ADDIS ABABA UNIVERSITY
SCHOOL OF GRADUATE STUDIES
FACULTY OF TECHNOLOGY
DEPARTMENT OF
ELECTRICAL AND COMPUTER ENGINEERING

Channel Estimation for Wireless MIMO-OFDM Systems Using Kalman Filtering

A thesis submitted to the school of graduate studies of Addis Ababa University
in partial fulfillment of the requirements for the degree of
Masters of Science in Electrical Engineering

By
Mulugeta Araya

Advisor: Dr.-Ing. Hailu Ayele

August 2007

ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

"Channel Estimation for Wireless MIMO-OFDM systems using Kalman Filtering"

By

Mulugeta Araya

FACULTY OF TECHNOLOGY

APPROVAL BY BOARD OF EXAMINERS

_____	_____
Chairman Department of Graduate Committee	Signature
_____	_____
Advisor	Signature
_____	_____
Internal Examiner	Signature
_____	_____
External Examiner	Signature

Declaration

I, the undersigned, declare that this thesis work is my original work, has not been presented for a degree in this or any other universities, and all sources of materials used for the thesis work have been fully acknowledged.

Name: Mulugeta Araya Signature: _____

Place: Addis Ababa

Date of submission:

This thesis has been submitted for examination with my approval as a university advisor.

Dr. Ing. Hailu Ayele Signature: _____

Advisor's Name

CHAPTER ONE
INTRODUCTION

1.1 Overview

Wireless communication has become an essential part of everyday communication for the last two decades. The growing acceptance and dependence on wireless communication has driven the need for higher data rates and stricter reliability requirements. To support emerging multimedia applications researchers must look for better ways of exploiting the limited radio spectrum while maintaining the quality of service required by users. For these and other requirements of the wireless communication systems Third generation (3G) wireless services are currently being deployed. While 3G networks outperform their second generation (2G) predecessors, they still may not be sufficient to meet the requirements for future high data rate applications like multimedia, video streaming, wireless teleconferencing and web browsing. Hence, there exists the need for a wireless network technology that improves 3G throughput performance by one order of magnitude at least. Major requirements of future beyond third generation (B3G) and fourth generation (4G) wireless communication systems are thus higher user data rates, improved coverage and spectral efficiency, as well as enhanced user mobility. Low latency and improved radio link quality are also desired. The future trend is the convergence to all digital IP-based packet networks, with both voice and data capabilities. All those enhancements will allow bringing to mobile users Internet services and multimedia applications that consume lot of resources [2]. Recent developments suggest that the use of multiple transmit and receive antennas can significantly enhance the performance and reliability of a wireless communication system. Multiple input and multiple output (MIMO) systems take the advantage of spatial diversity obtained through spatially separated antennas in a dense multipath scattering environment [1]. Theoretical studies indicate that the capacity of MIMO systems grows linearly with the number of transmit antennas used. The multiple antennas configuration exploits the multipath effect to accomplish the additional spatial diversity. However, the multipath effect also causes the negative effect of frequency selectivity of the channel. Orthogonal Frequency

Division Multiplexing (OFDM) is a promising multi-carrier modulation scheme that shows high spectral efficiency and robustness to frequency selective channels. In OFDM, a frequency-selective channel is divided into a number of parallel frequency-flat subchannels (narrowband channels), thereby reducing the receiver signal processing of the system. OFDM has been adopted in many wireless standards such as digital audio broadcasting (DAB) , digital video broadcasting (DVB) , HIPERLAN/2 , IEEE 802.11a wireless local area networks (WLAN) , IEEE 802.16a metropolitan area network (MAN) , and a potential candidate for fourth-generation (4G) mobile wireless systems. The combination of OFDM and MIMO is a promising technique to achieve high bandwidth efficiencies and system performance. In fact, MIMO-OFDM is being considered for the upcoming IEEE 802.11n standard, a developing standard for high data rate WLAN [2].

