Multiple Model Estimation for Channel Equalization and Space-Time Block Coding

MULTIPLE MODEL ESTIMATION FOR CHANNEL EQUALIZATION AND SPACE-TIME BLOCK CODING

By

ZIAUDDIN M KAMRAN, M. Eng.
Nanyang Technological University, Singapore, 2001

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfilment of the Requirements
for the Degree
Master of Applied Science

McMaster University

MASTER OF APPLIED SCIENCE (2004) MCMASTER UNIVERSITY

(Electrical and Computer Engineering) Hamilton, Ontario, Canada

TITLE: **Multiple Model Estimation for Channel Equalization and Space-Time Block Coding**

AUTHOR: Ziauddin M Kamran

M. Eng.

Nanyang Technological University

Singapore, 2001

SUPERVISOR: Dr. T. Kirubarajan

NUMBER OF PAGES: xii, 114

To my family

Abstract

This thesis investigates the application of multiple model estimation algorithms to the problem of channel equalization for digital data transmission and channel tracking for space-time block coded systems with non-Gaussian additive noise. Recently, a network of Kalman filters (NKF) has been reported for the equalization of digital communication channels based on the approximation of the *a posteriori* probability density function of a sequence of delayed symbols by a weighted Gaussian sum. A serious drawback of this approach is that the number of Gaussian terms in the sum increases exponentially through iterations. In this thesis, firstly, we have shown that the NKF-based equalizer can be further improved by considering the interactions between the parallel filters in an efficient way. To this end, we take resort to the Interacting Multiple Model (IMM) estimator widely used in the area of multiple target tracking. The IMM is a very effective approach when the system exhibits discrete uncertainties in the dynamic or measurement model as well as continuous uncertainties in state values. A computationally feasible implementation based on a weighted sum of Gaussian approximation of the density functions of the data signals is introduced. Next, we present an adaptive multiple model blind equalization algorithm based on the IMM estimator to estimate the channel and the transmitted sequence corrupted by intersymbol interference and noise. It is shown through simulations that the proposed IMM-based equalizer offers

substantially improved performance relative to the blind equalizer based on a (static or non-interacting) network of extended Kalman filters. It obviates the exponential growth of the state complexity caused by increasing channel memory length. The proposed approaches avoid the exponential growth of the number of terms used in the weighted Gaussian sum approximation of the plant noise making it practical for real-time processing.

Finally, we consider the problem of channel estimation and tracking for space-time block coded systems contaminated by additive non-Gaussian noise. In many practical wireless channels in which space-time block coding techniques may be applied, the ambient noise is likely to have an impulsive component that gives rise to larger tail probabilities than is predicted by the Gaussian model. Although Kalman filters are often used in practice to track the channel variation, they are notoriously sensitive to heavy-tailed outliers and model mismatches resulting from the presence of impulsive noise. Non-Gaussian noise environments require the modification of standard filters to perform acceptably. Based on the coding/decoding technique, we propose a robust IMM algorithm approach in estimating time-selective fading channels when the measurements are perturbed by the presence of impulsive noise. The impulsive noise is modeled by a two terms Gaussian mixture distribution. Simulations demonstrate that the proposed method yields substantially improved performance compared to the conventional Kalman filter algorithm using the clipping or localization approaches to handle impulses in the observation. It is also shown that IMM-based approach performs robustly even when the prior information about the impulsive noise is not known exactly.

Acknowledgements

Firstly, it is a delight to be able to express profuse and sincere thanks to my supervisor, Dr. T. Kirubarajan, for his continuous support and invaluable guidance during my research work. He has been instrumental in the successful completion of this work and without his diligent research, intense supervision, and unfathomable friendship this thesis would simply not have been possible. I am also indebted to Dr. Alex Gershman for his fruitful discussion and suggestions pertinent to this work.

My stay at Estimation, Tracking and Fusion Laboratory is one of the most memorable time of my life because of the academic atmosphere and well-established public relationship. I must also voice my heartfelt thanks to all my friends at McMaster for making my stay here an enjoyable one.

Additionally, I would like to express my respect and appreciation to all the staff members of our department for their dedication to work and kind assistance. Special thanks goes to Cheryl Gies and Helen Jachna for their help and support.

The financial support of the Ontario Graduate Scholarship program provided by the ministry of Training, Colleges and Universities, Canada is gratefully acknowledged.

Finally, enormous thanks are due to my beloved wife Nimun for her constant love and aspiration. I am widely indebted to my parents and my family for their untiring support and encouragement.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The global demand for wireless communications has grown apace over the last decade and the growth is anticipated to continue over the next decade. Future generation wireless communication technology promises to have increased data rate capability, larger channel capacity, improved spectral efficiency for the need of a greater range of services: video conferencing, internet services, data networks and multimedia, extended range coverage, and smaller and portable terminals that allow for low-power operations. In addition, these future systems should be reliably operated in both indoor and outdoor environments.

Wireless channels with delay spread exceeding the symbol period introduce frequency-selectivity. The growth in demand on the data rate capabilities of wireless systems necessitates an increase in bandwidth and signaling rate. As we go for higher rates, the multipath distortion or frequency-selective fading caused by physical medium becomes worse. The distortion caused by the resulting overlap of received signals is called intersymbol interference (ISI), which, if left uncompensated, drastically deteriorates the received signal. As a consequence, it is necessary for the receiver to compensate or remove these impairments from the received message in order to recover the transmitted message. Such a compensator for

the ISI is called an equalizer. The delay spreads (in symbol durations) of multipath wireless channels critically depends on the signaling bandwidth and physical environment. For example, in an urban environment, a global system for mobile communication (GSM) system with a symbol rate of 270.83 Kbaud has a typical delay spread of $5\mu s$. On the other hand, a digital terrestrial TV broadcast channel uses a symbol rate of 10.78 Mbaud and can have worst case delay spreads in the order of $50\mu s$. The baseband digital impulse response for a digital TV channel has a span of about 600 symbol taps. Hence, the equalizer plays an important role to the design of a receiver.

Fading, which arises mainly from destructive addition of multipaths in the propagation medium, is a major performance-limiting impairment on wireless channels. Diversity techniques mitigate fading by transmitting multiple correlated replicas of the same information signal through independently fading channel realization that are much likely to fade simultaneously than each individually. In recent years, space-time coding (STC), which is a technique that exploits the combination of spatial and temporal diversity, has been shown to be very effective in combating fading and increasing channel capacity significantly without necessarily sacrificing bandwidth efficiency. Space-time codes provide both diversity and coding gain when using multiple transmit antennas to increase spectral efficiency over wireless communication systems. There are two different schemes in space-time codes, namely, space-time trellis coding (STTC) and space-time block coding (STBC). STTC operates on one input symbol at a time, resulting a sequence of vector symbols whose length represents transmit antennas. Like traditional trellis coded modulation (TCM) for a single-antenna channel, STTC provides coding gain. Disadvantage of STTCs is that they are difficult to design and require high complexity in encoders and decoders. On the other hand, STBC operates on a block of input symbols, resulting a coding matrix whose columns represent antennas and rows represent time. Between these two schemes, STBC looks very attractive

because of its simplicity in decoding scheme especially when complexity is a main issue of the system design. STBC achieves the same amount of diversity as STTC even though it does not give as much coding gain as STTC.

1.1 Background

In this Section, a summary of some of the previous work related to the areas of equalization techniques and space-time coding schemes is presented. This is not meant to be a comprehensive literature survey, but will provide the readers with the background references on some of the ongoing research in the subject areas.

1.1.1 Equalization Techniques

Various approaches to data detection can be broadly classified into symbol-by-symbol and sequence estimation [57]. The first class contains linear and decision-feedback detectors. The other approach to data detection is the maximum likelihood sequence estimations (MLSE) [26]. The most effective detection technique is MLSE. In fact, it is the optimum equalizer for digital signals corrupted by ISI and buried with additive noise in the sense of minimizing the probability of sequence error, given knowledge of the channel impulse response [17,58]. The trellis-based Viterbi algorithm [27] solves the MLSE problem recursively when the memory of the channel is finite. However, the complexity of the algorithm is proportional to the number of states in the trellis which grows exponentially with the channel memory length. For instance, if the size of the symbol alphabet is $\mathcal{A}$ and the number of interfering symbols contributing to ISI is L, the Viterbi algorithm computes $\mathcal{A}^{L+1}$ branch metrics for each new received symbol. When the channel memory becomes large or the alphabet size is big, then the MLSE approach becomes impractical to track the fast fading

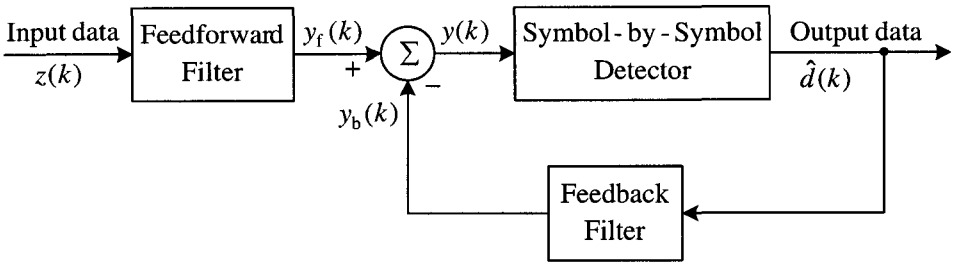

Figure 1.1: Block diagram of a decision feedback equalizer.

channels in high-speed mobile communication [7, 17, 57]. The computational burden can be eased by reducing the number of states and sequences in the MLSE detector as suggested by some schemes, e.g., reduced-state sequence estimation (RSSE) [24, 25] and delayed decision-feedback sequence estimation (DDFSE) [23]. These approaches assume some past decisions as correct while estimating several most recent symbols.

One suboptimal scheme that is widely used in practice for its simplicity in implementation and analysis is the linear equalizer (LE). This approach employs a linear transverse filter and has a computational complexity that is a linear function of the channel dispersion length L. For a channel introducing mild interference, the performance achieved by a conventional LE is often adequate. However, in suppressing the ISI, it causes noise enhancement at channel spectral nulls and it is often not suitable for applications where frequency-selective fading takes place [57]. Nonlinear equalizers show better performance than linear equalizers in applications where the channel distortion is severe. A popular nonlinear equalizer is the decision feedback equalizer (DFE). The DFE outperforms a linear equalizer of equivalent complexity [19]. Compared to MLSE, decision feedback equalization is a suboptimum and less complex equalization technique. As depicted in Figure 1.1, the DFE consists of a linear feedforward filter (FFF) which is used to suppress the effects of noise and precursors caused

by the future symbol and a feedback filter (FBF) whose task is to properly weight the decision of the previous symbol so that the postcursors caused by the previous symbol can be cancelled out. The result is then applied to a threshold device to determine the current symbol decision. The FFF enhances the noise, but the noise gain is not as severe as in the case of a linear equalizer. The DFE design is typically been carried out assuming that the past decisions are error free, thus simplifying the mathematics involved. However, when an error is made by the receiver, the output of the FBF is no longer the desired value and the probability of subsequent errors is increased resulting in errors that tend to occur in burst. The first residual-induced decision error, called a primary error, is fed back by the FBF causing secondary errors and creating an error burst. This phenomenon, known as *error propagation*, is more severe when tap weights and/or the number of the feedback taps are large. Some of these DFEs are modified to meet the need of a specific application. Belfiore *et al.* proposed a distortion predictive DFE whose FBF is a predictor, driven by the difference between the outputs of the FFF and the decision device [13]. As a result, the noise and the residual ISI at the output of the FFF can be predicted by the FBF and subtracted from the FFF output.

As an alternative to classical equalizers, Kalman filtering-based equalizers have been reported in [47,54] under the assumption that the plant noise is Gaussian, which is not valid in the context of data channel equalization. It has been shown that Kalman filter outperforms the LE. Recently, an equalizer based on a network of Kalman filters operating in parallel has been proposed in [50] which yields improved performance compared to the DFE and the classical Kalman equalizer. However, it did not consider the interaction between the parallel filters during the estimation process.

Blind equalization has attracted much research interest in wireless communications and related fields over the past few decades. Blind equalization refers to the reconstruction

of transmitted symbols based on the noise-corrupted channel output without knowing the underlying finite impulse response (FIR) channel where the transmitted digital symbols are distorted by the base-band equivalent discrete-time linear FIR channel, causing ISI. Since the performance of a linear equalizer in a severe ISI is not satisfactory, nonlinear blind equalization techniques have been addressed recently. For instance, the algorithms in [35,62] apply a sequence estimator and a bank of channel estimators and alternatively optimize with respect to data and channel. Sequence estimation is performed by a blind search of modified trellis and channel estimation is accomplished by conditioning on survivor sequences in the trellis and constructing the corresponding maximum likelihood (ML) or minimum mean-square error (MMSE) channel estimate. Blind equalization approaches based on symbol-by-symbol maximum *a posteriori* probability (MAP) have also been reported in the literature, e.g., see [36] and reference therein. Another set of equalizers employ a hidden Markov model (HMM) formulation for blind (or semi blind) equalization for input sequences governed by Markov chains. To maximize the Kullback-Leibler measure to calculate the HMM model, they use either off-line [40] or on-line expectation-maximization (EM) algorithm [45]. A stochastic ML blind channel estimation scheme is proposed in [21] based on the HMM formulation of the problem. Although on-line methods subdue the memory and computational cost involved in the off-line EM algorithm based method, they still need to use some kind of state reduction algorithm to reduce the state complexity of the state trellis. Recently, a blind Network of Extended Kalman Filters (NEKF) has been proposed for linear channel equalization in [4]. The network relies on the assumption that the density functions of the data signals can be represented by a weighted Gaussian sum as suggested in [3,50]. It has been shown that the NEKF-based blind approach achieves good performance compared to the blind Bayesian algorithm proposed by Iltis *et al.* in [36], with a lower complexity. For a excellent review on channel estimation and equalization, readers are encouraged to see the

seminal paper by Tugnait *et al.* [69].

1.1.2 Space-Time Block Coding

The interest in the multiple transmitter system started with the base station simulcast problem studied in [72]. The delay diversity scheme was suggested as a solution to this problem. In this scheme, replicas of one antenna are transmitted from the other antennas separated in time. Delay diversity [72, 73] and other related schemes [61, 71] are among the first techniques presented to exploit transmit diversity. Delay diversity can be viewed as a special case of STTC, later proposed by Tarokh *et al.* [65]. The generalized approach combines TCM with transmit diversity techniques. Although decoding complexity of these codes (measured by the number of trellis states at the decoder) increases exponentially as a function of the diversity level and transmission rate, they perform very well in slow fading environments. The rank and determinant criteria emerged from this work become a benchmark in space-time code design. Ensuring full diversity, in [30], a more structured method is presented.

In addressing the issue of decoding complexity, Alamouti [2] discovered an ingenious STBC scheme for transmission with two antennas. This coding scheme, remarkable for having an elegant and simple linear receiver, becomes a paradigm in STBC. However, the simple decoding rule of this scheme is valid only for a flat-fading channel where the channel gain is constant over two consecutive symbols. The Alamouti's STBC has been generalized using the theory of orthogonal designs in [66,67], which have full diversity and employ linear ML detectors that decouple the transmitted symbols. Recently, the space-time block code has been further generalized in [31] to achieve a higher capacity. Naguib *et al.* [55] presents a comprehensive review of space-time coding schemes.

1.2 Motivation

The objective in a signal processing problem is to process a finite number of data samples and extract important information which may be "hidden" in the data. This objective is usually achieved by combining the development of mathematical formulations with their algorithmic implementations (either in software or hardware) and their applications to real data. Various conflicting figures of merit are associated with digital signal processing techniques, e.g., quality of the estimates, computational complexity, data throughput rate, cost of implementation, finite word-length effects, and structural properties.

In this thesis, we focus on the application of the multiple model approaches for data channel equalization and channel estimation for space-time block coded systems with additive non-Gaussian noise. Specifically, we employ the Interacting Multiple Model (IMM) estimator which is based on the Kalman filtering techniques and incorporates a Bayesian framework dealing with dynamic situation of switching factors. The IMM estimator is a very effective approach when the system exhibits discrete uncertainties in the dynamic or measurement model as well as continuous uncertainties in state values. In the IMM method, several filters are used in parallel to estimate the state of a dynamic system with several modes of operation. In addition, the IMM algorithm is *decision free* [10] in the sense that at each time only the probabilities (conditioned on the available data) of each model being the prevailing one are evaluated.

It has been shown that the NKF-based equalizer [50] outperforms the other more classical equalizers like the linear transversal equalizer, the decision feedback equalizer, and the classical Kalman filter equalizer while NEKF-based blind approach [4] achieves good performance compared to the blind Bayesian algorithm reported in [36]. The NKF and NEKF equalizers rely on a weighted sum of Gaussian approximation of the density functions of the data signals. However, a serious drawback of these approaches is that the number of

Gaussian terms in the sum increases exponentially through iterations. These approaches compute the state estimate that accounts for each possible current model without considering any possible switching between the models. In this thesis, it is demonstrated that both the NKF and NEKF equalizers can be further improved by using IMM algorithm which handles the interactions between parallel filters in an efficient way. The proposed approaches avoid the exponential growth of the number of terms used in the weighted Gaussian sum approximation of the plant noise making it practical for real-time processing.

For many wireless communication systems, the estimation of time-varying channel with high accuracy is a challenging task in the receiver design and adaptive wireless channel tracking is an important way to obtain the required channel state information. Time-varying multipath fading is a major performance-limiting impairment on wireless channels which makes the wireless transmission difficult. Recently, STBC has been studied extensively as a method to alleviate the detrimental effects of wireless fading channels because of its simplicity and the feasibility to have multiple antennas at the base station. In space-time block coded systems, the channels are normally assumed to remain block stationary. In real life, this is not the situation and negligence of channel variation through the block results in degradation in performance. With the assumption of additive white Gaussian noise (AWGN), the problem of time-selective fading channel tracking for STBC has been investigated in [49] by using a single Kalman filter. Thus far, most of the work on this area assumes that the channel ambient noise is Gaussian. However, the noise in many physical channels such as urban and indoor radio channels [11, 12, 51] and underwater acoustic channels [52] exhibits Gaussian as well as non-Gaussian characteristics due to the impulsive phenomena of radio-frequency interference. For recent measurement results of impulsive noise in outdoor/indoor mobile and portable radio communications, see [12] and references therein. It is well known that filtering techniques based on Gaussian assumption, which underlies aforementioned channel tracking

algorithm, is not an appropriate one due to its nonrobustness against "outlier" introduced by the impulsive noise [59]. This raises the issue of devising robust wireless channels tracking algorithms that account the maybe impulsive behavior of the channel noise. Therefore, to resolve this problem, several "robust to impulsive noise" estimation algorithms have been studied (see [42,59] and references therein). The typical method of the removal of impulsive noise is the Huber's min-max approach. The performances of these algorithms depend on the threshold value which is generally unknown. In this thesis, we present an alternative robust approach utilizing the IMM algorithm to tackle the ambient impulsive channel noise. It is shown that proposed approach is less vulnerable to the impulsive noise and exhibits superior performance compared to the KF approach with a threshold scheme.

1.3 Organization of This Thesis

This thesis is concerned with digital channel equalization and channel estimation in space-time coded systems which are generic problems in the areas of signal processing and communications. The thesis is organized as follows.

In the first Chapter, we have motivated our work in the context of channel equalization and transmit diversity techniques for wireless communications and the need for methods based on the IMM estimator. We also gave a brief review of the available literature on the problems of channel equalization and STBC. In Chapter 2, we provide fundamental concepts on digital channel equalization and STBC schemes.

In Chapter 3, multiple model estimation techniques to estimate the state of the systems whose models vary with time are reviewed. We also discuss that the IMM estimator is considered to be the best compromise between performance and complexity.

A computationally feasible implementation for the problem of channel equalization based

on a weighted sum of Gaussian approximation of the density functions of the data signals is introduced in Chapter 4. We also present an adaptive blind equalization technique to estimate the channel and the transmitted sequence corrupted by ISI and noise. Numerical examples demonstrate the effectiveness of the proposed approaches. We simulate the performance of the algorithm and the results confirm that the new equalization techniques yield superior performance compared to existing algorithms.

Chapter 5 introduces a novel multiple model based time-selective fading channels estimation approach for space-time block coded systems with non-Gaussian ambient noise. The proposed adaptive channels tracking scheme is based on a state-space representation of the communication system and the prior information of the measurement noise. Again, computer simulations demonstrate the excellence of the incorporation of IMM estimator for solving a dynamic system with several behavior modes which can "switch" from one to another.

Finally, we conclude in Chapter 6 with a summary of the thesis and avenues for further research.

1.4 Related Publications

1. Z. M. Kamran, T. Kirubarajan, and A. B. Gershman, "Blind estimation and equalization of time-varying channels using the interacting multiple model estimator," in *Proceedings of 2004 IEEE International Symposium on Circuits and Systems*, Vancouver, BC, vol. 5, pp. 17–20, May 2004.

2. Z. M. Kamran, T. Kirubarajan, and A. B. Gershman, "Tracking time-selective fading channels for space-time coding in impulsive noise," in *Proceedings of the IEEE Canadian Conference on Electrical and Computer Engineering*, Niagara Falls, ON, vol. 2, pp. 685–688, May 2004.

Chapter 2

Channel Equalization and Space-Time Block Coding

This Chapter deals with the basic concepts pertaining to channel equalization and space-time block coded systems.

2.1 Channel Equalization

Signals that pass through a channel undergo various forms of distortion. A fundamental limiting factor in the performance of digital communication systems is the intersymbol interference. ISI is caused by interference from the symbols that were transmitted before and after the given symbol. ISI-induced errors can cause the receiver to misinterpret the received samples. In a band-limited digital communication system (e.g., in telephone channels), the transmitted digital symbols are perturbed by the base-band equivalent discrete-time linear FIR channel, causing ISI. ISI can also arise from frequency-selectivity (fading or multipath propagation) in digital microwave radio and in mobile communication systems. Figure 2.1 shows a mobile wireless propagation scenario. Since an impulse response corresponding to a

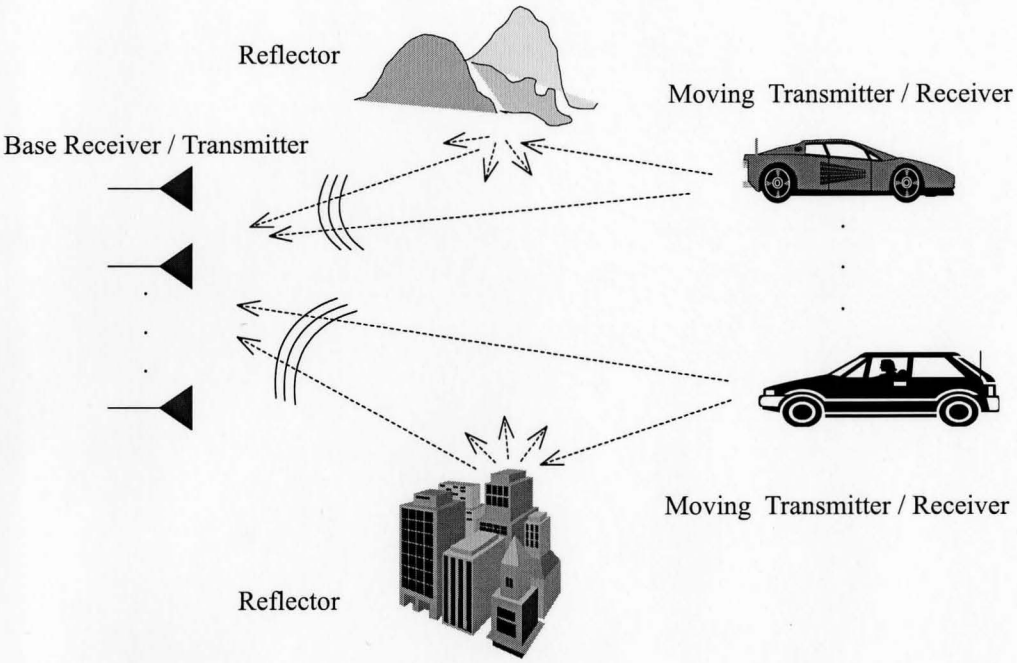

Figure 2.1: A mobile radio propagation scenario.

lack of ISI contains a single spike of width less than the time between symbols, the presence of ISI is readily observable in the sampled channel impulse response [39]. An example of a $T/2$-spaced (symbol spacing T) terrestrial microwave channel impulse response [obtained from the Rice University Signal Processing Information Base (SPIB)[1]] is depicted in Figure 2.2. A channel introducing such distortions is modeled by a finite-memory tapped delay line whose memory accounts for ISI and is followed by additive white Gaussian noise [57].

The *equalizer* is an important part of a modern digital communications receiver. It filters the received data to minimize the impairments from the received message in order to recover the transmitted message.

[1]This microwave channel database can be obtained form `http://spib.rice.edu/spib/microwave.html`.

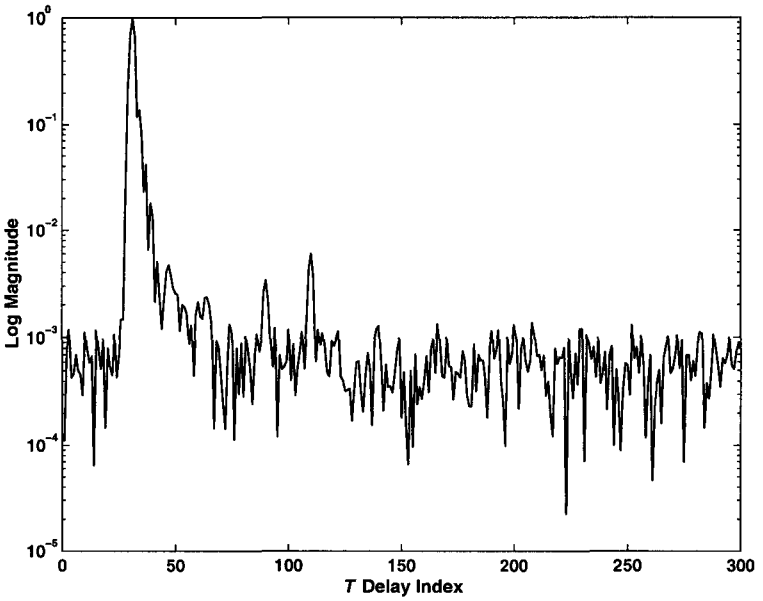

Figure 2.2: Magnitude of a terrestrial microwave channel impulse response with $1/T = 30 \times 10^6$ symbols/sec, (SPIB channel #9).

