers, "On the validity of the slow and moderate fading models for matched filter detection of Rayleigh fading signals", *Canadian Journal of Electrical and Computer Engineering*, vol. 17, pp. 183-189, October 1992.

[33] J.K. Cavers and P. Ho, "Analysis of the error performance of Trellis-coded modulations in Rayleigh fading channels," *IEEE Transactions on Communications*, Vol. 40, pp. 74-83, 1992.

[34] M.E. Rollins, M.E. and S.J. Simmons "Error performance analysis of MLSE for frequency selective Rayleigh fading channels with Kalman channel estimation", *International Conference on Communications*, pp. 321-326, 1994.

[35] M. Schwartz, W. Bennett, and S. Stein, *Communication systems and techniques* New York, McGraw-Hill, 1966.

[36] T. Chan and P. Ho, "Bit Error probability of uncoded QPSK transmitted over a 2-ray frequency selective Rayleigh fading channel", *International Conference on Communications*, pp. 0311-0315, 1992.

[37] B.D.O. Anderson and J.B. Moore, "Optimal filtering", Prentice Hall, 1979.

[38] Y.L. Tong, "The Multivariate Normal Distribution", *Springer-Verlag New York Inc.*, 1990.

[39] P. Hoeher, "A Statistical Distrete-Time Model for the WSSUS Multipath Channel", *IEEE Trans. on Vehicular Technology*, Vol. 41, No. 4, pp. 461-468, November 1992.