1.2 Motivations

MIMO-OFDM has the potential to meet the increasing high speed and reliability demands of the future. In order for this technology to truly succeed in commercial deployment there are still several technical obstacles that must be tackled A major impediment in MIMO-OFDM is the complicated receiver signal processing. For example, the complexity of a maximum likelihood detector increases exponentially with the number of transmit antennas. Spatial equalizers and space-time coding has been proposed to simplify the detection for MIMO-OFDM systems [3]. Note that coherent detection requires knowledge of the channel; therefore, accurate channel estimation is crucial in realizing the full potential of MIMO-OFDM. Channel estimation for OFDM has been well researched in literature. The extension of the results to MIMO-OFDM channel estimation is substantially more complicated. In a MIMO system, multiple channels have to be estimated simultaneously. The increased number of channel unknowns significantly increases the efficiency and computational complexity of the channel estimation algorithm. Previous works have investigated the problem of channel estimation in MIMO-OFDM [4]. The most common approach is training-based estimation, where a known pilot sequence is transmitted and used at the receiver to determine the channel. The least square (LS) approach and the minimum

mean square error (MMSE) approach are the usual methods for training-based estimation. The LS and MMSE solutions are relatively simple compared to other estimation techniques such as blind estimation. However, both solutions still require complex matrix inversions, which are undesirable in real time implementation. In [5], specific training sequences design and pilot placement patterns are used to obtain the channel frequency response (CFR) of the channel in attempt to reduce the estimation complexity. Note that the number of unknowns of the CFR is usually significantly greater than the number of unknowns in the channel impulse response (CIR), which the writer in the second reference (Timo Roman) also agrees. In [8], it is proven that computational complexity can be reduced by estimating the CIR as opposed to the CFR. The proposed solution reduces the number of unknowns to be solved, but the solution still requires a less complex, and efficient methods. In the thesis, the problem of tracking channel coefficients in MIMO-OFDM systems is addressed. The main objective of our research is to explore methods for reducing the complexity of the channel estimation for MIMO-OFDM and an efficient one. We propose the use of Kalman Filtering to solve for the channel unknowns. In using Kalman filtering we exploit many of its outstanding properties in dynamic systems or time-varying systems, which suit our system (see chapter four). Hence the channel estimation in MIMO-OFDM system will be done in time domain.

1.3 Thesis Contribution

The main contribution of this thesis includes the following.

- A Kalman filtering method is used for tracking of the channel and also MSE performance for a time domain MIMO-OFDM channel. High efficiency was achieved as compared to LS and MMSE estimators.
- The Kalman filter based channel estimation was investigated for computational complexity and made open for researchers to work on complexity reduction with possible comment.
- A less number of recursion to track the channels is also considered to be a merit for the thesis work. This is due to the modeling of the method with Gauss –Markov model and from the properties of Kalman filtering.

1.4 Thesis Outline

In chapter 2, wireless communications channels will be dealt. The chapter also briefs about components of our system OFDM and MIMO systems. The important characteristics of these two systems will be emphasized which helps us to appreciate their capability of fulfilling the demands of wireless communication. Finally presentation of the complete MIMO-OFDM system model is presented.

In chapter 3, literature survey of previous works on channel estimation for OFDM and MIMO-OFDM systems is presented. Estimation theories are discussed. Derivation of LS and MMSE estimators are also presented. These results are adapted to the OFDM and MIMO-OFDM channel estimation. Performance analysis will be done for LS and MMSE methods for estimating the CIR in the basis of MSE for MIMO-OFDM systems and computational complexity analysis.

In chapter 4, review of Kalman filtering is provided. The different kinds of Kalman filtering and their application are also presented. In addition the Gauss-Markov process is briefed along with the vector Kalman filtering equations. Properties of Kalman filtering (equations) and their applications to our system are discussed.