In this Section, the models that are used to characterize the digital communication channels are discussed. We also give a brief discussion of the various equalizer structures depending on the state-space representation of communication systems that are used to combat the signal distortions caused by channel and the need for blind approach.

2.1.1 Communication Channels

Signals propagating through wireless channels (indoors or outdoors) usually arrive at the destination through a number of different paths, referred to as multipaths. These paths arise from scattering, reflection, refraction, or diffraction of the radiated energy off of objects that lie in the environment. Multipath propagation results in a received signal that is a superposition of several delayed and constructively and destructively interfering copies of

the transmitted signal giving rise to frequency-selective fading. The environment around the transmitter and receiver can change over time, particularly in a mobile setting, leading to variations in the channel response with time. This gives rise to time-selective fading. Also, the channels may have a dominant path (direct path in line-of-sight channels) in addition to several secondary paths or they may be characterized as having multiple "random" paths with no single dominant path.

2.1.1.1 Frequency-Selective and Time-Nonselective Channel

The baseband channel consisting of the modulation, a time-invariant propagation channel, and the receiver filters can typically be modeled by a time-invariant discrete-time FIR filter as

$$\begin{aligned} z(k) &= h(q^{-1})d(k) + n(k) \\ &= \left(h_0 + h_1 q^{-1} + \ldots + h_{L-1} q^{L-1}\right) d(k) + n(k) \\ &= h_0 d(k) + h_1 d(k-1) + \ldots + h_{L-1} d(k-L+1) + n(k) \\ &= \sum_{l=0}^{L-1} h_l d(k-l) + n(k) \end{aligned} \tag{2.1}$$

where k is an integer representing the discrete time, $z(k)$ is the sampled receive signal, $d(k)$ is the sequence of transmitted symbols, $\{h_l\}$ are the coefficients of the FIR filter $h(q^{-1})$ of degree $L-1$ representing the channel for the desired signal, and $n(k)$ is AWGN with zero mean and variance σ_n^2. Model (2.1) represents a frequency-selective linear channel with no time selectivity. A tapped delay line structure for this model is shown in Figure 2.3. It is the most commonly used model for receiver design [69]. In many models, noise is assumed to be independent and identically distributed (i.i.d.). In each received symbol there is some contribution from a number of the past symbols causing ISI. Equalizers are designed to

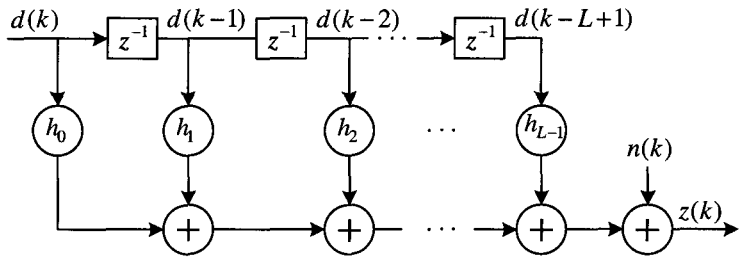

Figure 2.3: Tapped delay line model for frequency-selective channel with FIR.

compensate the effects of ISI resulting from frequency-selective channels.

If in (2.1) the coefficients $\{h_l\}$ are constant, this channel is called *time-invariant* channel.

2.1.1.2 Time- and Frequency-Selective Channel

The multipath propagation model for the channel results in *signal fading*. Fading channels with time variations in the phase and amplitude are typical examples of time-varying channels. For the fading channels, the impulse response of a channel model is no longer fixed constants.

The channel model having time-varying linear system response can be represented as

$$z(k) = \sum_{l=0}^{L-1} h_l(k)d(k-l) + n(k) \tag{2.2}$$

where $h_l(k)$ is the channel response at time k to a unit input at time $k-l$. Model (2.2) represents a time- and frequency-selective linear channel. If in Equation (2.2) the channel coefficients vary with time, this channel is called *time-varying* channel.

For a slowly (compared to the baud rate) time-varying system, Equation (2.2) is often simplified to a time-invariant system represented by Equation (2.1) where $h_l = h_l(0)$ is the time-invariant channel response to a unit input at time 0.

2.1.1.3 Time-Selective and Frequency-Nonselective Channel

If the coherence time of the channel is small compared to the duration of the received signal (symbol duration plus the channels delay spread) but the coherence bandwidth of the channel is large compared to the message bandwidth, then we have

$$h_l(k) = h(k)\delta(l, 0) \tag{2.3}$$

where $\delta(l, 0)$ is the Kronecker delta located at 0, i.e., $\delta(l, 0) = 1$ for $l = 0$ and $\delta(l, 0) = 0$ for $l \neq 0$. Such a channel can be modeled as

$$z(k) = h(k)d(k) + n(k). \tag{2.4}$$

Model (2.4) represents a time-selective and frequency-nonselective linear channel.

2.1.1.4 Time-Nonselective and Frequency-Nonselective Channel

If the channel is flat in both time and frequency, the received signal sample can be represented with the following channel model

$$z(k) = hd(k) + n(k) \tag{2.5}$$

where h is a random variable (or a constant). Model (2.5) represents a time-nonselective and frequency-nonselective linear channel.

All of the above channel response functions may be modeled as deterministic or random [69].

2.1.2 Equalizer Structures

Many problems in science and engineering require estimation of the state of a system that changes over time using a sequence of noisy observed data. In this thesis, we concentrate on the state-space description of the communication systems and the focus is on the discrete-time formulation of the problem. Hence, the evolution of the system with time can be modeled using difference equations and observed data are assumed to be available at discrete times. In order to analyze and make inherence about a dynamic system, at least two following models are required [10, 32].

- *System model:* a model which describes the evolution of the state with time.
- *Measurement model:* a model which associates the noisy measurements to the state.

It is assumed that these models for the underlying communication systems are available in a probabilistic form.

We consider the transmission of a digital data over a baseband channel. A digital communication channel with ISI may be approximately modeled by an equivalent discrete time transversal filter with additive white noise [57]. Assuming perfect carrier and symbol synchronization at the receiver, the baseband time invariant channel output $z(k)$, at time k, can be represented by the following state-space model

$$\mathbf{D}(k) = \mathbf{F}\mathbf{D}(k-1) + \mathbf{G}d(k) \tag{2.6}$$

$$z(k) = \mathbf{H}^T\mathbf{D}(k) + n(k) \tag{2.7}$$

where

- F is the $L \times L$ one-step transition matrix with all elements zero except those in positions $(i+1, i)$, $i = 1, \ldots, L-1$, which are 1,

- $\mathbf{G} = [1, 0, \ldots, 0]^T$ is the $L \times 1$ vector,
- $\mathbf{D}(k) = [d(k), d(k-1), \ldots, d(k-L+1)]^T$ consists of the L last transmitted symbols, taking values from a finite set,
- $\mathbf{H} = [h(0), h(1), \ldots, h(L-1)]^T$ is the parameter vector of the channel impulse response of memory length L that represents the combined effect of the transmitting filter, the channel, the matched filter, the sampler, and the discrete-time noise whitening filter [57], and
- $n(k)$ is the measurement noise which represents AWGN sequence independent of the input sequence $d(k)$ having variance σ_n^2.

According to this state-space model, the equalization is equivalent to estimate the state vector $\mathbf{D}(k)$ from the observation of the channel output $Z^k = \{z(k), z(k-1), \ldots, z(0)\}$. The estimation of $d(k)$ can be obtained at some delayed time $(k-r)$, where $0 \leq r \leq (L-1)$.

One possible equalizer structure consists of a linear filter section followed by a nonlinear slicer or decision device. The linear filter is designed to minimize mean square error (MSE) between the filter output and

$$s(k) = d(k-r). \tag{2.8}$$

In order to derive the optimal transversal equalizer the MSE is minimized subject to the constraint that the impulse response is finite, causal, and stable. For a less rigorous condition (e.g., if the impulse response is causal and stable), the solution to the minimization of the problem is provided by infinite impulse response (IIR) Wiener filter [54].

The minimum phase spectral factorization of the power spectrum of the observed data is the major obstacle to the design of the IIR Wiener equalizing filter. It has been shown in [54]

that by using a Kalman filter, the spectral factorization problem can be solved indirectly. The Kalman equalizer and the solution of the IIR Wiener filter are equivalent, if all processes are stationary and measurement noise is white [5, 54]. An attractive feature of the Kalman equalizer is that it yields an on-line minimum variance estimate of the complete state vector $\mathbf{D}(k)$.

Kalman equalizer is full of promise in terms of performance and complexity if the plant noise, the observation noise, and the initial estimate of the state are Gaussian and mutually independent. The information sequence $\{d(k)\}$ is typically assumed to be an i.i.d. sequence. Since it can take only finitely many values, it is clearly non-Gaussian. Hence, the plant noise, $\mathbf{G}d(k)$ of (2.6) is not-Gaussian in the context of data channel equalization.

The Gaussian assumption for the plant noise impairs the performance of the Kalman equalizer due to the occurrence of model mismatch of the Kalman filter. Given all available received data, knowledge of the probability density function (pdf) of the state provides the most complete possible description of the state and from this pdf any of the common types of estimates (e.g., MMSE or MAP) can be determined. It is extremely difficult to determine this density function except in the linear Gaussian case. For data channel equalization, in [50], a weighted sum of Gaussian probability density functions is used to approximate arbitrarily closely another pdf that permits the explicit calculation of the *a posteriori* density from the Bayesian recursion relations illustrated in [3]. The idea of using a weighted sum of Gaussian density functions for approximation of the *a posteriori* density function has been suggested in [3]. Consequently, the solution to the minimum mean square error equalization problem gives rise to a linear channel equalizer which consists of the convex combination of the output of several Kalman filters operating in parallel. A detailed discussion can be found in [50].

These equalizers need the knowledge of the channel parameters or assumed to have been estimated earlier. When the channel is unknown and/or time-varying, which is the real-life situation, significant trouble will be encountered with these equalizers. To resolve this problem, the state vector of the Kalman filter (KF) is augmented with the coefficients of the unknown channels. The result of this formulation is a nonlinear estimation problem to which an extended Kalman filter (EKF) is applied instead of the Kalman filter. Since the estimated state is the concatenation of the estimations of the symbol vector and channel vector, it corresponds to the blind estimation approach.

2.1.3 Why Blind?

Blind techniques have been actively studied by numerous researchers over the past 20 years. Blind techniques estimate the channel and/or the signals based only on the channel output. On the other hand, non-blind techniques require the transmitter to periodically send signals that are known to the receiver in order to enable channel identification. Although the use of training sequence is probably the most robust way to estimate the channel, there are several reasons for studying blind algorithms.

- Bandwidth is conserved by eliminating or reducing training sets.
- Training is inefficient for rapidly time-varying channels.
- Several multipath fading during the training period can lead to poor channel estimates.
- Training for interference is often not accessible.
- Training requires synchronization, which may not be feasible in multi-user scenarios.
- Training signals are not available in military applications.

- In distributed networks, sending training signals each time a new communication link is to be set up may not be feasible.

Yet almost all current wireless cellular systems, such as the North American time division multiple access (TDMA) Digital Cellular System (IS-54) and the European GSM, embed training signals in the transmitted data. However, the channel may change within a time slot and blind algorithms can help track the time-varying channel. This scenario is likely to occur in IS-54 where each time slot is relatively long in duration. In GSM, 20% of the bandwidth is devoted to training and blind algorithms can be used to reduce the length of the training sequence. Blind methods are also more robust if the signal undergoes severe fading during training. These are some of the issues that make blind techniques worth research attention.

2.2 Space-Time Block Coded Systems

Recent advances in communications are driven by the requirements of next generation wireless systems with reliable high data rate transmission. Although the information capacity of wireless communication systems increases dramatically by employing multiple transmit and receive antennas, increasing the quality or reducing the effective error rate in a multipath fading channel is still a challenging task. STTC for transmit diversity reported in [65] is very effective and of good performance. Due to its extremely high decoding complexity, Alamouti's STBC scheme [2] appeared as a simple way to obtain transmit diversity. Alamouti's STBC has been adopted in several wireless applications due to its main attractive features including the following.

- It achieves full diversity as the well known maximal-ratio receiver combining (MRRC) at full transmission rate for any (real or complex) signal constellation.

- It does not require channel state information at the transmitter.
- It increases transmission rate with ML detection based on the linear processing at the receiver (due to the orthogonal code structure).
- It not only increases the effective data rate and system capacity, but also improves the error performance without any bandwidth expansion or any feedback from the receiver to the transmitter.

However, the main demerits of Alamouti's STBC scheme are

- Unlike STTC, it does not provide any coding gain.
- The decoding scheme is valid only for a flat-fading channel.

In STBC, a vector of symbols in the time domain is mapped to a matrix of symbols in the space-time domain. Three parameters (N, B, P) characterize a space-time block code. A symbol vector $\mathbf{s}$ of size $P \times 1$ turns into a space-time coded matrix $\mathcal{M}(\mathbf{s})$ of size $N \times B$, where N represents the number of transmit antennas and B is the number of coded symbols in the time. Thus the code rate[2] can be defined as

$$\zeta = \frac{P}{B} \tag{2.9}$$

with a full transmit diversity. We consider Alamouti's STBC for which $N = B = P = 2$. Let the kth symbol block be

$$\begin{aligned} \mathbf{s}(k) &\triangleq [s(Pk) \;\; s(Pk+1) \;\; \ldots \;\; s(Pk+P-1)]^T \\ &= [s(2k) \;\; s(2k+1)]^T \end{aligned} \tag{2.10}$$

[2]Code rate can be defined in different manners, if the full transmit diversity is not assumed.

where $s(k) \in \mathcal{S}$ is an element of the symbol alphabet $\mathcal{S}$. $s(k)$ is assumed to be a symbol sequence of i.i.d. complex random variables from $\mathcal{S}$ and $\mathbb{E}[s^2(k)] = 0$, where $\mathbb{E}[\cdot]$ is the statistical expectation.

Let $\mathbf{g}(k)$ denote the $N \times 1$ space-time block coded signal vector. The ith element of $\mathbf{g}(k)$, i.e., $g_i(k)$, is transmitted through ith transmit antenna. Then the space time block code of $(N, B, P) = (2, 2, 2)$ in [2] is defined by the mapping given by

$$\begin{aligned} [\mathbf{g}(2k) \;\; \mathbf{g}(2k+1)] &= \begin{bmatrix} g_1(2k) & g_1(2k+1) \\ g_2(2k) & g_2(2k+1) \end{bmatrix} \\ &= \mathcal{M}(\mathbf{s}(k)) \\ &= \begin{bmatrix} s(2k) & s(2k+1) \\ -s^*(2k+1) & s^*(2k) \end{bmatrix}. \end{aligned} \tag{2.11}$$

Another space time block code of $(N, B, P) = (4, 4, 3)$ in [20] is defined by the following mapping with a code rate of 3/4

$$\mathcal{M}\left([s_1 \;\; s_2 \;\; s_3]^T\right) = \begin{bmatrix} s_1 & -s_2^* & \frac{s_3^*}{\sqrt{2}} & \frac{s_3^*}{\sqrt{2}} \\ s_2 & s_1^* & \frac{s_3^*}{\sqrt{2}} & -\frac{s_3^*}{\sqrt{2}} \\ \frac{s_3}{\sqrt{2}} & \frac{s_3}{\sqrt{2}} & \frac{-s_1-s_1^*+s_2-s_2^*}{2} & \frac{s_2+s_2^*+s_1-s_1^*}{\sqrt{2}} \\ \frac{s_3}{\sqrt{2}} & -\frac{s_3}{\sqrt{2}} & \frac{-s_2-s_2^*+s_1-s_1^*}{2} & -\frac{s_1+s_1^*+s_2-s_2^*}{\sqrt{2}} \end{bmatrix} \tag{2.12}$$

where $s_i = s(3k + i - 1)$, $i = 1, 2, 3$.

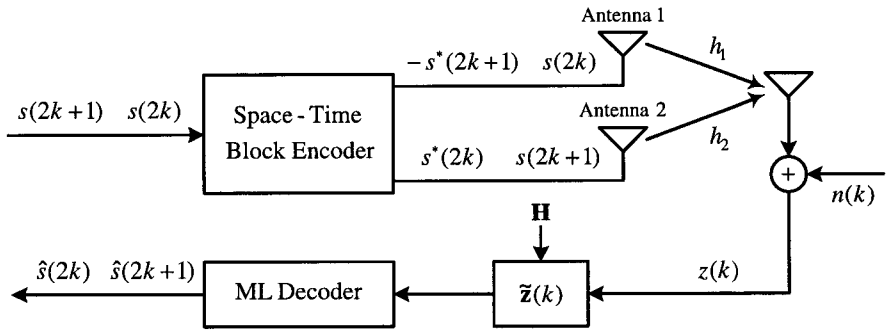

Figure 2.4: STBC system with 2 transmit antennas and 1 receive antenna.

2.2.1 STBC for Systems with One Receive Antenna

2.2.1.1 Coding

The baseband representation for STBC with two transmit antennas and one receive antenna is shown in Figure 2.4, where we consider Alamouti's space-time code mapping as shown in (2.11). Let h_i, $i = 1, 2$ represent the one-tap impulse response from the ith transmit antenna to the receive antenna. It is assumed that the channel is flat-fading channel where the channel gain is constant over two consecutive symbols, i.e., $h_i(2k) = h_i(2k+1)$.

During the time interval $2k$ and $2k+1$, we receive two consecutive samples $z(2k)$ and $z(2k+1)$ given by

$$z(2k) = h_1 s(2k) + h_2 s(2k+1) + n(2k) \tag{2.13}$$

$$z(2k+1) = -h_1 s^*(2k+1) + h_2 s^*(2k) + n(2k+1) \tag{2.14}$$

where $n(k)$ is the AWGN and modeled as i.i.d. complex Gaussian random variable with zero mean and power spectral density $N_0/2$ per dimension. Now, we define the following vectors

- The receive signal vector $\mathbf{z}(k) \triangleq [z(2k) \;\; z^*(2k+1)]^T$.

- The code symbol vector $\mathbf{s}(k) \triangleq [s(2k) \;\; s(2k+1)]^T$.
- The noise vector $\mathbf{n}(k) \triangleq [n(2k) \;\; n^*(2k+1)]^T$.

Hence, we have the following matrix/vector form

$$\mathbf{z}(k) = \mathbf{H}\mathbf{s}(k) + \mathbf{n}(k) \tag{2.15}$$

where the channel matrix $\mathbf{H}$ is given by

$$\mathbf{H} = \begin{bmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{bmatrix} \tag{2.16}$$

and $\mathbf{n}(k)$ is a complex random vector with zero mean and covariance $N_0\mathbf{I}_2$.

2.2.1.2 Decoding

If we define $\mathcal{C}$ as the code set for all possible symbol vectors $\mathbf{s}(k)$ and assume that the symbol pair is equally likely to occur, then the optimum ML decoder can be characterized [55] as follows

$$\hat{\mathbf{s}}(k) = \arg\left\{ \min_{\hat{\mathbf{s}}(k)\in C} \|\mathbf{z}(k) - \mathbf{H}\hat{\mathbf{s}}(k)\|^2 \right\}. \tag{2.17}$$

By realizing the fact that the channel matrix $\mathbf{H}$ is a orthogonal matrix, the ML decoding rule in (2.17) can be further simplified. Therefore, $\mathbf{H}^{\mathcal{H}}\mathbf{H} = \rho\mathbf{I}$ where $(\cdot)^{\mathcal{H}}$ denotes Hermitian transpose and $\rho = |h_1|^2 + |h_2|^2$. Pre-multiplying (2.15) by matrix $\mathbf{H}^{\mathcal{H}}$, we obtain

$$\tilde{\mathbf{z}}(k) = \begin{bmatrix} \tilde{z}(2k) \\ \tilde{z}(2k+1) \end{bmatrix} = \mathbf{H}^{\mathcal{H}}\mathbf{z}(k) = \rho\mathbf{s}(k) + \tilde{\mathbf{n}}(k) \tag{2.18}$$

where $\tilde{\mathbf{n}}(k) = \mathbf{H}^{\mathcal{H}}\mathbf{n}(k)$. Using (2.18), the ML decoding rule for this scenario can be reduced into two separate and simple decoding rules for $s(2k)$ and $s(2k+1)$, given by

$$\hat{s}(2k) = \arg\left\{\min_{\hat{s}(2k)\in S} \|\tilde{z}(2k) - \rho\hat{s}(2k)\|^2\right\} \tag{2.19}$$

$$\hat{s}(2k+1) = \arg\left\{\min_{\hat{s}(2k+1)\in S} \|\tilde{z}(2k+1) - \rho\hat{s}(2k+1)\|^2\right\} \tag{2.20}$$

where $\mathcal{S}$ denotes the code set for all possible symbols. Apparently, finding $\hat{s}(2k)$ and $\hat{s}(2k+1)$ requires much reduced computational complexity compared with that required by (2.17). In fact, for the above STBC, only two complex multiplications and one complex addition per symbol are required for decoding. A diversity gain of order 2 can be achieved by integrating this STBC scheme into the basic communication systems [55].

2.2.2 STBC for Systems with Multiple Receive Antennas

2.2.2.1 Coding

We consider for a system with two transmit antennas and M receive antennas. Let the received signal vector $\mathbf{z}^m(k)$ at the mth receive antenna is given by

$$\mathbf{z}^m(k) = \mathbf{H}^m\mathbf{s}(k) + \mathbf{n}^m(k) \tag{2.21}$$

where $\mathbf{n}^m(k)$ is the noise vector for the channel between the transmit antennas and the mth receive antenna and

$$\mathbf{H}^m = \begin{bmatrix} h_{1m} & h_{2m} \\ h_{2m}^* & -h_{1m}^* \end{bmatrix} \tag{2.22}$$

is the channel matrix for the same channel with h_{im}, $i = 1, 2$ denoting the channel impulse response between the ith transmit antenna and the mth receive antenna.

2.2.2.2 Decoding

In this case, the optimum ML decoding rule turns into

$$\hat{\mathbf{s}}(k) = \arg\left\{ \min_{\hat{\mathbf{s}}(k)\in C} \sum_{m=1}^{M} \|\mathbf{z}^m(k) - \mathbf{H}^m \hat{\mathbf{s}}(k)\|^2 \right\}. \tag{2.23}$$

As before, the pre-multiplication of (2.18) by $\mathbf{H}^{m\mathcal{H}}$ gives us

$$\tilde{\mathbf{z}}^m(k) = \begin{bmatrix} \tilde{z}^m(2k) \\ \tilde{z}^m(2k+1) \end{bmatrix} = \mathbf{H}^{m\mathcal{H}} \mathbf{z}^m(k) = \rho^m \mathbf{s}(k) + \tilde{\mathbf{n}}^m(k) \tag{2.24}$$

where $\rho^m = |h_{1m}|^2 + |h_{2m}|^2$ and $\tilde{\mathbf{n}}^m(k) = \mathbf{H}^{m\mathcal{H}} \mathbf{n}^m(k)$. Accordingly, the objective function in (2.23) can be reduced to

$$\begin{aligned} \sum_{m=1}^{M} \|\tilde{\mathbf{z}}^m(k) - \rho^m \hat{\mathbf{s}}(k)\|^2 &= \sum_{m=1}^{M} \{ \|\tilde{z}^m(2k) - \rho^m \hat{s}(2k)\|^2 \\ &\quad + \|\tilde{z}^m(2k+1) - \rho^m \hat{s}(2k+1)\|^2 \} \end{aligned} \tag{2.25}$$

and the ML decoding rule in (2.23) can be replaced by the following two separate and simple decoding rules given by

$$\hat{s}(2k) = \arg\left\{ \min_{\hat{s}(2k)\in S} \sum_{m=1}^{M} \|\tilde{z}^m(2k) - \rho^m \hat{s}(2k)\|^2 \right\} \tag{2.26}$$

$$\hat{s}(2k+1) = \arg\left\{ \min_{\hat{s}(2k+1)\in S} \|\tilde{z}^m(2k+1) - \rho^m \hat{s}(2k+1)\|^2 \right\}. \tag{2.27}$$

Under the assumption that the channel response for each propagation path are mutually independent, the diversity order provided by this scheme is $2M$ [55].

Chapter 3

The Multiple Model Estimation Approaches

In this Chapter, the Multiple Model (MM) estimators are reviewed. Since a single Kalman filter assumes a fixed model for the state evolution, it results in significant performance degradation in terms of estimation error when it is used to estimate the state of the systems whose models vary with time. In such a scenario, MM estimators can handle the potential source of model-mismatch. The MM estimators assume that the system behaves according to one of a finite number of models. The models can differ in noise levels or their structure. In this approach, a *Bayesian framework* is used: starting with prior probabilities of each model being correct (i.e., the system in a particular mode), the corresponding posterior probabilities are obtained.

A multiple model estimator can be *static* or *dynamic*. First the static case in which the model the system obeys is fixed, i.e., no switching from one model to another occurs during the estimation process (time-invariant mode) is considered. Then the optimal *dynamic* estimator which accounts for switching from one model to another model according to a Markov

chain is presented. Since the optimal dynamic estimator is not practical for implementation, two suboptimal approaches, one called generalized pseudo-Bayesian (GPB) and the other the interacting multiple model (IMM), are also presented.

3.1 The Static Multiple Model Estimator

A static multiple model is static from the point of view of the assumed evolution of the models. The model, assumed to be in effect throughout the entire process, is one of q possible models (i.e., the system is in one of q modes) given by

$$M \in \{M_j\}_{j=1}^{q}. \tag{3.1}$$

The prior probability that M_j is correct (i.e., the system is in mode j) is

$$\mu_j(0) = P\{M_j|Z^0\} \quad j = 1, \ldots, q \tag{3.2}$$

where Z^0 is the prior information and

$$\sum_{j=1}^{q} \mu_j(0) = 1 \tag{3.3}$$

since the correct model is among the assumed q possible models.

Using Bayes' formula, the posterior probability of model j being correct, given the measurements up to time k, $Z^k \triangleq \{z(j)\}_{j=1}^k$, is given by the recursion

$$\begin{aligned}\mu_j(k) &\triangleq P\{M_j|Z^k\} = P\{M_j|z(k), Z^{k-1}\} = \frac{p[z(k)|Z^{k-1}, M_j]P\{M_j|Z^{k-1}\}}{p[z(k)|Z^{k-1}]} \\ &= \frac{p[z(k)|Z^{k-1}, M_j]P\{M_j|Z^{k-1}\}}{\sum_{i=1}^q p[z(k)|Z^{k-1}, M_i]P\{M_i|Z^{k-1}\}} \\ &= \frac{p[z(k)|Z^{k-1}, M_j]\mu_j(k-1)}{\sum_{i=1}^q p[z(k)|Z^{k-1}, M_i]\mu_i(k-1)} \quad j = 1, \ldots, q \end{aligned} \tag{3.4}$$

starting with the given prior probabilities (3.2).

The term $p[z(k)|Z^{k-1}, M_j]$ in (3.4) is the likelihood function of mode j at time k, which, under linear-Gaussian assumptions, is given by

$$\Lambda_j(k) \triangleq p[z(k)|Z^{k-1}, M_j] = p[\nu_j(k)] = \mathcal{N}[\nu_j(k); 0, S_j(k)] \tag{3.5}$$

where ν_j and S_j are the innovation and its covariance from the mode matched filter corresponding to mode j, respectively. In a nonlinear and/or non-Gaussian situation, the same Gaussian likelihood functions are used.

Thus a Kalman filter[1] matched to each mode is set up yielding *mode-conditioned* state estimates and the associated *mode-conditioned* covariances. The probability of each mode being correct – the mode probability – is obtained according to (3.4) based on its likelihood function (3.5) relative to the other filters' likelihood functions.

A schematic diagram of the static multiple model estimator, which is a bank of two filters, is shown in Figure 3.1.

The outputs of each mode-matched filter are the mode conditioned state estimate $\hat{x}^j$, the associated covariance P^j and the mode likelihood function Λ_j.

[1]In a nonlinear situation the filters are extended Kalman filters instead of Kalman filters.

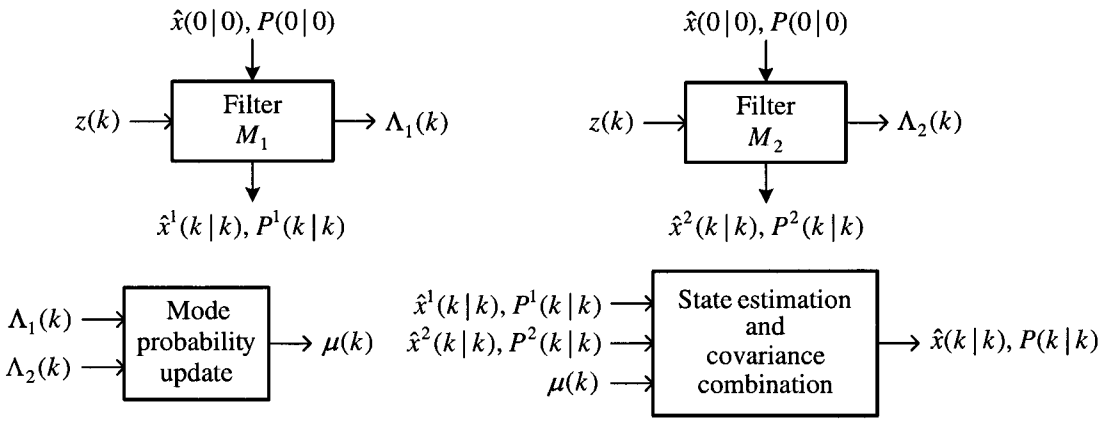

Figure 3.1: The static multiple model estimator for two models.

Once the filters are initialized, they run recursively *on their own estimates.* Their likelihood functions are used to update the mode probabilities. The latest mode probabilities are used to combine the mode-conditioned estimates and covariances.

Under the above assumptions, the pdf of the state of the system is a Gaussian mixture with q terms given by

$$p[x(k)|Z^k] = \sum_{j=1}^{q} \mu_j(k)\mathcal{N}[x(k); \hat{x}^j(k|k), P^j(k|k)]. \tag{3.6}$$

Therefore, the combination of the mode-conditioned estimates is obtained as follows

$$\hat{x}(k|k) = \sum_{j=1}^{q} \mu_j(k)\hat{x}^j(k|k) \tag{3.7}$$

and the error covariance matrix of the combined estimate is given by

$$P(k|k) = \sum_{j=1}^{q} \mu_j(k)\{P^j(k|k) + [\hat{x}^j(k|k) - \hat{x}(k|k)][\hat{x}^j(k|k) - \hat{x}(k|k)]^T\}. \tag{3.8}$$

The above discussion is valid under the following assumptions:

1. The correct model is among the set of models considered.

2. The same model has been in effect from the initial time.

Assumption 1 can be considered a reasonable approximation; however, assumption 2 is obviously not true if a maneuver has started at some time within the interval $[1, k]$, in which case a model change – *mode jump* – has occurred.

If the mode set includes the correct one and no mode jump occurs, then the probability of the true mode will converge to unity, i.e., this approach yields consistent estimates of the system parameters. Otherwise the probability of the model "nearest" to the correct one will converge to unity.

The following ad hoc modification can be made to the static MM estimator for estimating the state in the case of switching models: an artificial lower bound is imposed on the model probabilities (with a suitable renormalization of the remaining probabilities).

A shortcoming of the static MM estimator when used with switching models is that, in spite of the above ad hoc modification, the mismatched filters' errors can grow to unacceptable levels. Thus, reinitialization of the filters that are mismatched is, in general, needed. This is accomplished by using the estimate from filter corresponding to the best matched model in the other filters.

It should be pointed out that the above "fixes" are automatically (and rigorously) built into the dynamic MM estimation algorithms to be discussed in the subsequent sections. First the optimal dynamic MM estimator is described. Since this is not practical for real-time processing, sub-optimal dynamic MM estimators have to be considered. Three of them are described in Section 3.3.

3.2 The Dynamic Multiple Model Estimator

In this case, the mode the system is in can undergo switching in time. The dynamics and measurements of such a system are assumed linear and modeled as

$$x(k) = F[M(k)]x(k-1) + v[k-1, M(k)] \tag{3.9}$$

$$z(k) = H[M(k)]x(k) + w[k, M(k)] \tag{3.10}$$

where $x(k)$ is the system state at time k, $z(k)$ is the measurement at k, and $F[M(k)]$ and $H[M(k)]$ are the parameters describing system structure when the mode or model $M(k)$ is in effect during the sampling period ending at k (i.e., the sampling period $(k-1, k]$). $v[\cdot]$ and $w[\cdot]$ are the process noise and measurement noise, respectively. Such systems are also called jump-linear systems. This approach can also be used for nonlinear systems via linearization. The mode at time k is assumed to be among the possible q modes

$$M(k) \in \{M_j\}_{j=1}^{q}. \tag{3.11}$$

The structure of the system and/or the statistics of noises might be different from model to model. For example,

$$F[M_j] = F_j \tag{3.12}$$

$$v(k-1, M_j) \sim \mathcal{N}[u_j, Q_j]. \tag{3.13}$$

The optimal approach to filtering the sate of the system in (3.9) and (3.10) requires that every possible sequence of models from the initial observation through the most recent measurement be considered. Thus for q models, optimal filtering requires q^k filters for processing

Table 3.1: Mode histories for $q = 2$ models at time $k = 2$.

l	$i_{1,l}$	$i_{2,l}$
1	1	1
2	1	2
3	2	1
4	2	2

the kth observation $z(k)$ and estimating the state $x(k)$.

The lth mode history (sequence of models) through time k is denoted as

$$M^{k,l} = \{M_{i_{1,l}}, \ldots, M_{i_{k,l}}\} \quad l = 1, \ldots, q^k \tag{3.14}$$

where $i_{\kappa,l}$ is the model index at time κ from history l and

$$1 \leq i_{\kappa,l} \leq q \quad \kappa = 1, \ldots, k. \tag{3.15}$$

Note that the number of histories increases *exponentially with time.* For example, with $q = 2$ at time $k = 2$ one has the $q^k = 4$ possible sequences (histories) as presented in Table 3.1.

It is assumed that the mode switching is a Markov process (Markov chain) with known mode transition probabilities given by

$$\pi_{ij} \triangleq P\{M(k) = M_j | M(k-1) = M_i\}. \tag{3.16}$$

These mode transition probabilities will be assumed time-invariant and independent of the base state $x(k)$.

The event that model j is in effect at time k is denoted as

$$M_j(k) \triangleq \{M(k) = M_j\}. \tag{3.17}$$

The conditional probability of the lth history

$$\mu^{k,l} \triangleq P\{M^{k,l}|Z^k\} \tag{3.18}$$

is evaluated next.

The lth sequence of models through time k can be written as

$$M^{k,l} = \{M^{k-1,s}, M_j(k)\} \tag{3.19}$$

where sequence s through $k-1$ is its *parent sequence* and M_j is its last element.

Then, in view of the Markov property,

$$\pi_{ij} \triangleq P\{M_j(k)|M^{k-1,s}\} = P\{M_j(k)|M_i(k-1)\} \tag{3.20}$$

where $i = s_{k-1}$, the index of the last model in the parent sequence s through $k-1$.

The conditional pdf of the state at k is obtained using the total probability theorem with respect to the mutually exclusive and exhaustive set of events (3.14), as a Gaussian mixture with an exponentially increasing number of terms

$$p[x(k)|Z^k] = \sum_{l=1}^{q^k} p[x(k)|M^{k,l}, Z^k] P\{M^{k,l}|Z^k\}. \tag{3.21}$$

Since to each mode sequence one has to match a filter, it can be seen that an exponentially increasing number of filters are needed to estimate the (base) state, which makes the optimal

approach impractical.

The probability of a mode history is obtained using Bayes' formula as

$$\begin{aligned}
\mu^{k,l} &= P\{M^{k,l}|Z^k\} = P\{M^{k,l}|z(k), Z^{k-1}\} \\
&= \frac{1}{c}\, p[z(k)|M^{k,l}, Z^{k-1}]P\{M^{k,l}|Z^{k-1}\} \\
&= \frac{1}{c}\, p[z(k)|M^{k,l}, Z^{k-1}]P\{M_j(k), M^{k-1,s}|Z^{k-1}\} \\
&= \frac{1}{c}\, p[z(k)|M^{k,l}, Z^{k-1}]P\{M_j(k)|M^{k-1,s}, Z^{k-1}\}\mu^{k-1,s} \\
&= \frac{1}{c}\, p[z(k)|M^{k,l}, Z^{k-1}]P\{M_j(k)|M^{k-1,s}\}\mu^{k-1,s}
\end{aligned} \tag{3.22}$$

where c is the normalization constant.

If the current mode depends only on the previous one (i.e., it is a Markov chain), then

$$\begin{aligned}
\mu^{k,l} &= \frac{1}{c}\, p[z(k)|M^{k,l}, Z^{k-1}]P\{M_j(k)|M_i(k-1)\}\mu^{k-1,s} \\
&= \frac{1}{c}\, p[z(k)|M^{k,l}, Z^{k-1}]\pi_{ij}\mu^{k-1,s}
\end{aligned} \tag{3.23}$$

where $i = s_{k-1}$ is the last model of the parent sequence s.

Equation (3.23) shows that *conditioning on the entire past history* is needed even if the random parameters are Markov.

3.3 Practical Algorithms

Since the optimal approach involves an exponentially increasing number of hypotheses as the number of observation increases, the optimal approach is not practical for real-time processing and suboptimal techniques have to be considered for the efficient management of the multiple hypotheses.

One approach for managing the larger number of hypotheses in the filtering problem is the generalized pseudo-Bayesian (GPB) algorithms [1, 10, 18, 33]. In the GPB algorithms, the hypotheses that differ only in the *older* models are combined. The first order GPB algorithm, denoted as GPB1, is an approximation that considers only the possible models over the most recent sampling period. The second-order GPB algorithm, denoted as GPB2, is an approximation that considers only the possible models over the two most recent sampling periods. Finally, the IMM estimation algorithm is discussed.

3.3.1 The GPB1 Estimator

The GPB1 estimator runs q filters, conditioned on each possible state model, in parallel. At the end of each cycle of the algorithm, the estimates from these q filters are combined with weights $\mu_j(k)$. In other words, *the q hypotheses are merged into a single hypothesis at the end of each cycle.* The weights $\mu_j(k)$ are calculated based on the likelihood functions of these q filters, during each cycle. After the filters are initialized, they run recursively using the previous combined estimate.

3.3.1.1 The Algorithm

One cycle of the algorithm as shown in Figure 3.2 can be explained by the following steps.

1. ***Mode-matched filtering*** $(j = 1, \ldots, q)$: Starting with $\hat{x}(k-1|k-1)$, one computes $\hat{x}^j(k|k)$ and the associated error covariance $P^j(k|k)$ through a filter matched to $M_j(k)$. The likelihood functions

$$\Lambda_j(k) = p[z(k)|M_j(k), Z^{k-1}] \tag{3.24}$$

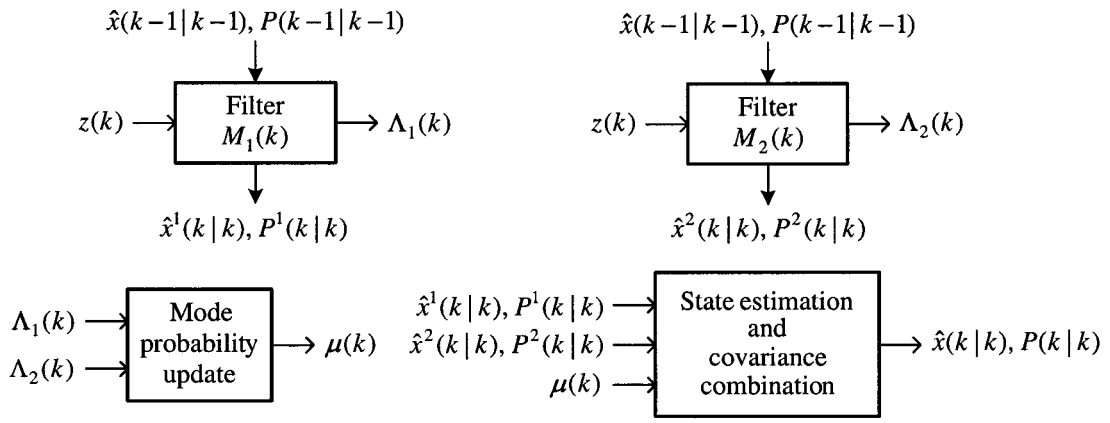

Figure 3.2: One cycle of the GPB1 MM estimator for two switching models.

corresponding to these q filters are evaluated as

$$\Lambda_j(k) = p[z(k)|M_j(k), \hat{x}(k-1|k-1), P(k-1|k-1)]. \tag{3.25}$$

2. ***Mode probability update***: This probability that the mode j $(j = 1, \ldots, q)$ is in effect is updated as follows

$$\begin{aligned} \mu_j(k) &\triangleq P\{M_j(k)|Z^k\} = P\{M_j(k)|z(k), Z^{k-1}\} \\ &= \frac{1}{c}\, p[z(k)|M_j(k), Z^{k-1}]P\{M_j(k)|Z^{k-1}\} \\ &= \frac{1}{c}\,\Lambda_j(k)\sum_{i=1}^{q} P\{M_j(k)|M_i(k-1), Z^{k-1}\}P\{M_i(k-1)|Z^{k-1}\} \end{aligned} \tag{3.26}$$

which yields with π_{ij} the known *mode transition probabilities*

$$\mu_j(k) = \frac{1}{c}\,\Lambda_j(k)\sum_{i=1}^{q}\pi_{ij}\mu_i(k-1) \tag{3.27}$$

where c is the normalization constant

$$c = \sum_{j=1}^{q} \Lambda_j(k) \sum_{i=1}^{q} \pi_{ij}\mu_i(k-1). \tag{3.28}$$

3. ***State estimate and covariance combination***: Finally, combination of the mode-conditioned estimates and covariances are done using the following mixture equations

$$\hat{x}(k|k) = \sum_{j=1}^{q} \hat{x}^j(k|k)\mu_j(k) \tag{3.29}$$

$$P(k|k) = \sum_{j=1}^{q} \mu_j(k)\{P^j(k|k) + [\hat{x}^j(k|k) - \hat{x}(k|k)][\hat{x}^j(k|k) - \hat{x}(k|k)]^T\}. \tag{3.30}$$

3.3.2 The GPB2 Estimator

If the system obeys one of q possible modes, the GPB2 runs q^2 filters simultaneously. There are q filters conditioned on each mode. Estimates from these q filters are merged. The q merged estimates are combined to obtain the latest state estimate. Also the merged estimate corresponding to a particular filter is the input to one of the q filters corresponding to each of the q modes.

3.3.2.1 The Algorithm

One cycle of the algorithm as shown in Figure 3.3 can be explained by the following steps.

1. ***Mode-matched filtering*** $(i, j = 1, \ldots, q)$: Starting with $\hat{x}^i(k-1|k-1)$, one computes $\hat{x}^{ij}(k|k)$ and the associated covariance $P^{ij}(k|k)$ through a filter matched to $M_j(k)$. The likelihood functions corresponding to these q^2 filters

$$\Lambda_{ij}(k) = p[z(k)|M_j(k), M_i(k-1), Z^{k-1}] \tag{3.31}$$

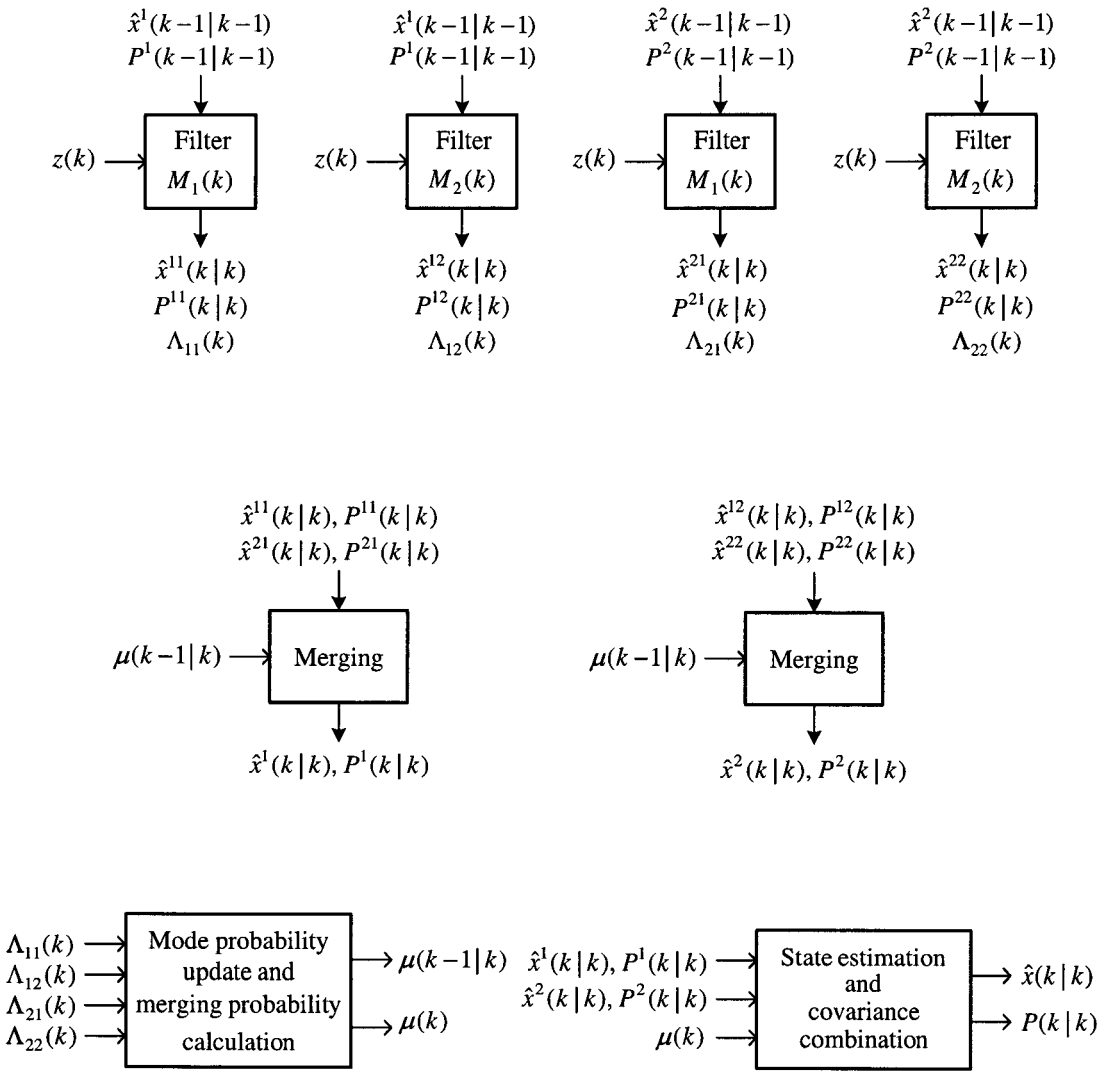

Figure 3.3: One cycle of the GPB2 MM estimator for two models.

are evaluated as

$$\Lambda_{ij}(k) = p[z(k)|M_j(k), \hat{x}^i(k-1|k-1), P^i(k-1|k-1)] \quad i,j = 1,\ldots,q. \tag{3.32}$$

2. ***Calculation of the merging probabilities*** $(i,j = 1,\ldots,q)$: The probability that mode i was in effect at $k-1$ if mode j is in effect at k conditioned on Z^k is

$$\begin{aligned}
\mu_{i|j}(k-1|k) &\triangleq P\{M_i(k-1)|M_j(k), Z^k\} = P\{M_i(k-1)|z(k), M_j(k), Z^{k-1}\} \\
&= \frac{1}{c_j} P[z(k), M_j(k)|M_i(k-1), Z^{k-1}] P\{M_i(k-1)|Z^{k-1}\} \\
&= \frac{1}{c_j} p[z(k)|M_j(k), M_i(k-1), Z^{k-1}] P\{M_j(k)|M_i(k-1), Z^{k-1}\} \\
&\quad \cdot P\{M_i(k-1)|Z^{k-1}\}
\end{aligned} \tag{3.33}$$

where $P[\cdot]$ denotes a mixed pdf-probability. Thus, the merging probabilities are

$$\mu_{i|j}(k-1|k) = \frac{1}{c_j} \Lambda_{ij}(k) \pi_{ij} \mu_i(k-1) \quad i,j = 1,\ldots,q \tag{3.34}$$

where

$$c_j = \sum_{i=1}^{q} \Lambda_{ij}(k) \pi_{ij} \mu_i(k-1). \tag{3.35}$$

The mode transition probabilities π_{ij} are assumed to be known — their selection is part of the algorithm design process.