In chapter 5, Kalman filtering channel estimation for MIMO-OFDM systems is presented. The assumptions and simulation parameters are also presented. The Kalman tracking of the channels in time domain are displayed. With this, we show the tracking capability of the Kalman filters. Comparison of the LS and MMSE estimators with Kalman filters in the sense of MSE is also shown computational complexity for Kalman filtering is also shown.

In chapter 6, conclusion of the work is presented along with some possible future works suggestions to point out extensions of this research.

CHAPTER TWO

BACKGROUND

In this chapter, basic wireless channel propagation theory is provided. The impacts of small scale-fading, which our system will rely on, on wireless channel, are discussed. In section 2.2 an introduction to MIMO system is given. In section 2.3 we discuss about fundamentals of OFDM modulation. Finally in section 2.4 system model for MIMO - OFDM is described.

2.1 Wireless Channel Properties

Signal multipath occurs when the transmitted signal arrives at the receiver via multiple propagation paths. Each of these paths may have a separate phase, attenuation, delay and Doppler frequency associated with it. This propagation paths places fundamental limitations on the performance of the wireless communication system because the transmitted signal travels through different paths and interact with objects in the environment. These interactions include reflection, refraction, diffraction, and scattering which cause attenuation and variations in the received signal power and phase shift of the transmitted signal. Due to this random phase shift associated with each received signal, they might add up destructively, resulting a phenomenon called

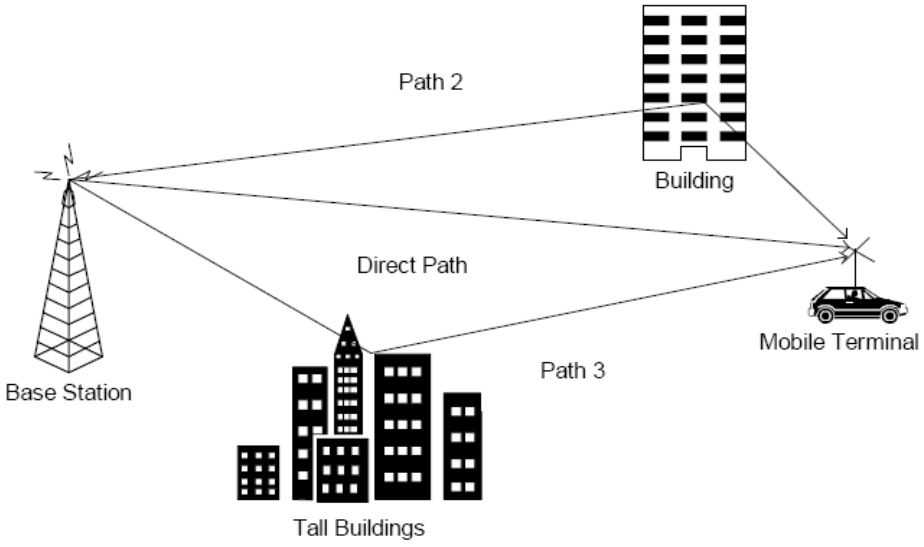


Figure 2.1: Multipath Effects in an Urban Scenario

fading [5]. The Doppler shift, which is the relative movement between the transmitter and the receiver, also impact the fading characteristics of the signal.

The effects of the wireless environment can be categorized as path loss or attenuation, large-scale (long-term fading), and small-scale (short-term) fading.

2.1.1 Multipath Effect and Time Varying Nature of the Channel

In this section we consider channels for time-variant multipath channels. Their characterization serves as a model for signal transmission over many radio channels. The time-variant impulse responses of these channels are a consequence of the constantly changing physical characteristics of the media.