3. ***Merging*** $(j = 1,\ldots,q)$: The state estimate corresponding to $M_j(k)$ is obtained by

combining the estimates as follows

$$\hat{x}^j(k|k) = \sum_{i=1}^{q} \hat{x}^{ij}(k|k)\mu_{i|j}(k-1|k) \quad j = 1, \ldots, q. \tag{3.36}$$

The corresponding covariance is

$$P^j(k|k) = \sum_{i=1}^{q} \mu_{i|j}(k-1|k)\{P^{ij}(k|k) + [\hat{x}^{ij}(k|k) - \hat{x}^j(k|k)][\hat{x}^{ij}(k|k) - \hat{x}^j(k|k)]^T\} \tag{3.37}$$

4. ***Mode probability updating*** $(j = 1, \ldots, q)$: This is done as follows

$$\begin{aligned}
\mu_j(k) &\triangleq P\{M_j(k)|z(k), Z^{k-1}\} = \frac{1}{c} P[z(k), M_j(k)|Z^{k-1}] \\
&= \frac{1}{c} \sum_{i=1}^{q} P[z(k), M_j(k)|M_i(k-1), Z^{k-1}]P\{M_i(k-1)|Z^{k-1}\} \\
&= \frac{1}{c} \sum_{i=1}^{q} p(z(k)|M_j(k), M_i(k-1), Z^{k-1}) \\
&\quad \cdot P\{M_j(k)|M_i(k-1), Z^{k-1}\}\mu_i(k-1).
\end{aligned} \tag{3.38}$$

Hence, the updated mode probabilities are

$$\mu_j(k) = \frac{1}{c} \sum_{i=1}^{q} \Lambda_{ij}(k)\pi_{ij}\mu_i(k-1) = \frac{c_j}{c} \quad j = 1, \ldots, q \tag{3.39}$$

where c_j is the expression from (3.36) and c is the normalization constant

$$c = \sum_{j=1}^{q} c_j. \tag{3.40}$$

5. ***State estimate and covariance combination***: The latest state estimate and

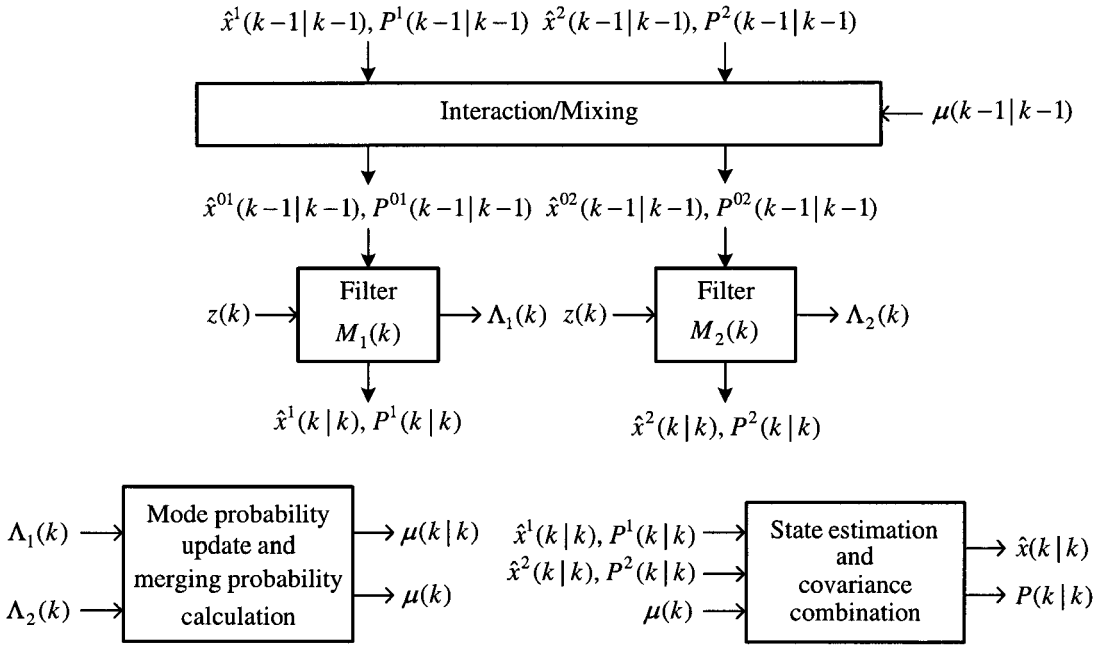

Figure 3.4: One cycle of the IMM estimator for two models.

covariance *for output only* are obtained according to

$$\hat{x}(k|k) = \sum_{j=1}^{q} \hat{x}^j(k|k)\mu_j(k) \tag{3.41}$$

$$P(k|k) = \sum_{j=1}^{q} \mu_j(k)\{P^j(k|k) + [\hat{x}^j(k|k) - \hat{x}(k|k)][\hat{x}^j(k|k) - \hat{x}(k|k)]^T\}. \tag{3.42}$$

3.3.3 The IMM Estimator

The IMM estimator is traditionally used in the field of multiple target tracking especially when the system can be described by a bank of multiple state space models [6]. Beginning with [9, 14, 16], the IMM estimator has been shown to handle both maneuvering and

non-maneuvering targets very effectively. The use of IMM algorithm for the problem of target tracking with glint noise is considered in [22]. The suboptimal approaches to the fixed-interval smoothing problems for Markovian switching systems using IMM estimator are presented in [33,34]. Recently, the application of IMM estimator has drawn much attention to solve problems in areas other than the target tracking. In [43], an efficient recursive algorithm based on the IMM estimator for enhancing speech signal degraded by additive white or colored noise is developed. The compensation of slowly evolving environment processes in speech recognition is presented in [44]. In [48], localization of mobile phones using IMM estimator in CDMA environment is addressed. The feasibility of employing the IMM estimator for CDMA multiuser detector is studied in [8]. Further, it is employed in [60] to develop a secure chaotic communication system. As the IMM algorithm handles the interactions between parallel filters in an efficient way, the performance of the system having model-mismatch is improved without much increase in computational complexity. A concise presentation of the IMM estimation algorithm developed in [15] is illustrated below.

3.3.3.1 The Algorithm

The IMM approach computes the state estimates that accounts for each possible current model using a suitable mixing of the previous model-conditioned estimates depending on the current model. Figure 3.4 describes this algorithm, which consists of two interacting filters operating in parallel. One cycle of IMM estimator is summarized below.

1. ***Calculation of the mixing probabilities*** $(i, j = 1, \ldots, q)$: This is the probability that mode M_i was in effect at $k-1$ given that M_j is in effect at k conditioned on Z^{k-1}

and given by

$$\begin{aligned}\mu_{i|j}(k-1|k-1) &\triangleq P\{M_i(k-1)|M_j(k), Z^{k-1}\} \\ &= \frac{1}{\bar{c}_j} P\{M_j(k)|M_i(k-1), Z^{k-1}\} P\{M_i(k-1)|Z^{k-1}\} \\ &= \frac{1}{\bar{c}_j} \pi_{ij}\mu_i(k-1) \quad i,j = 1,\ldots,q \end{aligned} \tag{3.43}$$

where the normalizing constants are

$$\bar{c}_j = \sum_{i=1}^{q} \pi_{ij}\mu_i(k-1) \quad j = 1,\ldots,q. \tag{3.44}$$

Note that the conditioning in (3.43) is Z^{k-1} whereas the conditioning in (3.33) is Z^k. This is what makes it possible to carry out the mixing at the *beginning* of the cycle, rather than the standard merging at the *end* of the cycle.

2. ***Mixing*** $(j = 1,\ldots,q)$: With the mixing probabilities as weights, estimates of all the filters at time $k-1$ are mixed to produce the initial estimates for each filter. The mixed initial condition for the filter matched to $M_j(k)$ is given by

$$\hat{x}^{0j}(k-1|k-1) = \sum_{i=1}^{q} \hat{x}^i(k-1|k-1)\mu_{i|j}(k-1|k-1) \quad j = 1,\ldots,q \tag{3.45}$$

where $\hat{x}^i(k-1|k-1)$ is the estimate of the filter matched to model M_i at time $k-1$ and the associated covariance is given by

$$\begin{aligned} P^{0j}(k-1|k-1) &= \sum_{i=1}^{q} \mu_{i|j}(k-1|k-1)\Big\{P^i(k-1|k-1) \\ &\quad + [\hat{x}^i(k-1|k-1) - \hat{x}^{0j}(k-1|k-1)] \\ &\quad \cdot [\hat{x}^i(k-1|k-1) - \hat{x}^{0j}(k-1|k-1)]^T\Big\} \quad j = 1,\ldots,q \end{aligned} \tag{3.46}$$

3. ***Mode-matched filtering*** $(j = 1, \ldots, q)$: The estimate (3.45) and covariance (3.46) are used as input to the filter matched to $M_j(k)$, which uses $z(k)$ to yield $\hat{x}^j(k|k)$ and $P^j(k|k)$.

   The likelihood functions corresponding to the q filters

$$\Lambda_j(k) = p[z(k)|M_j(k), Z^{k-1}] \tag{3.47}$$

   are computed by the mixed initial condition (3.45) and the associated covariance (3.46) as

$$\Lambda_j(k) = p[z(k)|M_j(k), \hat{x}^{0j}(k-1|k-1), P^{0j}(k-1|k-1)] \quad j = 1, \ldots, q. \tag{3.48}$$

4. ***Mode probability update*** $(j = 1, \ldots, q)$: The probability that the mode j is in effect is updated as follows

$$\begin{aligned} \mu_j(k) &\triangleq P\{M_j(k)|Z^k\} = \frac{1}{c}\, p[z(k)|M_j(k), Z^{k-1}] P\{M_j(k)|Z^{k-1}\} \\ &= \frac{1}{c}\, \Lambda_j(k) \sum_{i=1}^{q} P\{M_j(k)|M_i(k-1), Z^{k-1}\} P\{M_i(k-1)|Z^{k-1}\} \\ &= \frac{1}{c}\, \Lambda_j(k) \sum_{i=1}^{q} \pi_{ij}\mu_i(k-1) = \frac{1}{c}\, \Lambda_j(k)\bar{c}_j \quad j = 1, \ldots, q \end{aligned} \tag{3.49}$$

   where the normalization constant c is given by

$$c = \sum_{j=1}^{q} \Lambda_j(k)\bar{c}_j. \tag{3.50}$$

5. ***State estimate and covariance combination***: Finally, combination of the mode-conditioned estimates and covariances are done using the following mixture equations

$$\hat{x}(k|k) = \sum_{j=1}^{q} \hat{x}^j(k|k)\mu_j(k) \tag{3.51}$$

$$P(k|k) = \sum_{j=1}^{q} \mu_j(k)\{P^j(k|k) + [\hat{x}^j(k|k) - \hat{x}(k|k)][\hat{x}^j(k|k) - \hat{x}(k|k)]^T\}. \tag{3.52}$$

This combination is not part of the algorithm recursions whereas it is used only for output purposes.

It is important to note that even though the IMM estimator does not make a hard decision, that is it does not give unity probability for the mode that is active, but the mode probability corresponding to the active mode will be the highest. Hence, based on the mode probabilities one can make the decision of which model is being active at a particular time.

3.4 Conclusions

This Chapter has introduced multiple model approaches which assume the system to be in one of a finite number of modes. First the static case in which the model the system obeys is fixed, that is, no switching from one mode to another occurs during the estimation process. While the model that is in effect stays fixed, each model has its own dynamics, so the overall estimator is dynamic. The static multiple model estimator is useful if there is an ambiguity in the system model, but it does not switch. The typical ad hoc modification of the static model approach to handle switching models is to impose a lower bound on the probability of each model.

For systems that undergo changes in their mode during their operation, one can obtain

Table 3.2: Comparison of complexities of the MM algorithms.

	Static	GPB1	GPB2	IMM
Number of filters	q	q	q^2	q
Number of combinations of q estimates and covariances	1	1	$q+1$	$q+1$
Number of probability calculations	q	q	q^2+q	q^2+q

the optimal multiple model estimator which, however, consists of an exponentially increasing number of filters. Thus, suboptimal algorithms are necessary for the (realistic) mode transition situation.

The GPB1 MM approach computes the state estimate accounting for each possible current model. On the other hand, the GPB2 MM approach computes the state estimate accounting for

- each possible current model;
- each possible model of the previous time.

The interacting multiple model estimator computes the state estimate that accounts for each possible current model using a suitable mixing of the previous model-conditioned estimates depending on the current model. A comparison of the complexities of different MM approaches are presented in Table 3.2. As can be seen from above, the static algorithm has the same requirements as the GPB1. The IMM has only slightly higher requirements than the GPB1, but clearly significantly lower than GPB2. In view of this, the modifications of the static algorithm for the switching situation are considered obsolete.

From the point that the IMM performs significantly better than GPB1 and almost as well as GPB2, the IMM is considered to be the best compromise between complexity and performance.

Chapter 4

Data Channel Equalization

In this Chapter, our main concern is the problem of detection of transmitted digital data in presence of ISI and additive noise. We assume that the continuous time received signals are sampled at the baud rate (or symbol rate) after some processing (e.g., matched filtering). This results a discrete time model of the channel.

4.1 Problem Formulation

4.1.1 Channel Model

This Chapter considers the transmission of digital data over a baseband channel. A simplified block diagram of the discrete-time model of a digital communication system is depicted in Figure 4.1. The baseband equivalent channel of interest is composed of the transmitter, the physical channel, and the receiver. The channel is the physical media that connects the transmitter and the receiver. Examples of common physical channels include coaxial, fiber optic, or twisted-pair cable in wired communications and the atmosphere or ocean in wireless communications. At high enough data rates, all physical channels tend to exhibit ISI.

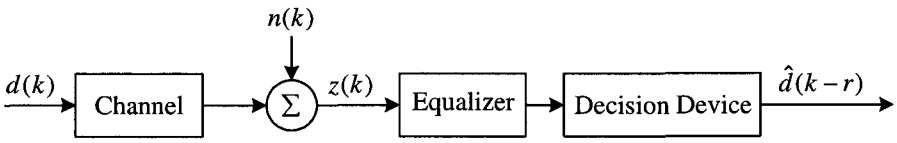

Figure 4.1: Discrete-time model of data transmission system.

Let $d(k)$ denote the symbol emitted by the digital source at time kT, where T is the symbol duration. The discrete time signal is modulated, filtered, sent through the communication channel, filtered and demodulated. The resulting continuous signal is given by

$$z(t) = \sum_{l=0}^{L-1} d(t - lT)\ h_l + n(t) \tag{4.1}$$

where T is the symbol period, $n(t)$ is the additive white noise independent from the emitted source signals, h_l is the composite time-invariant channel response encompassing the effects of the transmitting filter, reception filter, channel response and modulation/demodulation (which is assumed to be linear). The composite channel with a duration of approximately LT is assumed to be FIR. We consider symbol rate sampling[1] which results in an equivalent discrete time representation given by

$$z(k) = \sum_{l=0}^{L-1} d(k - l)\ h_l + n(k) \tag{4.2}$$

$$= \mathbf{HD}(k) + n(k) \tag{4.3}$$

where $z(k)$ is the received signal at time instant kT, $\mathbf{H} = [h_0, h_1, \ldots, h_{L-1}]$ is the parameter vector of the coefficients of the FIR channel impulse response, and $\mathbf{D}(k) = [d(k), d(k-1), \ldots, d(k-L+1)]^T$ consists of the L last transmitted symbols. The transmitted sequence

[1]This can easily be extended to multi rate sampling.

at the channel input $\{d(k)\}$ is composed of i.i.d. symbols from a finite alphabet $\gamma = \{d_i, i = 1, \ldots, q\}$ having variance σ_d^2 that is specific to the type of modulation and $n(k)$ represents a zero-mean AWGN sequence with variance σ_n^2. It is assumed that the noise sequence $n(k)$ is independent of the input symbol sequence $d(k)$. Although this model is suboptimal since the colored noise components introduced by the matched filter in the receiver are ignored, it is of interest in low noise applications where ISI is the limiting factor.

It is required to obtain the unbiased linear minimum error variance estimator of each transmitted symbol. To solve this problem, we make use of the state-space representation of stochastic processes and of results derived from the discrete KF theory.

4.1.2 State-Space Representation of the System

For convenience, the baseband channel output $z(k)$ is represented by the following state-space model

$$\mathbf{D}(k+1) = \mathbf{F}\mathbf{D}(k) + \mathbf{G}d(k+1) \tag{4.4}$$

$$z(k) = \mathbf{H}\mathbf{D}(k) + n(k) \tag{4.5}$$

where $\mathbf{F}$ is the $L \times L$ one-step transition matrix and $\mathbf{G}$ is the $L \times 1$ vector given by

$$\mathbf{F} = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 & 0 \\ 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 \end{bmatrix} \quad \mathbf{G} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{4.6}$$

Equations (4.4) and (4.5) which completely define the behavior of the channel allow us to interpret the channel as a linear state system. Equation (4.4) is the state transition equation where $\mathbf{D}(k)$ is the state vector whereas (4.5) represents the observation equation. We define the state or plant noise $\mathbf{u}(k) = \mathbf{G}d(k)$ in (4.4) and $n(k)$ as the observation noise in (4.5).

Lemma 1: The dynamic system defined by (4.4) and (4.5) is uniformly completely controllable and observable if $h_{L-1} \neq 0$.

Proof: A discrete time (deterministic) system is *completely controllable* if, given an *arbitrary destination point* in the state space, there is an input sequence that will bring the system from any initial state to this point in a finite number of steps. On the other hand, a system is completely observable if its initial state can be *fully and uniquely* recovered from a finite number of observations of its output and the knowledge of its input.

For a time-invariant discrete system described by the state-space model (4.4) and (4.5) is uniformly completely controllable and observable, if the pair $\{\mathbf{F}, \mathbf{G}\}$ is controllable, i.e., the *controllability matrix*

$$\mathcal{Q}_C \triangleq [\mathbf{G} \quad \mathbf{FG} \quad \mathbf{F}^2\mathbf{G} \quad \ldots \quad \mathbf{F}^{L-1}\mathbf{G}] \tag{4.7}$$

and the the pair $\{\mathbf{F}, \mathbf{H}\}$ is observable, i.e., the *observability matrix*

$$\mathcal{Q}_O \triangleq [\mathbf{H}^T \quad \mathbf{F}^T\mathbf{H}^T \quad (\mathbf{F}^T)^2\mathbf{H}^T \quad \ldots \quad (\mathbf{F}^T)^{L-1}\mathbf{H}^T] \tag{4.8}$$

have rank equal to the order of the system L [10].

For the state-space model represented by (4.4) and (4.5)

$$\mathbf{G} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \mathbf{FG} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \dots \quad \mathbf{F}^{L-1}\mathbf{G} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}. \tag{4.9}$$

Therefore, the controllability matrix $\mathcal{Q}_C$ is a $L \times L$ identity matrix and has rank equal to L. In addition,

$$\mathbf{H}^T = \begin{bmatrix} h_0 \\ h_1 \\ h_2 \\ \vdots \\ h_{L-1} \end{bmatrix} \quad \mathbf{F}^T\mathbf{H}^T = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ \vdots \\ 0 \end{bmatrix} \quad \dots \quad (\mathbf{F}^T)^{L-1}\mathbf{H}^T = \begin{bmatrix} h_{L-1} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{4.10}$$

Hence, the observability matrix

$$\mathcal{Q}_O = \begin{bmatrix} h_0 & h_1 & h_2 & \dots & h_{L-1} \\ h_1 & h_2 & h_3 & \dots & 0 \\ h_2 & h_3 & h_4 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ h_{L-1} & 0 & 0 & \dots & 0 \end{bmatrix} \tag{4.11}$$

has full rank L if $h_{L-1} \neq 0$.

4.1.3 The Kalman Observer

In system theory, the objective of an observer is to estimate the actual state vector using the knowledge of past inputs and the actual output. For the system defined by (4.4) and (4.5), the observer then estimates the state vector $\mathbf{D}(k)$ which contains the channel inputs from time k to $k-L+1$ from the observation sequences of the channel output up to and including time k, $Z^k = \{z(k), z(k-1), \ldots, z(0)\}$. Followed by a decision device as shown in Figure 4.1 to recover the input data symbols, the observer thus performs the equalization of the channel.

Since the output of the observer is an estimate of the last L inputs of the channel, the equalizer can recover the input channel at some delayed time $(k-r)$ where $0 \leq r \leq (L-1)$. This problem has been resolved in [47, 54] by applying Kalman observer as depicted in Figure 4.2 to the system (4.4) and (4.5) leading to the following equations.

Prediction:

$$\hat{\mathbf{D}}(k|k-1) = \mathbf{F}\hat{\mathbf{D}}(k-1|k-1) \tag{4.12}$$

$$\mathbf{P}(k|k-1) = \mathbf{F}\hat{\mathbf{P}}(k-1|k-1)\mathbf{F}^T + \mathbf{G}\mathbf{G}^T\sigma_d^2 \tag{4.13}$$

Estimation:

$$\mathbf{K}(k) = \frac{\mathbf{P}(k|k-1)\mathbf{H}^T}{\mathbf{H}\mathbf{P}(k|k-1)\mathbf{H}^T + \sigma_n^2} \tag{4.14}$$

$$\hat{\mathbf{D}}(k|k) = \hat{\mathbf{D}}(k|k-1) + \mathbf{K}(k)\big[z(k) - \mathbf{H}\hat{\mathbf{D}}(k|k-1)\big] \tag{4.15}$$

$$\mathbf{P}(k|k) = \mathbf{P}(k|k-1) - \mathbf{K}(k)\mathbf{H}\mathbf{P}(k|k-1) \tag{4.16}$$

where $\hat{\mathbf{D}}(k|k-1)$ and $\hat{\mathbf{D}}(k|k)$ are the predicted and estimated values of the state vector $\mathbf{D}(k)$, whereas $\mathbf{P}(k|k-1)$ and $\mathbf{P}(k|k)$ are the covariance matrices of the associated errors.

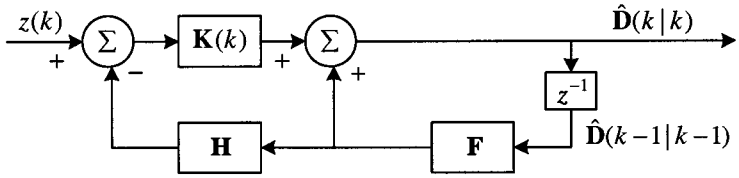

Figure 4.2: The Kalman observer.

$\mathbf{K}(k)$ is the Kalman gain.

The decision device performs an r-delayed estimation of the channel input $\hat{d}(k-r)$. It has been shown in [54] that the choice of $r = 0$ is sufficient for minimum phase channels whereas $r \neq 0$ may be interesting for non-minimum phase channels. For the non-minimum phase channels, if $r > (L-1)$, the state vector $\mathbf{D}(k)$ is merely augmented to contain $(r+1)$ elements while the parameter vector $\mathbf{H}$ is padded with zeros to give $(r+1)$ elements as follows

$$\mathbf{H} = [h_0, h_1, \ldots, h_{L-1}, 0, 0, \ldots, 0]. \tag{4.17}$$

In this case, $\mathbf{F}$ is a $(r+1) \times (r+1)$ one-step transition matrix with all elements being zero except those in positions $(i+1, i)$, $i = 1, \ldots, r$, which are 1 and $\mathbf{G}$ is a $(r+1) \times 1$ vector given by

$$\mathbf{G} = [1, 0, 0, \ldots, 0]^T. \tag{4.18}$$

Hence, the order of the Kalman equalizer is the maximum value between $(L-1)$ and r. This strategy was first introduced and analyzed in [47].

4.2 Weighted Gaussian Sum Approximation

The goal of channel equalization can be briefly stated as the estimation of the state vector $\mathbf{D}(k)$ from the observation sequences of the channel output Z^k. It is well known that if the plant noise $\mathbf{u}(k)$, the observation noise $n(k)$, and the initial estimate of the state are Gaussian and mutually independent, a standard Kalman filter yields the optimal estimate of the state [10, 32]. However, if the plant noise is non-Gaussian, which is valid in the context of data channel equalization, the Kalman filter performs poorly [10]. This problem has been resolved in [3] by approximating the *a posteriori* pdf of the state by a weighted Gaussian sum (WGS). The idea to compute the *a posteriori* pdf of a sequence of delayed symbols, in the context of channel equalization, using the approximation of the pdf by a WGS is reported in [50].

Since the state vector $\mathbf{D}(k)$ is a random variable, knowledge of the pdf of the state conditioned on all available measurement data, i.e., $p\big(\mathbf{D}(k)|Z^k\big)$ provides the most complete possible description of the state $\mathbf{D}(k)$. The recursion on $p\big(\mathbf{D}(k)|Z^k\big)$ is explicitly given by the following Bayes' relations [3, 5]

$$p\big(\mathbf{D}(k)|Z^k\big) = \frac{p\big(\mathbf{D}(k)|Z^{(k-1)}\big)p(z(k)|\mathbf{D}(k))}{p\big(z(k)|Z^{(k-1)}\big)} \tag{4.19}$$

$$p\big(z(k)|Z^{(k-1)}\big) = \int p(z(k)|\mathbf{D}(k))p\big(\mathbf{D}(k)|Z^{(k-1)}\big)d\mathbf{D}(k) \tag{4.20}$$

$$p\big(\mathbf{D}(k)|Z^{(k-1)}\big) = \int p(\mathbf{D}(k)|\mathbf{D}(k-1))p\big(\mathbf{D}(k-1)|Z^{(k-1)}\big)d\mathbf{D}(k-1) \tag{4.21}$$

where the initial condition for (4.19) is

$$p\big(\mathbf{D}(0)|Z^{-1}\big) \triangleq p\big(\mathbf{D}(0)\big). \tag{4.22}$$

The likelihood of the observation $p(z(k)|\mathbf{D}(k))$ and $p(\mathbf{D}(k)|\mathbf{D}(k-1))$ are determined from

(4.4) and (4.5) and *a priori* distributions for $d(k)$ and $n(k)$. Except when the *a priori* distributions are Gaussian, it is generally impossible to determine $p\big(\mathbf{D}(k)|Z^k\big)$ in a closed form from equations (4.19), (4.20) and (4.21), the Kalman filter being, then, the solution. To this end, in [50], a network of Kalman filters has been proposed to estimate the state $\mathbf{D}(k)$ which relies on the approximation of the *a posteriori* pdf by a WGS. The problem is that the number of Gaussian terms in the sum increases dramatically through iterations which is not practical for real-time processing. To avoid this problem, we are motivated to use IMM approach to estimate the state $\mathbf{D}(k)$ depending on the weighted Gaussian sum representation of the plant noise as described below.

We consider the *a priori* density function of the plant noise $\mathbf{u}(k)$ given by [50]

$$p(\mathbf{u}(k)) \triangleq \begin{cases} p_i & \text{if } d(k) = d_i \quad 1 \leq i \leq q \\ 0 & \text{otherwise} \end{cases} \tag{4.23}$$

where $d(k)$ can take on q possible values $\{d_i, i = 1, \ldots, q\}$ associated with the probabilities $\{p_i, i = 1, \ldots, q\}$. For simplicity we use binary transmission ($q = 2$) and symbols transmitted are either -1 or $+1$. The extension to the general case is straightforward. This density function is approximated by a weighted sum of Gaussian density functions centered on the discrete values $\{\mathbf{u}_i = \mathbf{G}d_i, i = 1, \ldots, q\}$ that the plant noise can take. This approximation leads to the following WGS representation of the plant noise

$$p(\mathbf{u}(k)) = \sum_{i=1}^{q} p_i \mathcal{N}\big[\mathbf{u}(k); \mathbf{u}_i, \mathbf{B}_i\big] \tag{4.24}$$

where

$$\mathcal{N}\big[\mathbf{u}(k); \mathbf{u}_i, \mathbf{B}_i\big] = |2\pi\mathbf{B}_i|^{-1/2} \exp\Big\{ -\frac{1}{2}[\mathbf{u}(k) - \mathbf{u}_i]^T \mathbf{B}_i^{-1} [\mathbf{u}(k) - \mathbf{u}_i] \Big\} \tag{4.25}$$

$$\sum_{i=1}^{q} p_i = 1 \quad \text{and} \quad p_i \geq 0 \quad \forall i. \tag{4.26}$$

The mean values $\mathbf{u}_i$ establish a grid in the region of state-space that contains probability mass equal to p_i (or at least the significant part of it). For convenience, all of the covariance matrices $\mathbf{B}_i$ are set equal to $\xi \mathbf{I}_L$ where ξ is a positive scalar. As ξ tends to zero, each term in WGS approaches a unit impulse function located at the mean values $\mathbf{u}_i$. Hence, ξ is selected small enough such that each term of the Gaussian sum is effectively equal to zero everywhere except in a small neighborhood of $\mathbf{u}_i$.

4.3 IMM Approach for Data Channel Equalization

The dynamic system of (4.4) and (4.5) is actually a special case of stochastic hybrid systems with additive (continuous) noise and discrete uncertainties in model evolution. Such a system can be described as

$$\mathbf{D}(k+1) = \mathbf{F}_j(k)\mathbf{D}(k) + \mathbf{u}_j(k+1) \tag{4.27}$$

$$z(k) = \mathbf{H}_j(k)\mathbf{D}(k) + n_j(k) \quad \forall j \in \mathcal{Q} \tag{4.28}$$

where $\mathcal{Q} = \{1, \ldots, q\}$ is the set of possible modes. This set of models can handle uncertainties in the system structure and/or parameters as well as in the noise statistics but in our case, as will be seen, uncertainties occur only in the plant noise $\mathbf{u}_j(k)$.

In this Section, a hybrid system for channel equalization with Markovian switching coefficients is presented by multiple models with a given probability of switching between the models (or modes). The model is one of t hypothesized models, $M_1, \ldots, M_t$ for the system and the event that model j is in effect during the sampling period ending at time k (i.e.,

the sampling period $(k-1,k])$ will be denoted by $M_j(k)$. Such systems are called hybrid systems [10], because they have both continuous noise uncertainties (i.e., plant and measurement noises) and discrete uncertainties (i.e., model uncertainties). As the plant noise in (4.24) is approximated by a weighted q Gaussian terms, it can be assumed that one is dealing with a hybrid system with q modes of operation

$$\mathbf{D}(k+1) = \mathbf{F}\mathbf{D}(k) + \mathbf{u}_j(k+1) \quad \forall j \in \mathcal{Q} \tag{4.29}$$

$$z(k) = \mathbf{H}\mathbf{D}(k) + n(k). \tag{4.30}$$

Such a hybrid system state is dependent on a Markovian switching process generated by the symbol sequence and can be estimated effectively using an IMM algorithm [10].

For q models, the IMM algorithm is implemented with q Kalman filters operating in parallel at each cycle. The structure of the IMM algorithm is

$$\left(N_e; N_f\right) = \left(q; q\right) \tag{4.31}$$

where N_e is the *number of estimates* at the start of the cycle of the algorithm and N_f is the *number of filters* in the algorithm. The switching between the models is assumed to be governed by a finite-state Markov chain according to

$$\pi_{ij} \triangleq P\left\{M_j(k)|M_i(k-1)\right\} \tag{4.32}$$

where π_{ij} are the transition probabilities of switching from model $M_i(k-1)$ to model $M_j(k)$. The transition probabilities π_{ij} are assumed known, time-invariant, and independent of the base state.

The IMM-based channel equalization for binary data sequence is shown in Figure 4.3.

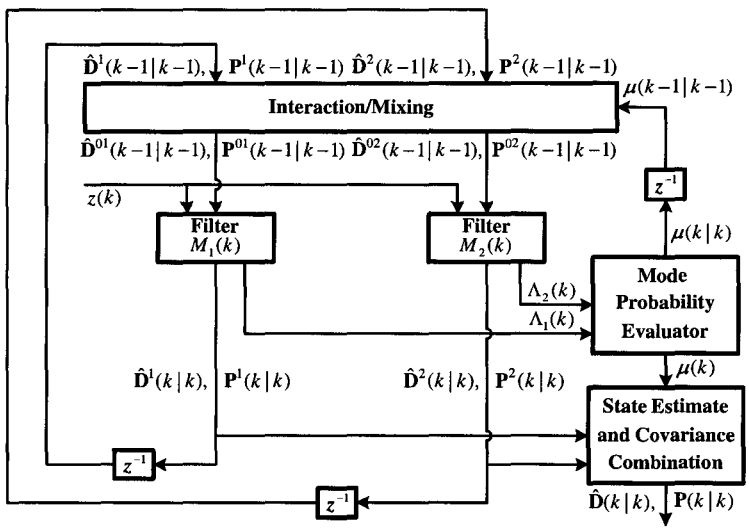

Figure 4.3: One cycle of the IMM estimator for channel equalization.

At every cycle, each Kalman filter will produce a model conditioned state estimate $\hat{\mathbf{D}}^j(k|k)$ and the associated covariance $\mathbf{P}^j(k|k)$ based on its input state $\hat{\mathbf{D}}^{0j}(k-1|k-1)$, covariance $\mathbf{P}^{0j}(k-1|k-1)$ and the current observation $z(k)$. The details of the algorithm are similar to and can be easily understood from the algorithm described in the next Section.

Under the assumption that the channel coefficients are perfectly known, an IMM approach is introduced for stationary channel equalization in this Section. Unfortunately, in many practical communication systems, the channel coefficients are unknown and/or time-varying. We describe next the blind estimation and equalization of time-varying channels using the interactive multiple model estimator.

4.4 Blind Channel Equalization

In this Section, a hybrid system for blind equalization with Markovian switching coefficients is presented by multiple models with a given probability of switching between the models (or modes).

4.4.1 Channel Model

To account for unknown coefficients, we augment the state vector to include the channel parameters as states. Denoting the unknown time-varying channel parameters as a vector $\mathbf{H}(k) = [h_0(k), \ldots, h_{L-1}(k)]^T$, we define the augmented state vector

$$\mathbf{Y}(k) \triangleq \left[\mathbf{D}^T(k) \;\; \mathbf{H}^T(k)\right]^T. \tag{4.33}$$

We adopt a (discrete time) Wiener process to model the time-varying channel given by

$$\mathbf{H}(k+1) = \mathbf{H}(k) + \mathbf{v}(k) \tag{4.34}$$

where $\mathbf{v}(k)$ is an i.i.d. zero-mean Gaussian vector with covariance matrix $\sigma_v^2\ \mathbf{I}_{L\times L}$. The augmented state equation is then

$$\mathbf{Y}(k+1) = \widetilde{\mathbf{F}}\mathbf{Y}(k) + \mathbf{w}(k) \tag{4.35}$$

where

$$\widetilde{\mathbf{F}} = \begin{bmatrix} \mathbf{F} & \mathbf{0}_{L\times L} \\ \mathbf{0}_{L\times L} & \mathbf{I}_{L\times L} \end{bmatrix} \tag{4.36}$$

and the augmented process noise

$$\mathbf{w}(k) = \begin{bmatrix} \mathbf{G}d(k+1) \\ \mathbf{v}(k) \end{bmatrix}. \tag{4.37}$$

The nonlinearity of this formulation arises in the observation equation

$$z(k) = f\big[\mathbf{Y}(k)\big] + n(k) \tag{4.38}$$

where $f\big[\mathbf{Y}(k)\big] = \sum_{j=1}^{L}\big[\mathrm{Y}_j(k)\mathrm{Y}_{j+L}(k)\big] = \mathbf{D}^T(k)\mathbf{H}(k)$ and Y_{j+L} is the channel coefficient corresponding to the jth input Y_j.

The non-Gaussian augmented process noise is approximated by the following WGS representation

$$p(\mathbf{w}(k)) = \sum_{i=1}^{q} p_i \mathcal{N}\big[\mathbf{w}(k); \mathbf{w}_i, \mathbf{B}_i\big] \tag{4.39}$$

where $\mathbf{w}_i = \widetilde{\mathbf{G}}d_i$, $\{d_i, i = 1, \ldots, q\}$ are the q values that $d(k)$ can take associated with the probabilities $\{p_i, i = 1, \ldots, q\}$, $\widetilde{\mathbf{G}} = \big[\mathbf{G}^T \;\; \mathbf{0}_L^T\big]^T$, $\mathbf{B}_i$ is now equal to $\xi\mathbf{I}_{2L}$, $\mathbf{I}_{2L}$ being the identity matrix, and ξ chosen small enough so that each Gaussian density function is located on a neighborhood of $\mathbf{w}_i$ with a probability mass equal to p_i.

4.4.2 IMM Estimator for Blind Channel Equalization

In this Section, a hybrid system for blind equalization with Markovian switching coefficients is presented by multiple models with a given probability of switching between the models (or modes). Since the non-Gaussian augmented process noise (4.39) is approximated by a weighted q Gaussian terms, it can be assumed that one is dealing with a hybrid system with

q modes of operation given by

$$\mathbf{Y}(k+1) = \widetilde{\mathbf{F}}\mathbf{Y}(k) + \mathbf{w}_j(k) \quad \forall j \in \mathcal{Q} \tag{4.40}$$

$$z(k) = f[\mathbf{Y}(k)] + n(k). \tag{4.41}$$

Therefore, such a hybrid system state is dependent on a Markovian switching process generated by the symbol sequence and can be estimated effectively using an IMM algorithm [10]. Similar to the previous method, the switching between the models is assumed to be governed by a finite-state Markov chain according to

$$\pi_{ij} \triangleq P\{M_j(k)|M_i(k-1)\} \tag{4.42}$$

where π_{ij} are the transition probabilities of switching from model $M_i(k-1)$ to model $M_j(k)$ and the transition probabilities π_{ij} are assumed known, time-invariant and independent of the base state.

Since the measurement equation (4.41) is nonlinear in nature, we can linearize $z(k)$ by the following approximation. To obtain the predicted measurement, the nonlinear function in (4.41) is expanded in Taylor series around the predicted state with terms up to first order to yield the first order EKF. Hence, for q models, the IMM algorithm is implemented with q EKFs operating in parallel at each cycle. At every cycle, each EKF will produce a model conditioned state estimate $\hat{\mathbf{Y}}^j(k|k)$ and the associated covariance $\mathbf{P}^j(k|k)$ based on its input state $\hat{\mathbf{Y}}^{0j}(k-1|k-1)$, covariance $\mathbf{P}^{0j}(k-1|k-1)$ and the current observation $z(k)$. We perform partial derivatives of the state matched to mode j $\{j = 1, \ldots, q\}$ evaluated in the

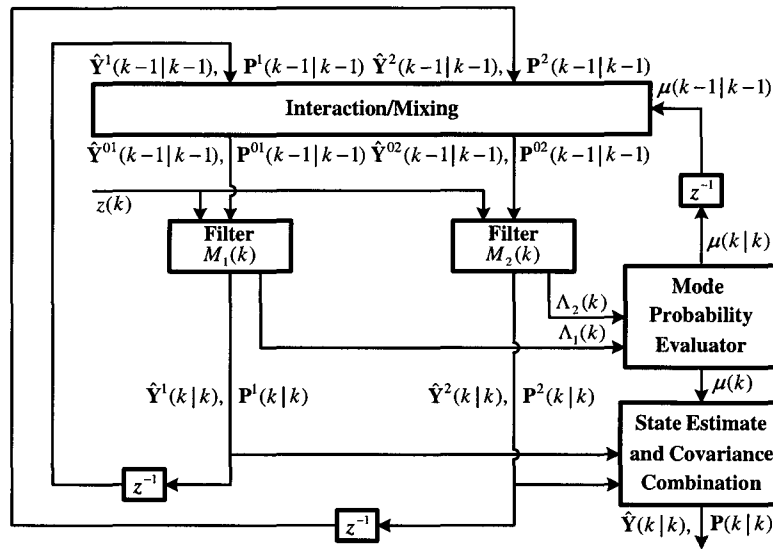

Figure 4.4: One cycle of the IMM estimator for blind equalization.

predicted states $\hat{\mathbf{Y}}^j(k|k-1) = \left[\hat{\mathbf{D}}^{j^T}(k|k-1) \;\; \hat{\mathbf{H}}^{j^T}(k|k-1)\right]^T$ as defined in (4.46) and denote

$$\mathbf{E}^j(k) = \left.\frac{\partial f[\mathbf{Y}(k)]}{\partial \mathbf{Y}(k)}\right|_{\mathbf{Y}(k)=\hat{\mathbf{Y}}^j(k|k-1)} \tag{4.43}$$

$$= \left[\hat{\mathbf{H}}^{j^T}(k|k-1) \;\; \hat{\mathbf{D}}^{j^T}(k|k-1)\right]^T. \tag{4.44}$$

Now, the Taylor series expansion of (4.41) up to first order term is given by

$$z(k) = f[\hat{\mathbf{Y}}^j(k|k-1)] + \mathbf{E}^{j^T}(k)[\mathbf{Y}(k) - \hat{\mathbf{Y}}^j(k|k-1)] + n(k) \tag{4.45}$$

which is the linearized model of (4.41) around the predicted state $\hat{\mathbf{Y}}^j(k|k-1)$.

Therefore, the structure of the IMM algorithm requires q EKFs while each filter matched to mode j $\{j = 1, \ldots, q\}$ will be implemented by the following steps.

Prediction:

$$\hat{\mathbf{Y}}^j(k|k-1) = \widetilde{\mathbf{F}}\hat{\mathbf{Y}}^{0j}(k-1|k-1) + \mathbf{w}_i \tag{4.46}$$

$$\mathbf{P}^j(k|k-1) = \widetilde{\mathbf{F}}\mathbf{P}^{0j}(k-1|k-1)\widetilde{\mathbf{F}}^T + \mathbf{B}_i \tag{4.47}$$

Estimation:

$$\mathbf{K}^j(k) = \frac{\mathbf{P}^j(k|k-1)\mathbf{E}^j(k)}{\mathbf{E}^{j^T}(k)\mathbf{P}^j(k|k-1)\mathbf{E}^j(k) + \sigma_n^2} \tag{4.48}$$

$$\hat{\mathbf{Y}}^j(k|k) = \hat{\mathbf{Y}}^j(k|k-1) + \mathbf{K}^j(k)\big[z(k) - \hat{\mathbf{H}}^{0j^T}(k-1|k-1)\hat{\mathbf{D}}^j(k|k-1)\big] \tag{4.49}$$

$$\mathbf{P}^j(k|k) = \mathbf{P}^j(k|k-1) - \mathbf{K}^j(k)\mathbf{E}^{j^T}(k)\mathbf{P}^j(k|k-1). \tag{4.50}$$

Figure 4.4 depicts the IMM-based blind equalization for binary data sequence. One cycle of the algorithm can be summarized below.

1. ***Calculation of the mixing probabilities***: This is the probability that the symbol corresponding to mode M_i was in effect at $k-1$ given that M_j is in effect at k conditioned on the measurements up to time $k-1$, Z^{k-1}, for all $i, j = 1, \ldots, q$. Here symbols corresponding to i and j takes on values in γ and the calculation is done for all q values in γ.

$$\begin{aligned}\mu_{i|j}(k-1|k-1) &\triangleq P\{M_i(k-1)|M_j(k), Z^{k-1}\} \\ &= \frac{1}{\bar{c}_j}P\{M_j(k)|M_i(k-1), Z^{k-1}\}P\{M_i(k-1)|Z^{k-1}\} \\ &= \frac{1}{\bar{c}_j}\pi_{ij}\mu_i(k-1) \qquad i,j = 1,\ldots,q\end{aligned} \tag{4.51}$$

where $\bar{c}_j = \sum_{i=1}^{q} \pi_{ij}\mu_i(k-1)$ is the normalizing constant and $\mu_i(k-1)$ is the posterior symbol probability at time $(k-1)$.

2. ***Mixing***: Starting with $\hat{\mathbf{Y}}^i(k-1|k-1)$, state estimate of the EKF matched to mode i at time $k-1$, the mixed initial condition for the filter matched to each symbol at time k is computed as follows

$$\hat{\mathbf{Y}}^{0j}(k-1|k-1) = \sum_{i=1}^{q} \hat{\mathbf{Y}}^i(k-1|k-1)\mu_{i|j}(k-1|k-1) \quad j = 1,\ldots,q \tag{4.52}$$

and the associated covariance is

$$\begin{aligned} \mathbf{P}^{0j}(k-1|k-1) \;=\; & \sum_{i=1}^{q} \mu_{i|j}(k-1|k-1)\Big\{\mathbf{P}^i(k-1|k-1) \\ & +\big[\hat{\mathbf{Y}}^i(k-1|k-1) - \hat{\mathbf{Y}}^{0j}(k-1|k-1)\big] \\ & .\big[\hat{\mathbf{Y}}^i(k-1|k-1) - \hat{\mathbf{Y}}^{0j}(k-1|k-1)\big]^T\Big\} \quad j = 1,\ldots,q. \end{aligned} \tag{4.53}$$

3. ***Mode-matched filtering***: The estimate (4.52) and covariance (4.53) are used as input to the filter matched to $M_j(k)$, which uses observation at time k, $z(k)$ to yield $\hat{\mathbf{Y}}^j(k|k)$ and $\hat{\mathbf{P}}^j(k|k)$. These are the outputs of the EKFs. The likelihood functions corresponding to the q filters are computed as

$$\begin{aligned} \Lambda_j(k) \;&\triangleq\; p\big[z(k)|M_j(k), Z^{k-1}\big] \\ &=\; p\big[z(k)|M_j(k), \hat{\mathbf{Y}}^{0j}(k-1|k-1), \mathbf{P}^{0j}(k-1|k-1)\big] \\ &=\; p\big[\nu_j(k)\big] = \mathcal{N}\big[\nu_j(k); 0, S_j(k)\big] \quad j = 1,\ldots,q \end{aligned} \tag{4.54}$$

where $\nu_j(k)$ and $S_j(k)$ are the innovation and its covariance from the mode-matched filter corresponding to mode j.

4. ***Mode probability update***: The probability that the mode j ($j = 1,\ldots,q$) is in effect

is updated as follows

$$\begin{aligned}
\mu_j(k) &\triangleq P\{M_j(k)|Z^k\} \\
&= \frac{1}{c}p\big[z(k)|M_j(k), Z^{k-1}\big]P\big\{M_j(k)|Z^{k-1}\big\} \\
&= \frac{1}{c}\Lambda_j(k)\sum_{i=1}^{q} P\big\{M_j(k)|M_i(k-1), Z^{k-1}\big\}P\big\{M_j(k-1)|Z^{k-1}\big\} \\
&= \frac{1}{c}\Lambda_j(k)\sum_{i=1}^{q}\pi_{ij}\mu_i(k-1) \\
&= \frac{1}{c}\Lambda_j(k)\bar{c}_j \quad j=1,\ldots,q
\end{aligned} \tag{4.55}$$

where the normalizing constant $c = \sum_{j=1}^{q}\Lambda_j(k)\bar{c}_j$.

5. ***Estimate and covariance combination***: Finally, combination of the mode-conditioned estimates and covariances is done using the following mixture equations

$$\hat{\mathbf{Y}}(k|k) = \sum_{j=1}^{q}\hat{\mathbf{Y}}^j(k|k)\mu_j(k) \tag{4.56}$$

$$\begin{aligned}
\mathbf{P}(k|k) = \sum_{j=1}^{q}\mu_j(k)\Big\{&\mathbf{P}^j(k|k) \\
&+\big[\hat{\mathbf{Y}}^j(k|k) - \hat{\mathbf{Y}}(k|k)\big]\big[\hat{\mathbf{Y}}^j(k|k) - \hat{\mathbf{Y}}(k|k)\big]^T\Big\}.
\end{aligned} \tag{4.57}$$

This combination is not part of the algorithm recursions whereas it is used only for output purposes. The estimated state $\hat{\mathbf{Y}}(k|k)$ is the concatenation of the so-resulted estimations of both the symbol vector $\hat{\mathbf{D}}(k|k)$ and the channel coefficients $\hat{\mathbf{H}}(k|k)$. It corresponds to a blind estimation of the channel and the transmitted data sequence corrupted by ISI and noise.

Table 4.1: Channel impulse responses.

Channel Number	Impulse Response	Classification
1	$0.2602 + 0.9298z^{-1} + 0.2602z^{-2}$	Nonminimum Phase Channel
2	$0.4950 + 0.0990z^{-1} - 0.4901z^{-2}$	Nonminimum Phase Channel
3	$0.6082 + 0.7603z^{-1} + 0.2280z^{-2}$	Minimum Phase Channel

4.5 Simulation Results

In this Section, we investigate the effectiveness of proposed approaches by computer simulations. In the simulations environment, the information symbol from independent and identically distributed binary phase shift keying (BPSK) signal is used. For the IMM-based approaches, the initial mode probabilities are

$$\mu(0) = [0.5 \;\; 0.5]^T \tag{4.58}$$

and the mode-switching probability matrix is given by

$$\begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}. \tag{4.59}$$

The observation noise $n(k)$ is generated as a stationary zero-mean Gaussian white process with variance σ_n^2. Decisions on the estimated transmitted symbols are taken with a delay r such that if the delay is greater than the length of the channel impulse response, the channel vector is padded with zeros in order to have a length that is greater than or equal to the delay r. We compute bit error rate (BER) as a measure of the performance of the equalizers. For each signal to noise ratio (SNR), the BER is computed over 100 Monte Carlo runs of length 10000 symbols. Equalization experiments are made up of two Subsections. In the

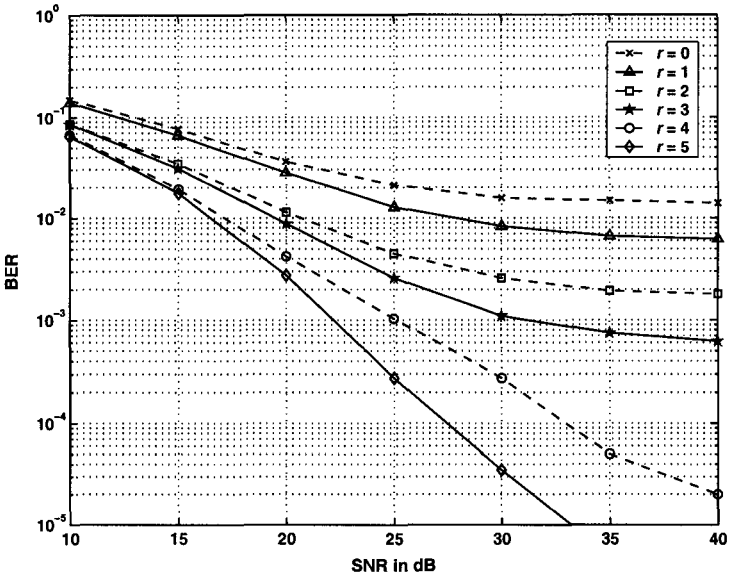

Figure 4.5: Performance of the Kalman equalizer.

first Subsection, the performance of the equalizers are evaluated for channels with ISI while the channel parameters are assumed to be known. The performance of the equalizers for reconstruction of transmitted symbols based on the noise-corrupted channel output without knowing the time-varying FIR channels is described in the second Subsection.

4.5.1 Performance Comparison of Channel Equalizers

The impulse responses of channel models used to evaluate the equalizers are summarized in Table 4.1. First of all, we study the equalizer based on the Kalman filter originally proposed in [47]. Channel 2, which is a nonminimum phase filter, of Table 4.1 is used to model the communication channel. The performance of the Kalman equalizer for different delays is shown in Figure 4.5. As expected, the introduction of a delay in the input estimation improves the BER for a nonminimum phase channel. However, this Kalman equalizer is

based on the assumption that the plant noise is Gaussian, which is not the case in the context of data channel equalization. Taking this fact into consideration, we show next that the overwhelming performance gain can be obtained by using a bank of Kalman filters instead of a single Kalman filter.

Figure 4.6 illustrates the zeros of channel 1 and its frequency response. Using this non-minimum phase filter to model the communication channel, in Figure 4.7, we compare the performance of the proposed IMM-based equalizer with that of the NKF-based equalizer reported in [50]. Decisions on the estimated transmitted symbols are taken with a delay $r = 4$.

For channel 2, which is also nonminimum phase as depicted in Figure 4.8, the performance comparison of these two equalizers is illustrated in Figure 4.9 for $r = 4$. Next, we take into account a minimum phase filter to model the communication channel. Such a channel is the channel 3 as shown in Table 4.1. The zero plot and frequency response of this channel are shown in Figure 4.10. In Figure 4.11, the proposed structure is simulated and compared with the NKF-based approach. From these figures, we notice that the equalizer based on IMM algorithm improves the performance significantly compared to the NKF-based equalizer.

4.5.2 Performance Comparison of Blind Channel Equalizers

In this Subsection, we simulate blind equalizers for time-varying channels. The nonstationary channel is modeled by a (discrete) Wiener process. A time-varying channel is selected for simulation with channel coefficients $\mathbf{H}(0) = [1.0 \quad 0.2 \quad 0.5]^T$ having $\sigma_v^2 = 5 \times 10^{-5}$. The resulting zero plot and frequency response of $\mathbf{H}(0)$ are depicted in Figure 4.12. A realization of this nonstationary channel modeled by the Wiener process is shown in Figure 4.13. In the WGS approximation, the covariance matrix $\mathbf{B}_i$ is taken the same for all i and equal to $\sigma_v^2 \, \mathbf{I}_{2L \times 2L}$. Decisions on the estimated transmitted symbols are done with a delay $r = 2$.

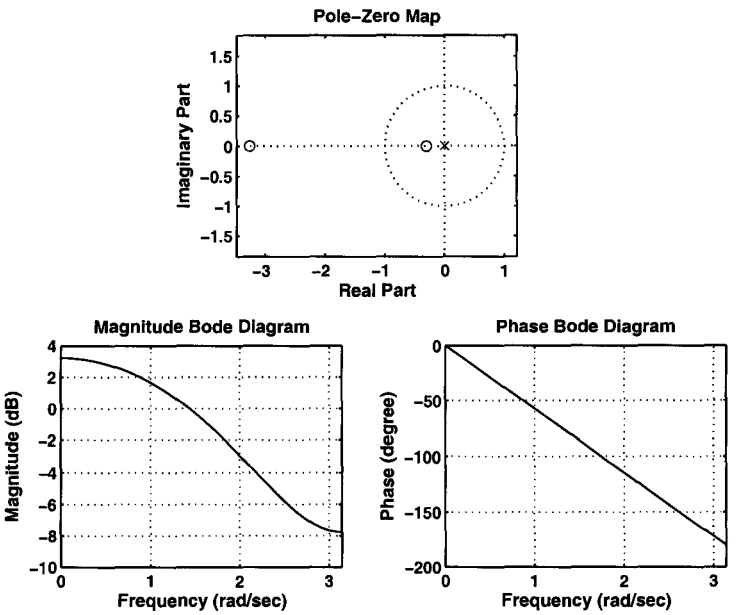

Figure 4.6: Zero plot and frequency response of channel 1.

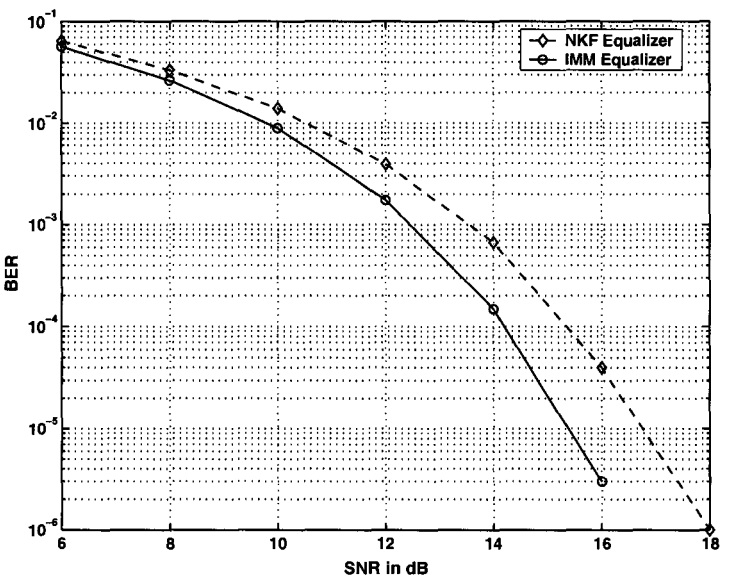

Figure 4.7: Performance of IMM-based and NKF-based equalizers for channel 1.

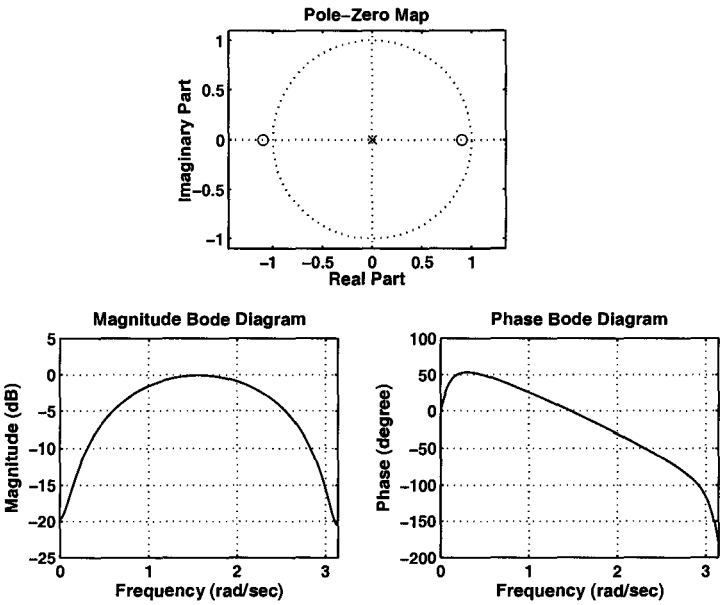

Figure 4.8: Zero plot and frequency response of channel 2.

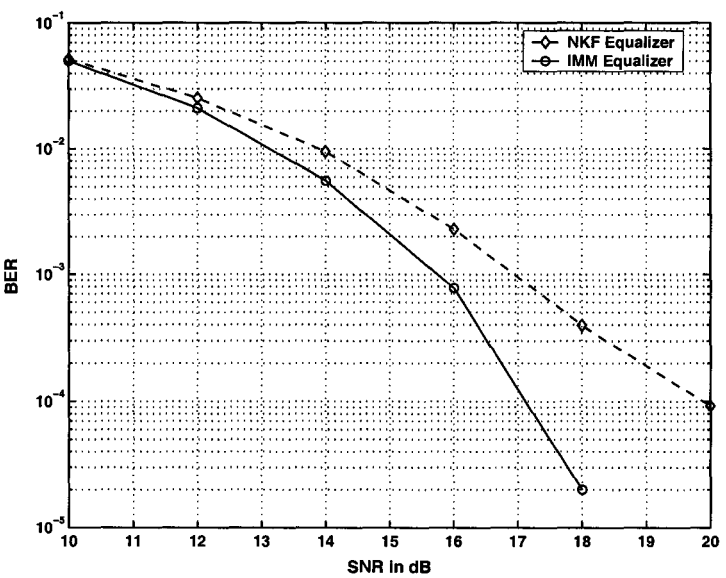

Figure 4.9: Performance of IMM-based and NKF-based equalizers for channel 2.

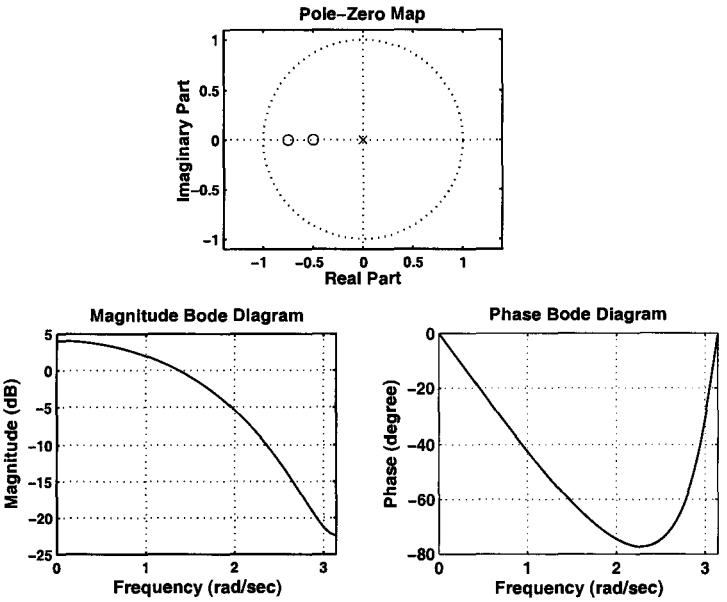

Figure 4.10: Zero plot and frequency response of channel 3.

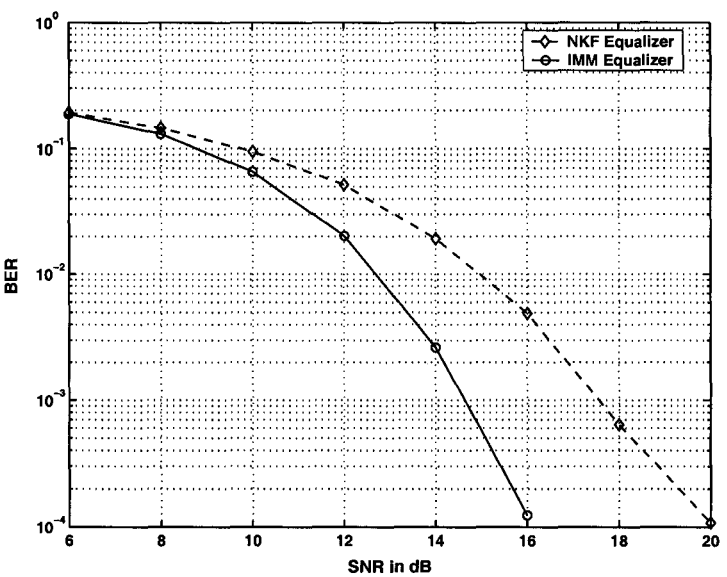

Figure 4.11: Performance of IMM-based and NKF-based equalizers for channel 3.

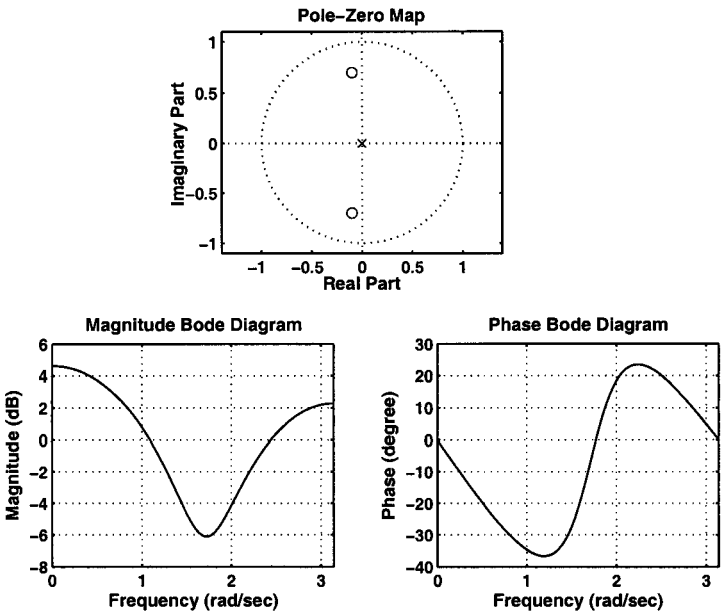

Figure 4.12: Zero plot and frequency response of the channel with initial channel parameters $\mathbf{H}(0) = [1.0 \;\; 0.2 \;\; 0.5]^T$.

In Figure 4.14, the performance of the proposed approach is compared with that of the NEKF-based equalizer by computing the corresponding BER. One hundred simulation runs of length 10000 symbols are used to obtain an average BER. This thus reflects the effects of estimator convergence.

In evaluating the performance of the channel estimation, we define mean square error as follows

$$\text{MSE } (k) = \frac{1}{N_m} \sum_{m=1}^{N_m} \left\{ \frac{1}{k} \sum_{i=1}^{k} \|\mathbf{H}(i) - \hat{\mathbf{H}}_m(i|i)\|^2 \right\} \tag{4.60}$$

where N_m is the number of Monte Carlo trials and $\hat{\mathbf{H}}_m(i|i)$ is the estimate of the channel vector $\mathbf{H}(i)$ at the mth trial. The convergence of channel coefficients is shown in Figure 4.15 where we compare the MSE of the channel estimates as a function of the number of iterations

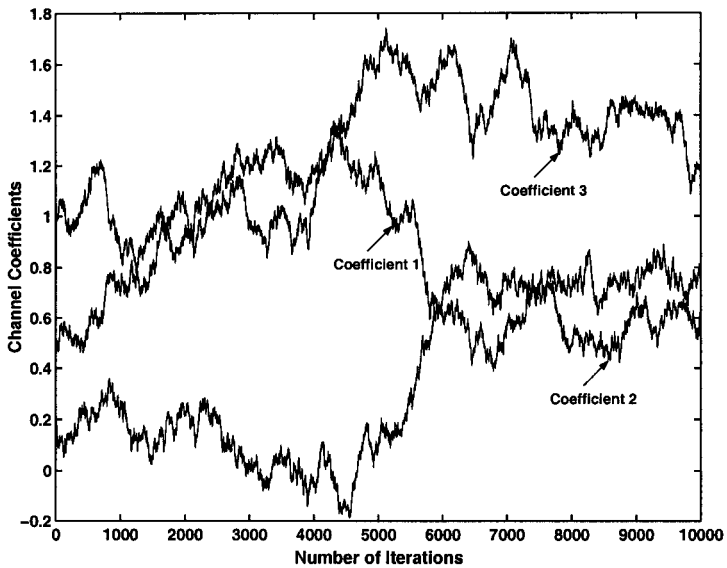

Figure 4.13: A realization of the nonstationary channel with initial channel parameters $\mathbf{H}(0) = [1.0 \;\; 0.2 \;\; 0.5]^T$.

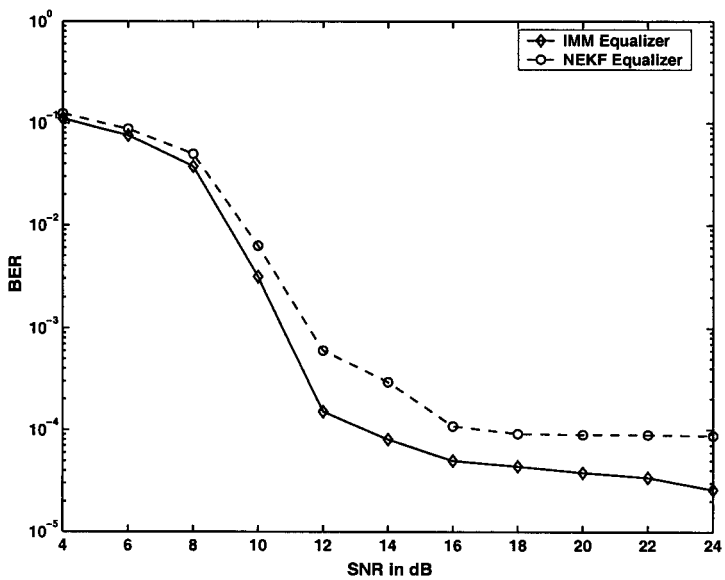

Figure 4.14: BER as a function of SNR.

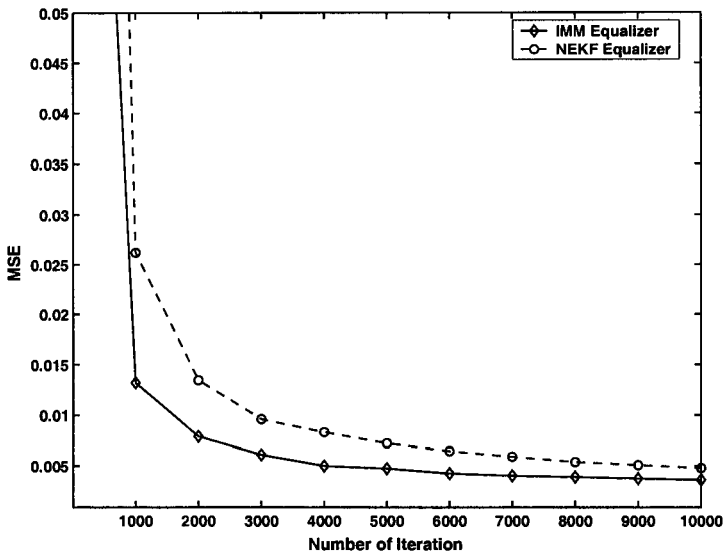

Figure 4.15: MSE with varying number of iterations.

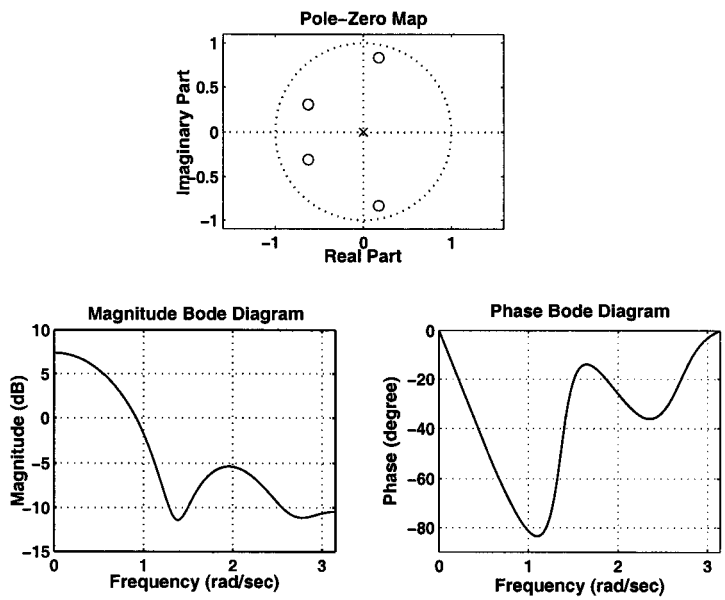

Figure 4.16: Zero plot and frequency response of the channel with initial channel parameters $\mathbf{H}(0) = [0.62 \quad 0.56 \quad 0.48 \quad 0.46 \quad 0.22]^T$.

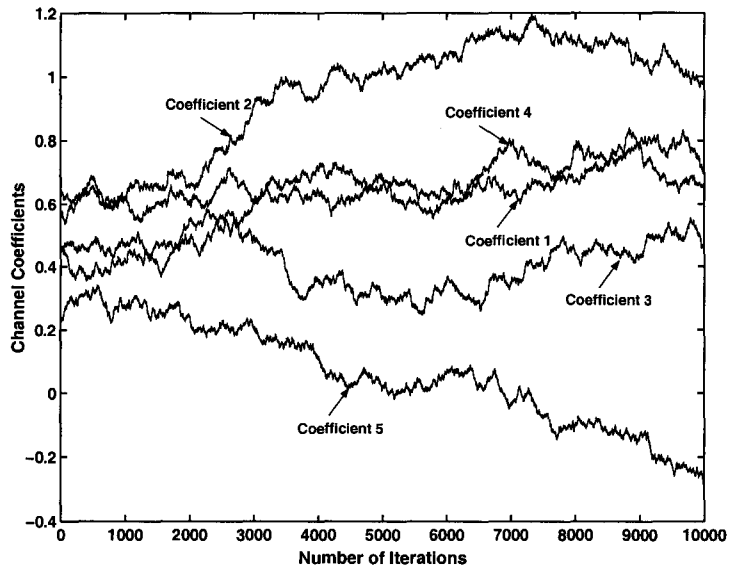

Figure 4.17: A realization of the nonstationary channel with initial channel parameters $\mathbf{H}(0) = [0.62 \quad 0.56 \quad 0.48 \quad 0.46 \quad 0.22]^T$.

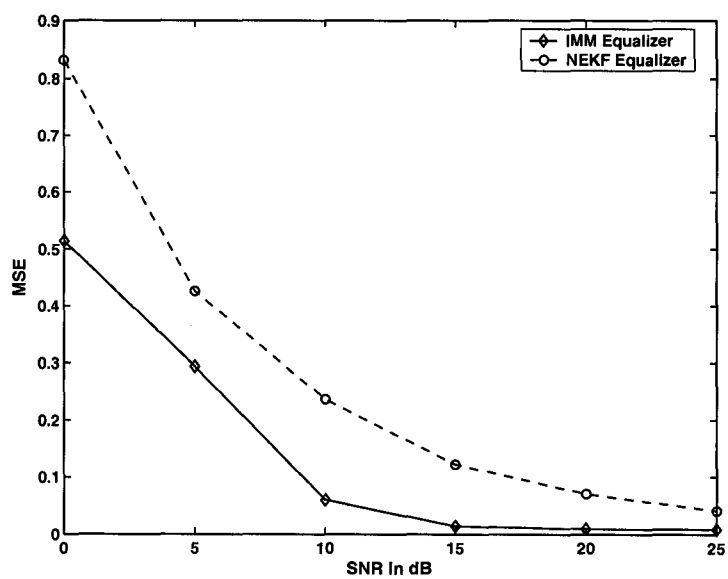

Figure 4.18: MSE as a function of SNR with $r = 4$.

for SNR = 20 dB.

Finally, we consider a channel with initial parameters $\mathbf{H}(0) = [0.62\ \ 0.56\ \ 0.48\ \ 0.46\ \ 0.22]^T$ of which the zero plot and frequency response are delineated in Figure 4.16. Figure 4.17 shows a realization of this channel. Figure 4.18 depicts the convergence of channel coefficients of this channel having $\sigma_v^2 = 6 \times 10^{-6}$ at different SNRs. As seen in these figures, the IMM-based equalizer performs better than the NEKF-based approach.

During the simulations, we have observed that the IMM equalizer is more stable at high SNR than the NEKF equalizer where instability arise when the real symbol predicted states tend more towards L-binary vectors. The superior performance of IMM equalizer is due to its "adaptive bandwidth" capability.

4.6 Conclusions

In this Chapter, we consider the application of the IMM algorithm for the problem of data channel equalization when the plant noise is non-Gaussian. As a solution of this problem, a bank of Kalman filters has been reported in the literature to achieve overwhelming gain over a single Kalman filter. It has been shown that the performance of the NKF-based equalizer can be further improved by considering efficient interactions among the parallel filters. The proposed equalizer based on the IMM algorithm can handle changes in the system structure as well as in the noise statistics.

We also introduce a novel interacting multiple model based nonstationary channel estimator and equalizer. This proposed blind equalizer results in superior performance compared with the previous equalizer consisting of a (static or non-interacting) network of extended Kalman filters. The major advantage of the IMM-based equalizer is that, unlike the NEKF-based equalizer, it avoids the number of terms in the WGS approximation of the plant noise

which increases dramatically through iterations. It also avoids the exponential growth of the state complexity caused by increasing channel memory length in [23]. The IMM-based blind equalizer is also more stable at high SNR.

Chapter 5

Tracking Channels for STBC in Non-Gaussian Noise

This Chapter investigates the problem of estimation and tracking of non-Gaussian channels for space-time block coded signals over a time-selective channel. Time-varying multipath fading is a fundamental phenomenon which makes wireless transmission difficult. In most scattering environments, the effect of multipath fading can be reduced by incorporating multiple antennas at transmitters and/or receivers [37].

One of the efficient ways to exploit the multiple antennas to enable high data rate communications is to use space-time coding [55]. Space-time coding is designed to exploit the multiple antennas by inducing spatial and temporal correlation in the signals transmitted over different antennas. STBC is one of the space-time coding techniques which improves the signal quality at the receiver by simple processing using two antennas at the transmitter. This type of coding/decoding scheme, introduced by Alamouti [2], is based on channel estimation at the receiver in order to decode the transmitted symbol.

In space-time block coded systems, the simple decoding rule is valid only for a flat-fading

channel where the channel gain is constant over two consecutive symbols. However, in reality, this is not the case and ignoring channel variation through the block results in performance degradation. In addition, the assumption of Gaussian noise in communication systems is not always valid as it has been shown in certain circumstances [11, 53], such as indoor communications, the signals are corrupted by non-Gaussian noise. Therefore, some form of channel tracking is needed to keep track of the channel variations even in the non-Gaussian environment if better performance is required.

Based on the Alamouti's coding/decoding technique, in this Chapter, we propose a robust IMM estimator approach for STBC in estimating time-selective fading channels when the measurements are perturbed by the presence of impulsive noise.

5.1 System Description

5.1.1 System Model

We consider a wireless communication system in a scenario with two transmit antennas and one receive antenna. The baseband discrete-time equivalent transmitter and receiver model is presented in Figure 5.1. At the transmitter, the information symbol sequence $s(k)$ at symbol rate $1/T$, where T is the symbol duration, is first parsed into blocks

$$\mathbf{s}(k) \triangleq [s(2k) \quad s(2k+1)]^T. \tag{5.1}$$

Using the Alamouti's space-time encoding scheme [2], the transmitted symbols are as listed in Table 5.1.

We assume that the channel delay spread is smaller than T but the channel coherence time is comparable to T. Under these assumptions, channels are frequency-flat but time-selective.

Table 5.1: Alamouti's space-time coding scheme.

Interval	*Transmitted symbol*	
	Antenna 1	Antenna 2
$2k$	$s(2k)$	$s(2k+1)$
$2k+1$	$-s^*(2k+1)$	$s^*(2k)$

Note that Alamouti's coding scheme achieves full diversity gains in flat-fading channels. We consider here how this coding scheme can be applied in more realistic time-selective channels tracking in many situations where non-Gaussian noise is dominant.

During the time interval $2k$ and $2k+1$, we receive two consecutive samples $z(2k)$ and $z(2k+1)$ given by

$$\begin{aligned} z(2k) &= h_1(2k)s(2k) + h_2(2k)s(2k+1) + w(2k) + b(2k) \\ z(2k+1) &= -h_1(2k+1)s^*(2k+1) + h_2(2k+1)s^*(2k) \\ &\quad + w(2k+1) + b(2k+1) \end{aligned} \tag{5.2}$$

where $h_i(k)$, $i = 1, 2$ denote the time-selective channel from the ith transmit antenna to the receive antenna, $w(k)$ is the AWGN with variance σ_w^2 and $b(k)$ is the impulsive noise. Our goal is to recover $\mathbf{s}(k)$ from two consecutive measurements $z(2k)$ and $z(2k+1)$ in the impulsive noise environment.

5.1.2 Non-Gaussian Noise Model

Traditionally, it has been assumed that communication systems are dominated by AWGN. However, substantial theoretical considerations supported by experimental evidence confirms that the noise in many wireless channels often exhibits non-Gaussian characteristics to some

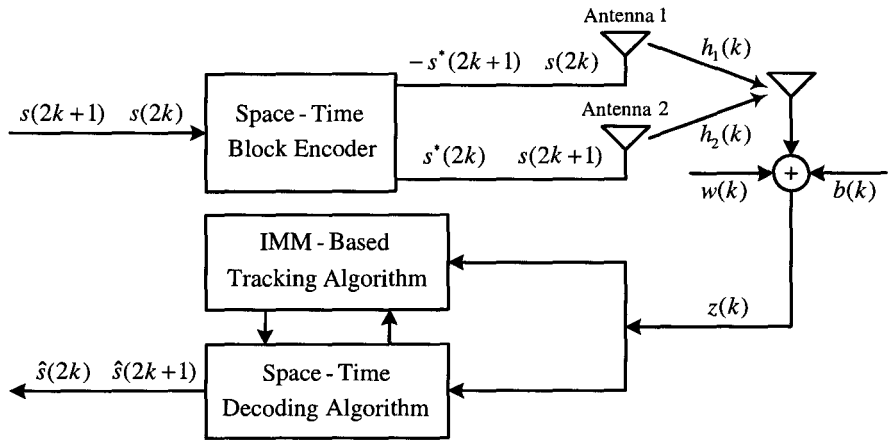

Figure 5.1: Space-time transmission and IMM-based tracking structure.

extent [28,63]. One type of these noises is the impulsive noise.

Impulsive noise can be modeled in a number of ways. A popular model is a mixture model, where the noise density is a linear combination of a nominal density (typically a Gaussian) and a contaminating density. Other models include the alpha-stable family of probability distributions and the elliptical family of distributions. It has been shown that alpha-stable distributions can be well approximated by a finite mixtures of Gaussian pdfs [46].

In the following, the impulsive noise can be modeled as in [29]

$$b(k) = \beta(k)g(k) \tag{5.3}$$

where $\{\beta(k)\}$ stands for Bernoulli process, i.e., an i.i.d. sequence of zeros and ones with $P(\beta(k) = 1) = \epsilon$, and $g(k)$ is a complex white Gaussian noise with zero mean and variance σ_b^2 such as $\sigma_b^2 \gg \sigma_w^2$. Herein, we consider $\sigma_b^2 = \kappa\sigma_w^2$ with $\kappa \gg 1$. Under this model, the

probability density of the channel noise

$$n(k) = w(k) + b(k) \tag{5.4}$$

can be expressed as a Gaussian mixture

$$p(n(k)) = (1-\epsilon)\mathcal{N}[0, \sigma_w^2] + \epsilon\mathcal{N}[0, \underbrace{(\kappa+1)}_{\simeq\kappa}\sigma_w^2] \tag{5.5}$$

where the parameter ϵ is the contamination constant or the probability that an impulse occurs and is a potential source of model mismatch. $\{b(k)\}$ is called an "ϵ-contaminated" noise sequence [42,59]. The model of (5.5) is a good model for many natural impulsive noise sources, such as low-frequency atmospheric noise, man-made impulsive noise, and noise sources occurring in urban and military radio networks [41,56]. Moreover, this model serves as an approximation to the more fundamental Middleton Class A noise model [53,74].

5.1.3 Channel Model

In wireless mobile communications, channel time-variations arise mainly due to Doppler shifts and carrier frequency offsets. There are several methods of modeling time-varying wireless communication channels. These include Jake's method [37,38] and autoregressive (AR) modeling [49,64]. The AR or state-space method is of the most interest as it can fit well into the framework of channel tracking. It is shown in [64,70] that a narrowband time-varying wireless channel can be approximated by the first-order autoregressive AR(1) model which is suitable to capture the channel variations. Therefore, we use such a channel model given by

$$h_i(k) = \alpha h_i(k-1) + v_i(k), \qquad i = 1,2 \tag{5.6}$$

where $v_i(k)$ is an i.i.d. circular complex Gaussian variable with variance σ_v^2 and is statistically independent of $h_i(k-1)$, and the α with $0 < \alpha \leq 1$ is the first-order AR coefficient which is assumed to have been estimated as illustrated in [68]. We assume that two channels $h_i(k),\ i = 1, 2$ are complex Gaussian processes with zero mean and unit covariance and have similar time-variations, i.e., $|\alpha|^2 \gg \sigma_v^2$. Under these assumptions, it can be written as

$$|\alpha|^2 + \sigma_v^2 = 1 \tag{5.7}$$

and

$$\alpha = \mathbb{E}[h_i(k)h_i^*(k-1)]. \tag{5.8}$$

The two channels will induce same career frequency offsets because two transmit antennas share the transmit oscillator. The different angles of arrival of the multipath components may cause different Doppler shifts for the two channels. However, this difference is negligible if the multipath components originate far away from the receiver. Note that, α captures the common part of time-variations in both channels and $v_i(k)$ stands for unmodeled differences in (5.6).

5.2 Space-Time Decoding

Given the received signal $z(k)$ and based on channel model (5.6), we present the space-time decoding technique [49] in this Section.

During the time interval $2k$ and $2k+1$, the received vector can be defined as

$$\mathbf{z}(k) \triangleq [z(2k) \quad z^*(2k+1)]^T. \tag{5.9}$$

Hence, we have the following matrix/vector form

$$\mathbf{z}(k) = \mathbf{H}_k \mathbf{s}(k) + \mathbf{n}(k) \tag{5.10}$$

where

$$\mathbf{n}(k) \triangleq [n(2k) \;\; n^*(2k+1)]^T \tag{5.11}$$

and the channel matrix

$$\mathbf{H}_k \triangleq \begin{bmatrix} h_1(2k) & h_2(2k) \\ h_2^*(2k+1) & -h_1^*(2k+1) \end{bmatrix}. \tag{5.12}$$

In order to decode $\mathbf{s}(k)$ from $\mathbf{z}(k)$, we compute the soft estimate $\mathbf{y}(k) \triangleq [y(2k) \;\; y(2k+1)]^T$ as follows

$$\begin{bmatrix} y(2k) \\ y(2k+1) \end{bmatrix} = \mathbf{H}_k^{\mathcal{H}} \begin{bmatrix} z(2k) \\ z^*(2k+1) \end{bmatrix}. \tag{5.13}$$

Since the outputs $y(2k)$ and $y(2k+1)$ are decoupled, the independent estimate of the transmitted symbols $\hat{s}(2k)$ and $\hat{s}(2k+1)$ is possible. This is due to the fact that under the assumptions imposed on the channel variation, the matrix $\mathbf{H}_k$ is near-unitary in the mean sense as shown in [49]. However, in time-varying channels, the orthogonality of $\mathbf{H}_k$ only holds if $|\alpha|^2 \approx 1$. The symbols $[\hat{s}(2k) \;\; \hat{s}(2k+1)]^T$ are estimated from the vector $[y(2k) \;\; y(2k+1)]^T$ using a decision device, actually this is the ML receiver in the case of flat-fading channels.

The orthogonality property of $\mathbf{H}_k$ is the key property that must be preserved both at the transmitter and receiver in order to achieve error free decoding. To attain this goal, good estimates of the channels must be available at the receiver even in presence of heavy-tailed non-Gaussian disturbances. To this end, in the next Section, we resort to the IMM estimator and develop a robust adaptive algorithm to track the channel variations by exploiting the

knowledge of the impulsive noise parameters (ϵ, κ).

5.3 Robust Recursive Channel Tracking Based on IMM Estimator

We start off with the initial channel responses which can be obtained by training sequences. During the training mode, the receiver knows the transmitted symbols while in decision-directed mode, the decoded symbols take their place. After training session it switches to decision-directed mode as discussed in [49]. In the sequel, we concentrate on the decision-directed mode only and assume that initial channel estimates are available.

The "ϵ-contaminated" channel model is here expressed via both the observation equation

$$z(k) = \tilde{\mathbf{s}}^T(k)\mathbf{h}(k) + n(k) \tag{5.14}$$

where

$$\tilde{\mathbf{s}}(k) \triangleq \begin{cases} [s(k) \quad s(k+1)]^T & \text{if } k \text{ is even} \\ [-s^*(k) \quad s^*(k-1)]^T & \text{if } k \text{ is odd} \end{cases} \tag{5.15}$$

and the following state equation

$$\mathbf{h}(k) = \mathbf{A}\mathbf{h}(k-1) + \mathbf{v}(k) \tag{5.16}$$

where

$$\mathbf{h}(k) = [h_1(k) \;\; h_2(k)]^T \tag{5.17}$$

$$\mathbf{A} = \text{diag}(\alpha, \; \alpha) \tag{5.18}$$

$$\mathbf{v}(k) = [v_1(k) \;\; v_2(k)]^T. \tag{5.19}$$

As can be seen from the state-space model described by (5.14) and (5.16), the estimation of the channel $\mathbf{h}(k)$ requires the knowledge of the decoded symbols $\tilde{\mathbf{s}}(k)$. On the other hand, the detection of $\tilde{\mathbf{s}}(k)$ depends on the good estimates of the channel $\mathbf{h}(k)$. To resolve this joint detection-estimation problem, we first obtain the coarse channel prediction given by

$$\hat{\mathbf{h}}(2k|2k-1) = \alpha\hat{\mathbf{h}}(2k-1|2k-1) \tag{5.20}$$

$$\hat{\mathbf{h}}(2k+1|2k-1) = \alpha^2\hat{\mathbf{h}}(2k-1|2k-1). \tag{5.21}$$

This gives a rough estimate of the correct channel assuming a zero mean process noise. Using this estimate, we can construct a coarse estimated channel matrix denoted by $\hat{\mathbf{H}}_k^{(c)}$ from (5.12) and use (5.13) to obtain coarse estimates of $\tilde{\mathbf{s}}(2k)$ and $\tilde{\mathbf{s}}(2k+1)$ denoted by $\hat{\tilde{\mathbf{s}}}^{(c)}(2k)$ and $\hat{\tilde{\mathbf{s}}}^{(c)}(2k+1)$, respectively. These coarse estimated symbols are then used in the appropriate filtering algorithm to obtain refined channel estimates $\hat{\mathbf{h}}(2k|2k)$ and $\hat{\mathbf{h}}(2k+1|2k+1)$. Finally, the refined channel estimates are used to give a more accurate set of transmitted symbol estimates, $\hat{\tilde{\mathbf{s}}}(2k)$ and $\hat{\tilde{\mathbf{s}}}(2k+1)$.

It is well known that if the plant noise $\mathbf{v}(k)$, the observation noise $n(k)$, and initial estimate of the states are Gaussian and mutually independent a standard KF yields the optimal estimate of the state $\mathbf{h}(k)$ using the knowledge of the decoded $\tilde{\mathbf{s}}(k)$ and the observation $z(k)$ [10]. However, the filtering techniques, based on an ordinary KF, completely lose their

optimal properties because of the spiky character of the impulsive noise which severely degrades the performance of a linear estimator [59]. The parameter ϵ in (5.5) determines the probability of contamination and is a potential source of model mismatch.

In order to tackle the impulsive noise, we develop a system for channel tracking with Markovian switching coefficients by multiple models with a given probability of switching between the models. The models are one of t hypothesized models, $M_1, \ldots, M_t$ for the system and the event that model j is in effect during the sampling period ending at time k will be denoted by $M_j(k)$. Since the measurement noise of the dynamic system described by (5.14) and (5.16) is approximated by the mixture of a Gaussian noise with moderate variance and a Gaussian noise with high variance and low occurrence probability, it can be assumed that one is dealing with a hybrid system with two modes of operation. Hence, such a hybrid system state is dependent on a Markovian switching process generated by the occurrence of impulsive noise and can be estimated effectively using an IMM algorithm [10]. Therefore, for this problem we consider an IMM algorithm with two Kalman filters. One filter is matched to the dynamic system with nominal ambient noise $\mathcal{N}[0, \sigma_w^2]$ and the other is matched to the same dynamic system but with impulsive component $\mathcal{N}[0, \sigma_b^2]$. We denote $\sigma_1^2 = \sigma_w^2$ and $\sigma_2^2 = \sigma_b^2$.

The switching between the models is assumed to be governed by a finite-state Markov chain according to the transition probabilities $\pi_{ij} \triangleq P\{M_j(k)|M_i(k-1)\}$ of switching from model $M_i(k-1)$ to model $M_j(k)$. One cycle of the IMM-based channel tracking as shown in Figure 5.2 is summarized below. At every cycle, each KF will produce a model conditioned state estimate $\hat{\mathbf{h}}^j(k|k)$ and the associated covariance $\mathbf{P}^j(k|k)$ based on its input state $\hat{\mathbf{h}}^{0j}(k-1|k-1)$, covariance $\mathbf{P}^{0j}(k-1|k-1)$ and the current observation $z(k)$.

1. ***Calculation of the mixing probabilities***: This is the probability that the noise component corresponding to mode M_i was in effect at $k-1$ given that M_j is in effect

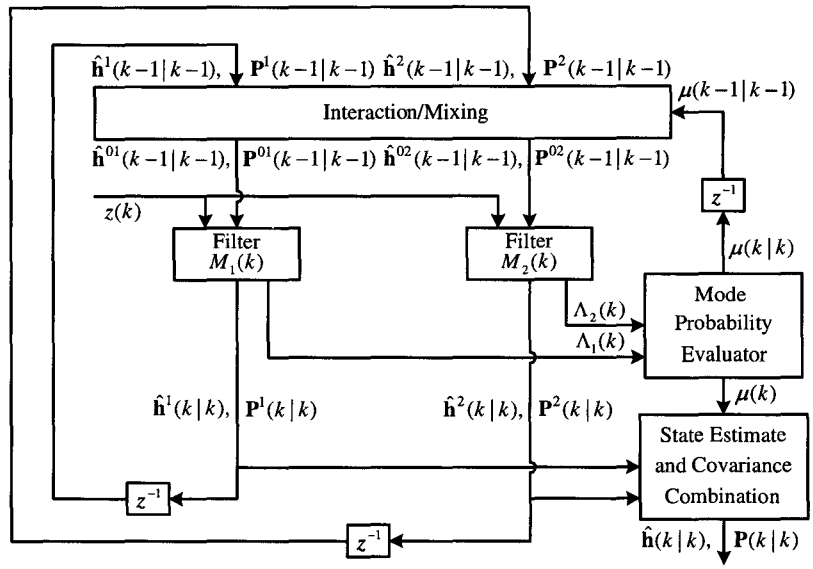

Figure 5.2: One cycle of the IMM estimator for STBC.

at k conditioned on the measurements up to time $k-1$, Z^{k-1}, for $i,j = 1,2$. The mixing probabilities are given by

$$\begin{aligned} \mu_{i|j}(k-1|k-1) &\triangleq P\{M_i(k-1)|M_j(k), Z^{k-1}\} \\ &= \frac{1}{\bar{c}_j}\pi_{ij}\mu_i(k-1) \qquad i,j = 1,2 \end{aligned} \tag{5.22}$$

where $\bar{c}_j = \sum_{i=1}^{2} \pi_{ij}\mu_i(k-1)$ is the normalizing constant and $\mu_i(k-1)$ is the posterior mode probability at time $(k-1)$.

2. ***Mixing***: Starting with $\hat{\mathbf{h}}^i(k-1|k-1)$, state estimate of the KF matched to mode i at time $k-1$, the mixed initial condition for the filter matched to each noise component

at time k is computed as follows

$$\hat{\mathbf{h}}^{0j}(k-1|k-1) = \sum_{i=1}^{2} \hat{\mathbf{h}}^{i}(k-1|k-1)\mu_{i|j}(k-1|k-1) \qquad j=1,2 \tag{5.23}$$

and the associated covariance is

$$\begin{aligned}\mathbf{P}^{0j}(k-1|k-1) \;=\; & \sum_{i=1}^{2} \mu_{i|j}(k-1|k-1)\Big\{\mathbf{P}^{i}(k-1|k-1) \\ & +\big[\hat{\mathbf{h}}^{i}(k-1|k-1) - \hat{\mathbf{h}}^{0j}(k-1|k-1)\big] \\ & .\big[\hat{\mathbf{h}}^{i}(k-1|k-1) - \hat{\mathbf{h}}^{0j}(k-1|k-1)\big]^{T}\Big\} \quad j=1,2. \end{aligned} \tag{5.24}$$

3. ***Mode-matched filtering***: The estimate (5.23) and covariance (5.24) are used as inputs to the filter matched to $M_j(k)$, which uses $z(k)$ to yield $\hat{\mathbf{h}}^{j}(k|k)$ and $\mathbf{P}^{j}(k|k)$. These are the outputs of the KF. Using the estimated information vector $\hat{\tilde{\mathbf{s}}}(k)$ according to time index k, KF equations for the mode j ($j=1,2$) can be written as follows [10,32].

Prediction:

$$\hat{\mathbf{h}}^{j}(k|k-1) \;=\; \mathbf{A}\hat{\mathbf{h}}^{0j}(k-1|k-1) \tag{5.25}$$

$$\mathbf{P}^{j}(k|k-1) \;=\; \mathbf{A}\mathbf{P}^{0j}(k-1|k-1)\mathbf{A}^{T} + \sigma_v^2\mathbf{I} \tag{5.26}$$

Estimation:

$$\mathbf{K}^{j}(k) \;=\; \frac{\mathbf{P}^{j}(k|k-1)\hat{\tilde{\mathbf{s}}}(k)}{\hat{\tilde{\mathbf{s}}}^{T}(k)\mathbf{P}^{j}(k|k-1)\hat{\tilde{\mathbf{s}}}(k) + \sigma_j^2} \tag{5.27}$$

$$\hat{\mathbf{h}}^{j}(k|k) \;=\; \hat{\mathbf{h}}^{j}(k|k-1) + \mathbf{K}^{j}(k)\big[z(k) - \hat{\tilde{\mathbf{s}}}^{T}(k)\hat{\mathbf{h}}^{j}(k|k-1)\big] \tag{5.28}$$

$$\mathbf{P}^{j}(k|k) \;=\; \mathbf{P}^{j}(k|k-1) - \mathbf{K}^{j}(k)\hat{\tilde{\mathbf{s}}}^{T}(k)\mathbf{P}^{j}(k|k-1) \tag{5.29}$$

where $\mathbf{K}^j(k)$ is the Kalman gain of the filter matched to $M_j(k)$. The likelihood functions corresponding to the two filters are computed as

$$\begin{aligned}\Lambda_j(k) &\triangleq p\big[z(k)|M_j(k), Z^{k-1}\big] \\ &= p\big[z(k)|M_j(k), \hat{\mathbf{h}}^{0j}(k-1|k-1), \mathbf{P}^{0j}(k-1|k-1)\big] \\ &= \mathcal{N}\big[\nu_j(k); 0, S_j(k)\big] \qquad j=1,2 \end{aligned} \tag{5.30}$$

where $\nu_j(k)$ and $S_j(k)$ are respectively the innovation and innovation variance given by the filter matched to $M_j(k)$.

4. ***Mode probability update***: The probability that the mode j ($j=1,2$) is in effect is updated as follows

$$\mu_j(k) \triangleq P\{M_j(k)|Z^k\} = \frac{1}{c}\Lambda_j(k)\bar{c}_j \qquad j=1,2 \tag{5.31}$$

where the normalizing constant $c = \sum_{j=1}^{2} \Lambda_j(k)\bar{c}_j$.

5. ***Estimate and covariance combination***: Finally, combination of the mode-conditioned estimates and covariances is done using the following mixture equations

$$\hat{\mathbf{h}}(k|k) = \sum_{j=1}^{2} \hat{\mathbf{h}}^j(k|k)\mu_j(k) \tag{5.32}$$

$$\begin{aligned}\mathbf{P}(k|k) = \sum_{j=1}^{2} \mu_j(k)\Big\{ &\mathbf{P}^j(k|k) \\ &+\big[\hat{\mathbf{h}}^j(k|k) - \hat{\mathbf{h}}(k|k)\big]\big[\hat{\mathbf{h}}^j(k|k) - \hat{\mathbf{h}}(k|k)\big]^T\Big\}. \end{aligned} \tag{5.33}$$

This combination is not part of the algorithm recursions whereas it is used only for output purposes.

The proposed IMM-based channel tracking algorithm can be summarized as follows:

- Wait until two consecutive measurements $z(2k)$ and $z(2k+1)$ become available.
- Obtain coarse estimates $\hat{\mathbf{h}}(2k|2k-1)$ and $\hat{\mathbf{h}}(2k+1|2k-1)$ of current and previous channel states using (5.20) and (5.21).
- Make initial estimates of the current transmitted symbols $\hat{\tilde{\mathbf{s}}}^{(c)}(2k)$ and $\hat{\tilde{\mathbf{s}}}^{(c)}(2k+1)$ using (5.13), the observation vector $\mathbf{z}(k)$, and $\hat{\mathbf{H}}_k^{(c)}$.
- Apply IMM estimator to estimate $\hat{\mathbf{h}}(2k|2k)$and $\hat{\mathbf{h}}(2k+1|2k+1)$ using $\hat{\tilde{\mathbf{s}}}^{(c)}(2k)$ and $\hat{\tilde{\mathbf{s}}}^{(c)}(2k+1)$.
- Construct $\hat{\mathbf{H}}_k$ from $\hat{\mathbf{h}}(2k|2k)$and $\hat{\mathbf{h}}(2k+1|2k+1)$.
- Re-estimate symbols $\hat{\tilde{\mathbf{s}}}(2k)$ and $\hat{\tilde{\mathbf{s}}}(2k+1)$ from $\hat{\mathbf{H}}_k$ using (5.13).

5.4 Simulation Results

Computer simulations are conducted to illustrate the effectiveness of the proposed approach. The system we simulate use a frame length of 150 information symbols. The transmitted sequence of quadrature phase shift keying (QPSK) modulation is employed. The channel impulse response used in simulations is a time-selective fading channel (i.e., time-varying but frequency-flat). The time-selective fading channels are generated by initializing $h_i(0)$ as a complex Gaussian variable with unit variance and taking $\alpha = 0.998$.

The modes of the hybrid system evolve according to a Markov chain. In this problem, the impulsive noise occurs independently in time. Therefore, the mode-switching probability

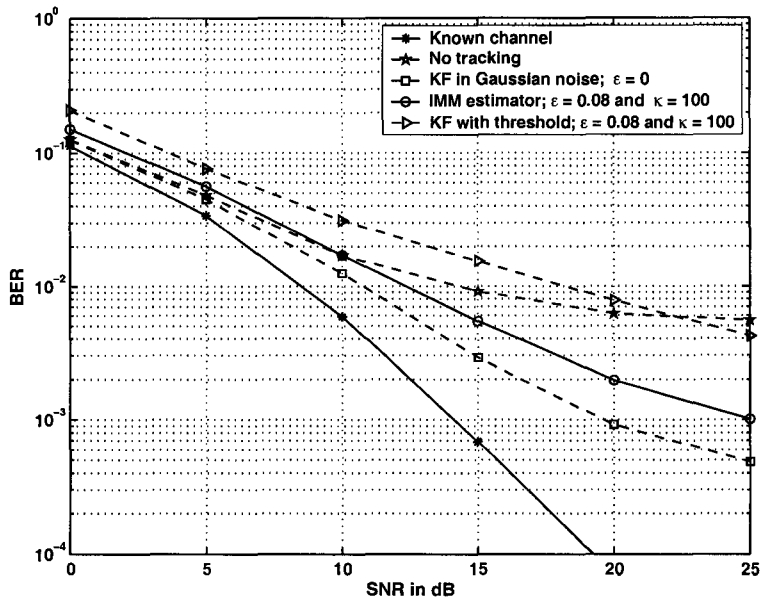

Figure 5.3: BER as a function of SNR.

matrix is selected as

$$\begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{bmatrix} = \begin{bmatrix} 1-\epsilon & \epsilon \\ 1-\epsilon & \epsilon \end{bmatrix}. \tag{5.34}$$

This yields a white, binary chain for the mode evolution with probabilities $1-\epsilon$ and ϵ for the two modes. The parameter of the Bernoulli sequence is $\epsilon = 8 \times 10^{-2}$ and the impulsive noise parameter $\kappa = 100$. In order to avoid divergence in Kalman filtering, we insert one pilot symbol every 12 symbols which incurs a loss of 8% bandwidth efficiency. To maintain fairness, when no channel tracking is used, we set the receiver with perfect current channel estimates every 12 symbols.

The simulated BER performance is depicted in Figure 5.3. For each SNR, the BER is computed over 5000 Monte Carlo runs of 150 iterations. This thus reflects the effects of

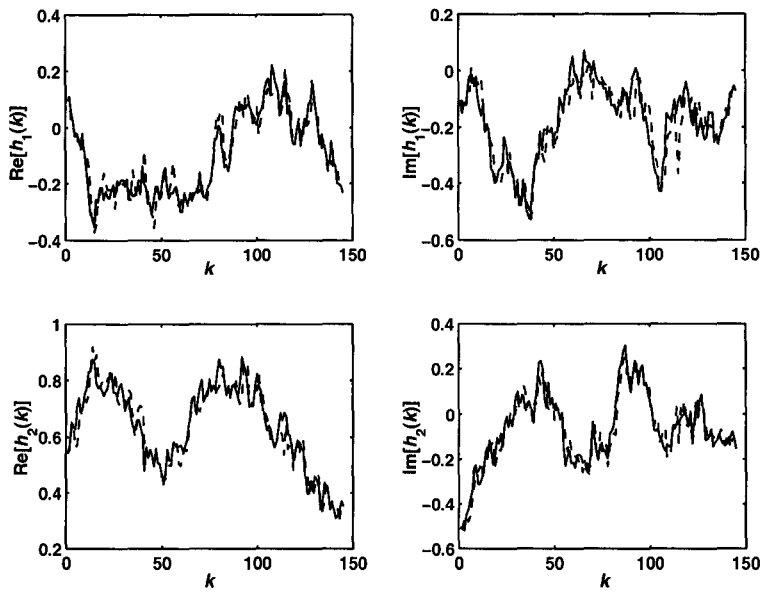

Figure 5.4: IMM-based channel tracking capability.

estimator convergence. From Figure 5.3, it can be noticed that IMM-based approach shows excellent behavior since the curve after cancellation of impulsive noise is only marginally different from the curve obtained by KF-based approach with Gaussian noise only (i.e., when $\epsilon = 0$) published in [49]. We can also remark that the proposed method outperforms the classical KF with a fixed threshold to tackle impulsive noise as reported in [42].

Figure 5.4 illustrates the migration of the channel coefficients through time and tracking capability of the proposed approach to follow the movement of the channel coefficients at SNR of 25 dB in impulsive noise environment. The solid lines in the graphs depict the true channel coefficients while the dotted lines are the estimated coefficients. We can observe that our IMM estimator yields excellent tracking capability in presence of impulsive noise by keeping fairly close to the true values throughout.

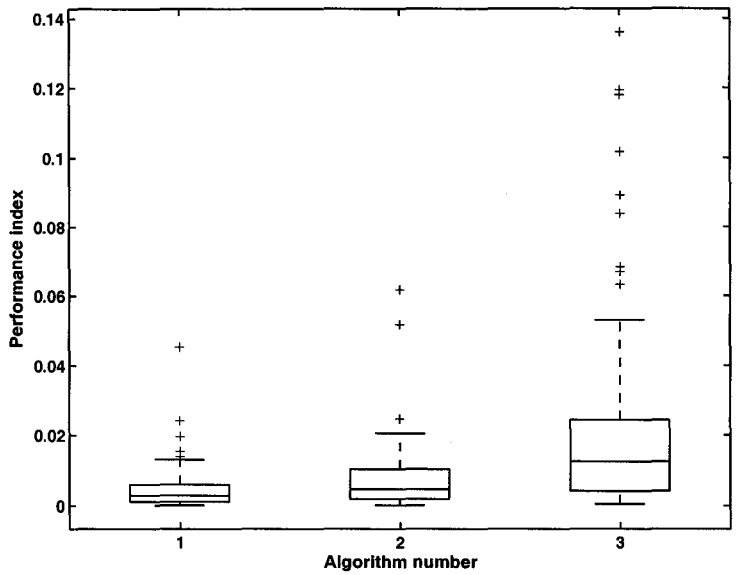

Figure 5.5: Performance index. Algorithm numbers correspond (in order): KF ($\epsilon = 0$), IMM estimator ($\epsilon = 8 \times 10^{-2}$ and $\kappa = 100$), and KF with threshold ($\epsilon = 8 \times 10^{-2}$ and $\kappa = 100$).

Alamouti's coding/decoding scheme is based on the orthogonality of the estimated channel matrix $\hat{\mathbf{H}}_k$, i.e., $\hat{\mathbf{H}}_k^{\mathcal{H}}\hat{\mathbf{H}}_k$ should be diagonal. To demonstrate the stability of the proposed approach, we define the performance index as

$$\mathcal{I}_{\text{perf}} \triangleq \sum_{q \neq p} \mathcal{I}_{pq} \tag{5.35}$$

where

$$\mathcal{I}_{pq} \triangleq \mathbb{E}[\|(\hat{\mathbf{H}}_k^{\mathcal{H}}\hat{\mathbf{H}}_k)_{pq}\|^2] \tag{5.36}$$

is called interference to signal ratio (ISR) which measures the ratio of the power of the interference of the qth symbol to the power of the pth symbol estimated as in (5.13). Each

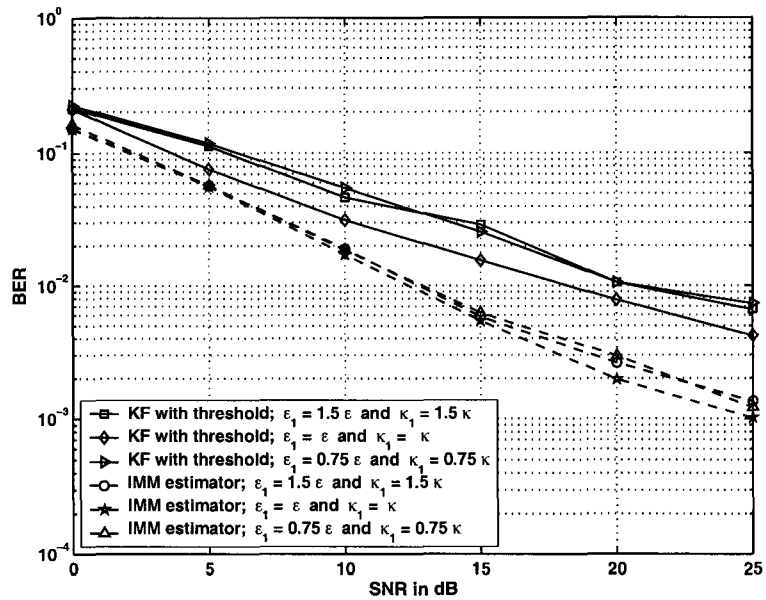

Figure 5.6: BER with varying SNR.

ISR is calculated for every simulation and averaged over 100 realizations. The boxplot of the performance indices is depicted in Figure 5.5. For each data set, the boxplot has a box whose top is at the upper quartile values and bottom is the lower quartile values. The median is plotted as a horizontal line inside the box. The lines extending from each end show the extent of the rest of the results with outlier presented as crosses. It can be inferred that the IMM-based approach is more stable than the channels tracking using KF with a fixed threshold. The IMM algorithm is also robust to inaccurate prior information of the impulsive noise. Suppose the exact values of the impulsive noise parameters (ϵ, κ) are not known. A reasonable assumption for the values of ϵ and κ will yield acceptable results because the IMM algorithm will adjust the mode probabilities properly. Figure 5.6 depicts the BER performance of the proposed approach for the same dynamic system with $(\epsilon = 8 \times 10^{-2}, \kappa = 100)$ while the algorithm assumes the impulsive noise parameters $(\epsilon_1 = 1.5\epsilon, \kappa_1 = 1.5\kappa)$. Figure 5.6 also

shows similar results for ($\epsilon_1 = 0.75\epsilon, \kappa_1 = 0.75\kappa$). It is seen that the BER in the IMM-based channels tracking is almost the same while the KF with fixed threshold has significantly performance degradation.

5.5 Conclusions

In many practical wireless channels in which space-time block coding techniques may be applied, the ambient noise is likely to have an impulsive component that gives rise to larger tail probabilities than is predicted by the Gaussian model. Although Kalman filters are often used in practice to track the channels variation, their performance degrades severely when the plant and/or measurement noises are not Gaussian.

In this Chapter, we address a multiple model based time-selective fading channels tracking approach for STBC with non-Gaussian ambient noise. The proposed adaptive channels tracking scheme is based on a state-space representation of the STBC communication systems and the prior information of the measurement noise which is corrupted by the impulsive noise. The pdf of the measurement noise is considered here as two Gaussian terms weighted by the probability of appearance of Gaussian and impulsive noises.

In this method, neither clipping nor impulses correction mechanism is applied. It has been shown that the proposed approach is less vulnerable to the impulsive noise and has superior performance compared to the KF with a threshold scheme. The IMM-based algorithm is consistent and also robust to inaccurate *a priori* information on the impulsive noise.

Chapter 6

Summary

This Chapter of the thesis concludes with a summary of results and avenues for further research work.

6.1 Conclusions

The focus of this thesis has been on the development and analysis of techniques for data channel equalization and channel estimation for space-time block coded systems with additive non-Gaussian noise which are generic problems in signal processing. To resolve these problems, we applied the multiple model estimator based approaches for parameter tracking. Specifically, we employed the Interacting Multiple Model algorithm which incorporates a Bayesian framework dealing with dynamic situation of switching factors. The IMM approach is based on the Kalman filtering techniques and is widely used in the fields of maneuvering target tracking especially when the system can be described by a bank of multiple state-space models.

Firstly, we proposed a practical application of the IMM estimator to the equalization of digital communications channels to combat the unfavorable channel effects such that the

transmitted signal can be preserved with highest integrity. Our approach is based on the approximation of the *a posteriori* probability density functions of the symbol sequences by a weighted sum of Gaussian density functions, where one filter is matched to the dynamic system with each Gaussian density function. Recently, a network of Kalman filters has been applied to the problem of data channel equalization. A serious drawback of this approach is that the number of Gaussian terms in the sum increases exponentially through iterations. In order to make it computationally feasible, the Gaussian sum in the equalizer is truncated after each iteration, which results in significant performance degradation. In addition, it did not consider the interaction between the parallel filters during the estimation process. In this thesis, it is demonstrated that the network of Kalman filter based solution can be further improved by using a *switching* or *interacting* multiple model estimation approach that computes the state estimate accounting for possible transitions in the models from one time to another.

This approach is based on the assumption that the channel coefficients are perfectly known. Unfortunately, in many practical communication systems, the channel coefficients are unknown and/or nonstationary. Next, to account for unknown channel coefficients, we proposed a novel interacting multiple model based nonstationary channel estimator and equalizer for systems with non-Gaussian plant noise. The proposed blind equalizer, which is completely blind towards any learning phase, can handle changes in the system structure. It has been shown that IMM based blind equalizer results in superior performance compared to the previous equalizer consisting of a (static or noninteracting) network of extended Kalman filters. As the IMM algorithm handles the interactions between parallel filters in an efficient way, performance is improved without much increase in complexity. The IMM equalizer is more stable in high SNR. It also avoids the exponential growth of the state complexity

caused by increasing memory length for blind equalization using parallel adaptive filtering reported in the literature. The major advantage of the proposed approaches is that, unlike the NKF and NEKF equalizers, proposed equalizers avoid the number of terms in the weighted Gaussian sum approximation of the plant noise which increases dramatically through iterations.

Finally, we considered the problem of channel estimation and tracking for space-time block coding contaminated by additive non-Gaussian noise. The proposed adaptive channels tracking scheme is based on a hybrid system which is the state-space representation of the communication system and the prior information of the measurement noise. The measurement noise of this hybrid system is approximated by the mixture of a Gaussian noise with moderate variance and a Gaussian noise with high variance and low occurrence probability. The state of such a system which is dependent on a Markovian switching process generated by the occurrence of impulsive noise is estimated by the IMM estimator. We have shown that the IMM based channel tracking is less vulnerable to the impulsive noise and has superior performance compared to the traditional method available in the literature using the Kalman filter with a threshold scheme to combat the impulsive noise. In addition, it has been shown that the IMM-based approach is consistent and robust to inaccurate *a priori* information on the non-Gaussian noise component. The excellence of the IMM estimator for the system with several behavior modes which can *switch* from one to another was proven again.

6.2 Future Directions

Although this thesis has laid a theoretical basis for use of the multiple model estimator for channel equalization with non-Gaussian plant noise and space-time block coded systems with

non-Gaussian additive noise, there are several issues that remain to be explored. We discuss some important areas where further studies could be fruitfully directed.

- The proposed channel equalizer considers the channels which are linear in nature. However, amplifier devices used in high speed data transmission and satellite communications operate in saturation region. These amplifier devices introduce in the transmitted message memoryless nonlinearities which, combined with the effects of the transmitter and receiver filters, become nonlinearities with memory. Such nonlinear channels can be modeled by Volterra series. Hence, a future challenge could be to modify the proposed equalizer structure for nonlinear channels.

- The equalizer designed by the network of Kalman filters is dependent on the *a priori* knowledge of the channel parameters and on the noise variance which is not a real life situation. To circumvent this obstacle, there may be some algorithms which can be used to make the structure adaptive.

- In the design of blind equalizer, to account for unknown channel coefficients, we augment the state vector to include the channel parameters. A detailed analysis could yield some useful results, such as the initialization and stability of such blind algorithm.

- The time-varying nature of wireless channels is mainly due to the Doppler shift. To model such a channel, it has been reported that a first order autoregressive process is suitable to capture the slow channel variations. In channel tracking for space-time block coded systems, it is assumed that the autoregressive coefficient has been estimated earlier. Hence, blind estimation of the autoregressive model parameters is an important area of further research.

- Carrier frequency-offsets and Doppler shifts give rise to time-selectivity in wireless channels, while high data rates and multipath propagation introduce frequency-selectivity

in wireless links. The combined time- and frequency-selective (or doubly-selective) channels affect critically the performance of communication systems. It is important and challenging task to mitigate these channels, but once acquired, they offer joint multipath-Doppler diversity gains. Therefore, the quality of channel acquisition has a major impact on the design of space-time coded systems. It remains to explore the techniques for tracking channels in fast time-varying fading, frequency-selective fading, and their extensions to doubly-selective fading environments.

Bibliography

[1] G. A. Ackerson and K. S. Fu, "On state estimation in switching environments," *IEEE Transaction on Automatic Control*, vol. AC-15, no. 1, pp. 10–17, February 1970.

[2] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Select Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, October 1998.

[3] D. L. Alspach and H. W. Sorenson, "Nonlinear Bayesian estimation using Gaussian sum approximations," *IEEE Transaction on Automatic Control*, vol. AC-17, no. 4, pp. 439–448, August 1972.

[4] R. Amara and S. Marcos, "A blind network of extended Kalman filters for nonstationary channel equalization," in *Proceedings IEEE ICASSP*, Salt Lake City, Utah, vol. 4, pp. 2117–2120, May 2001.

[5] B. D. O. Anderson and J. B. Moore, *Optimal Filtering.* Englewood Cliffs, NJ: Prentice Hall, 1979.

[6] A. Averbuch, S. Itzikowitz, and T. Kapon, "Radar target tracking - Viterbi versus IMM," *IEEE Transaction on Aerospace and Electronic Systems*, vol. 27, no. 3, pp. 550–563, May 1991.

[7] E. Baccarelli and R. Cusani, "Combined channel estimation and data detection using soft statistics for frequency-selective fast-fading digital links," *IEEE Transaction on Communications*, vol. 46, no. 4, pp. 424–427, April 1998.

[8] B. Balakumar, T. Kirubarajan, and A. Gershman, "Blind adaptive multiuser detection over time-varying time-dispersive channels," in *Proceedings IEEE ICSMC*, Washington D. C., USA, pp. 1922–1927, October 2003.

[9] Y. Bar-Shalom and K. C. Chang, "Tracking a maneuvering target using input estimate versus the interacting multiple model algorithm," *IEEE Transaction on Aerospace and Electronic Systems*, vol. AES-25, no. 2, pp. 296–300, March 1989.

[10] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation: Theory, Algorithms, and Software.* New York: John Wiley & Sons, Inc., 2001.

[11] K. L. Blackard, T. S. Rappaport, and C. W. Bostian, "Measurements and models of radio frequency impulsive noise for indoor wireless communications," *IEEE Journal on Selected Areas of Communications*, vol. 11, no. 7, pp. 991–1001, September 1993.

[12] T. K. Blankenship, D. M. Kriztman, and T. S. Rappaport, "Measurements and simulation of radio frequency impulsive noise in hospitals and clinics," in *Proceedings IEEE VTC'97*, Phoenix, AZ, vol. 3, pp. 1942–1946, May 1997.

[13] C. A. Belfiore and J. H. Park, Jr., "Decision feedback equalization," *Proceedings of the IEEE*, vol. 67, no. 8, pp. 1143–1156, August 1979.

[14] H. A. P. Blom, "A sophisticated tracking algorithm for ATC surveillance data," in *Proceedings International Conference on Radar*, Paris, France, pp. 393–398, May 1984.

[15] H. A. P. Blom, "An efficient filter for abruptly changing systems," in *Proceedings IEEE Conference on Decision and Control*, Las Vegas, Nevada, pp. 656–658, December 1984.

[16] H. A. P. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with Markovian switching coefficients," *IEEE Transaction on Automatic Control*, vol. 33, no. 8, pp. 780–783, August 1988.

[17] G. Castellini, F. Conti, E. Del Re, and L. Pierucci, "A continuously adaptive MLSE receiver for mobile communications: Algorithm and performance," *IEEE Transaction on Communications*, vol. 45, no. 1, pp. 80–89, January 1997.

[18] C. B. Chang and M. Athans, "State estimation for discrete systems with switching parameters," *IEEE Transaction on Aerospace and Electronic Systems*, vol. AES-14, no. 5, pp. 418–425, 1978.

[19] S. Chen, B. Mulgrew, and S. McLaughlin, "Adaptive Bayesian equalizer with decision feedback," *IEEE Transactions on Signal Processing*, vol. 41, no. 9, pp. 2918–2927, September 1993.

[20] J. Choi, "Equalization and semi-blind channel estimation for space-time block coded signals over a frequency-selective fading channel," *IEEE Transactions on Signal Processing*, vol. 52, no. 3, pp. 774–785, March 2004.

[21] H. A. Cipran and M. K. Tsatsanis, "Blind receivers for nonlinearly modulated signals in multipath," *IEEE Transactions on Signal Processing*, vol. 47, no. 2, pp. 583–586, February 1999.

[22] E. Daeipour and Y. Bar-Shalom, "An interacting multiple model approach for target tracking with glint noise," *IEEE Transaction on Aerospace and Electronic Systems*, vol. 31, no. 2, pp. 706–715, April 1995.

[23] A. Duel-Hallen and C. Heegard, "Delayed decision-feedback sequence estimation," *IEEE Transaction on Communications*, vol. 37, no. 5, pp. 428–436, May 1989.

[24] M. V. Eyuboğlu and S. U. H. Qureshi, "Reduced-state sequence estimation with set partitioning and decision feedback," *IEEE Transaction on Communications*, vol. 36, no. 1, pp. 13–20, January 1988.

[25] M. V. Eyuboğlu and S. U. H. Qureshi, "Reduced-state sequence estimation for coded modulation of intersymbol interference channels," *IEEE Journal on Selected Areas in Communications*, vol. 7, no. 6, pp. 989–995, August 1989.

[26] G. D. Forney, Jr., "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Transaction on Information Theory*, vol. 18, no. 3, pp. 363–378, May 1972.

[27] G. D. Forney, Jr., "The Viterbi algorithm," *Proceedings of the IEEE*, vol. 61, no. 3, pp. 268–278, March 1973.

[28] J. G. Gonzalez, "Robust techniques for wireless communication in non-Gaussian environment," Ph. D. dissertation, Dept. of Elect. Comput. Eng., University of Delware, December 1997.

[29] M. Ghosh, "Analysis of the effect of impulse noise on multicarrier and single carrier QAM systems," *IEEE Transaction on Communications*, vol. 44, no. 2, pp. 145–147, February 1996.

[30] A. R. Hammons, Jr. and H. E. Gamal, "On the theory of space-time codes for PSK modulation" *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 524–542, March 2000.

[31] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Transactions on Information Theory*, vol. 48, no. 7, pp. 1804–1824, July 2002.

[32] S. Haykin, *Adaptive Filter Theory.* Prentice Hall, 2001.

[33] R. E. Helmick and W. D. Blair, "Fixed-interval smoothing for Markovian switching systems," *IEEE Transaction on Information Theory*, vol. 41, no. 6, pp. 1845–1855, November 1995.

[34] R. E. Helmick and W. D. Blair, "Fixed-interval smoothing for Markovian switching systems," *IEEE Transaction on Automatic Control*, vol. 41, no. 7, pp. 1051–1056, July 1996.

[35] R. A. Iltis, "A Bayesian maximum-likelihood sequence estimation algorithm for *a priori* unknown channels and symbol timing," *IEEE Journal on Selected Areas in Communications*, vol. 10, no. 3, pp. 579–588, April 1992.

[36] R. A. Iltis, J. J. Shynk, and K. Giridhar, "Bayesian algorithms for blind equalization using parallel adaptive filtering," *IEEE Transaction on Communications*, vol. 42, no. 2/3/4, pp. 1017–1032, February/March/April 1994.

[37] W. C. Jakes, *Microwave Mobile Communications.* New York: John Wiley & Sons, 1974.

[38] M. C. Jeruchim, P. Balaban, and K. S. Shanmugan, *Simulation of Communication Systems.* New York: Plenum Press, 1992.

[39] C. R. Johnson, Jr., P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown, and R. A. Casas, "Blind equalization using the constant modulus criterion: A review," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 1927–1950, October 1998.

[40] G. K. Kaleh and R. Vallet, "Joint parameter estimation and symbol detection for linear or nonlinear unknown channels," *IEEE Transaction on Communications*, vol. 42, no. 7, pp. 2406–2413, July 1994.

[41] S. A. Kassam, *Signal Detection in Non-Gaussian Noise.* New York: Springer-Verlag, 1988.

[42] H. Kim, J. Lim, S. Baek, and K. Sung, "Robust Kalman filtering with variable forgetting factor against impulsive noise," *IEICE Transactions on Fundamentals*, vol. E84-A, pp. 363–366, January 2001.

[43] J. B. Kim, K. Y. Lee, and C. W. Lee, "On the applications of the interacting multiple model algorithm for enhancing noisy speech," *IEEE Transaction on Speech and Audio Processing*, vol. 8, no. 3, pp. 349–352, May 2000.

[44] N. S. Kim, "IMM-based estimation for slowly evolving environments," *IEEE Signal Processing Letters*, vol. 5, no. 6, pp. 146–149, June 1998.

[45] V. Krishnamurthy and J. B. Moore, "On-line estimation of hidden Markov model parameters based on the Kullback-Leibler information measure," *IEEE Transactions on Signal Processing*, vol. 41, no. 8, pp. 2557–2573, August 1993.

[46] E. E. Kuruoglu, C. Molina, and W. J. Fotzgerald, "Approximation of a α-stable probability densities using finite mixtures of Gaussians ," in *Proceedings EUSIPCO*, Rhodes, Greece, September 1998.

[47] R. E. Lawrence and H. Kaufman, "The Kalman filter for the equalization of a digital communications channel," *IEEE Transaction on Communication Technology*, vol. COM-19, no. 6, pp. 1137–1141, December 1971.

[48] J. Lee and H. Ko, "Tracking of mobile phone using IMM in CDMA environment," in *Proceedings IEEE ICASSP*, Salt Lake City, UT, USA, pp. 2829–2832, May 2001.

[49] Z. Liu, X. Ma, and G. B. Giannakis, "Space-time coding and Kalman filtering for time-selective fading channels," *IEEE Transactions on Communications*, vol. 50, no. 2, pp. 183–186, February 2002.

[50] S. Marcos, "A network of adaptive Kalman filters for data channel equalization," *IEEE Transactions on Signal Processing*, vol. 48, no. 9, pp. 2620–2627, September 2000.

[51] D. Middleton, "Man-made noise in urban environments and transportation systems: Models and measurements," *IEEE Transactions on Communications*, vol. COM-21, no. 11, pp. 1232–1241, November 1973.

[52] D. Middleton, "Channel modeling and threshold signal processing in underwater acoustics: An analytical overview," *IEEE Journal of Oceanic Engineering*, vol. OE-12, no. 1, pp. 4–28, January 1987.

[53] D. Middleton, "Non-Gaussian noise models in signal processing for telecommunications: New methods and results for class A and class B noise models," *IEEE Transactions on Information Theory*, vol. 45, no. 4, pp. 1122–1129, May 1999.

[54] B. Mulgrew and C. F. N. Cowan, "An adaptive Kalman equalizer: Structure and performance," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-35, no. 12, pp. 1727–1735, December 1987.

[55] A. F. Naguib, N. Seshadri, and A. R. Calderbank, "Increasing data rate over wireless channels," *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp. 76–92, May 2000.

[56] T. Öberg and M. Mettiji, "Robust detection in digital communications," *IEEE Transaction on Communications*, vol. 43, no. 5, pp. 1872–1876, May 1995.

[57] J. G. Proakis, *Digital Communications,* 3rd ed. New York: McGraw Hill, 1995.

[58] J. G. Proakis, "Adaptive equalization for TDMA digital mobile radio," *IEEE Transactions on Vehicular Technology*, vol. 40, no. 2, pp. 333–341, May 1991.

[59] R. Pupeikis, "On the optimization of an adaptive state estimation of dynamic systems in the presence of time-varying outliers," *Informatica*, vol. 13, no. 1, pp. 89–104, 2002.

[60] T. Sathyan and T. Kirubarajan, "Secure communication using chaotic systems and Markovian jump systems," in *Proceedings IEEE ICSMC*, Washington D. C., USA, pp. 1932–1937, October 2003.

[61] N. Seshadri and J. H. Winters, "Two signaling schemes for improving the error performance of frequency-division-duplex (FDD) transmission systems using transmitter antenna diversity," in *Proceedings IEEE VTC'93*, Secaucus, NJ, pp. 508–511, May 1993.

[62] N. Seshadri, "Joint data and channel estimation using blind trellis search techniques," *IEEE Transaction on Communications*, vol. 42, no. 2/3/4, pp. 1000–1011, February/March/April 1994.

[63] M. Shao and C. L. Nikias, "Signal processing with fractional lower order moments: Stable processes and their applications," *Proceedings of the IEEE*, vol. 81, no. 7, pp. 986–1010, July 1993.

[64] G. L. Stüber, *Principles of Mobile Communications.* Kluwer Academic Publisher, 1996.

[65] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, March 1998.

[66] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block coding for wireless communications: Performance results," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 3, pp. 451–460, March 1999.

[67] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, July 1999.

[68] M. K. Tsatsanis, G. B. Giannakis, and G. Zhou, "Estimation and equalization of fading channels with random coefficients," *Signal Processing*, vol. 53, no. 2/3, pp. 211–228, 1996.

[69] J. K. Tugnait, L. Tong, and Z. Ding, "Single-user channel estimation and equalization," *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp. 16–28, May 2000.

[70] H. S. Wang and P.-C. Chang, "On verifying the first-order Markovian assumption for a Rayleigh fading channel model," *IEEE Transactions on Vehicular Technology*, vol. 45, no. 2, pp. 353–357, May 1996.

[71] J. H. Winters, "The diversity gain of transmit diversity in wireless systems with Rayleigh fading," in *Proceedings IEEE ICC'94*, New Orleans, LA, vol. 2, pp. 1121–1125, May 1994.

[72] A. Wittneben, "Basestation modulation diversity for digital SIMULCAST," in *Proceedings IEEE VTC'91*, St. Louis, MO, vol. 1, pp. 848–853, May 1991.

[73] A. Wittneben, "A new bandwidth efficient transmit antenna modulation diversity scheme for linear digital modulation," in *Proceedings IEEE ICC'93*, Geneva, Switzerland, vol. 3, pp. 1630–1634, May 1993.

[74] S. M. Zabin and H. V. Poor, "Efficient estimation of the class A parameters via the EM algorithm," *IEEE Transactions on Information Theory*, vol. 37, no. 1, pp. 60–72, January 1991.