The multipath effect is a phenomenon that causes multiple versions of the transmitted signal to arrive at the receiver at different time delays. Reflecting objects and scatterers in the transmission environment generate multiple versions of the transmitted signal as shown in Figure (2.1). Each of the paths will have different characteristics, such as amplitude, phase, arrive time, and angle of arrival. The multiple signals may constructively or destructively add up at the receiver, thus creating the rapid fluctuations in the received signal envelope. When the signals add up constructively it will increase the signal power at the receiver, but destructive summation will cause fades in the received signal, which corresponds to the sudden drops in received power. Multipath does not only cause fluctuations in the received power, but it also affects the shape of the pulse as it is transmitted through the channel. The arrival of the multiple versions will broaden the transmitted signal. As illustrated in Figure (2.2), the transmitted signal arriving at different times will overlap with each other and lead to broadening of the envelope of the pulse.

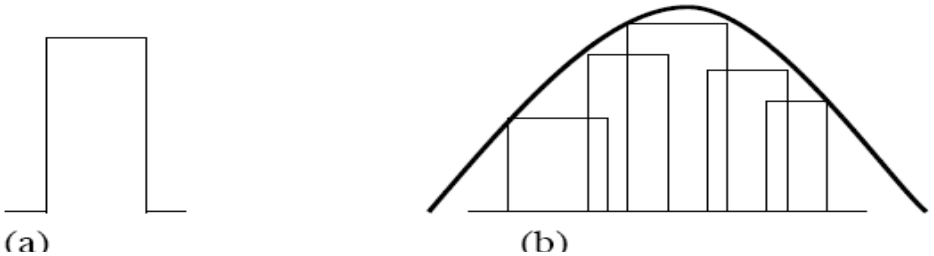


Figure 2.2 (a) transmitted signal (b) Multiple copies of the transmitted signals

The signal power and arrival times of the multipath signals are used to characterize the channel. If for example, we transmit an extremely short pulse, ideally an impulse, over a time-varying multipath channel the received signal may appear as a train of pulses, as shown in Figure (2.3). Hence, one characteristics of a multipath medium is the time spread introduced in the signal that is transmitted through the channel.

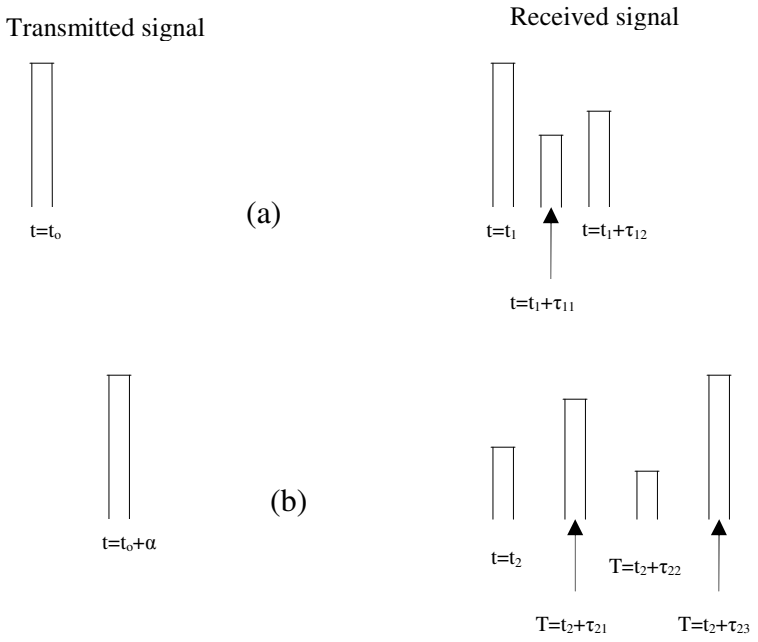


Figure 2.3 Example of the response of a time-invariant multipath channel to a very narrow pulse.

A second characteristic is due to the time variations in the structure of the medium. As a result of such time variations, the nature of the multipath varies with time [26]. That is, if we repeat the pulse-surrounding experiment over and over, we shall observe changes in the received pulse train, which will include changes in the sizes of the individual pulses, changes in the relative delays among the pulses, and, quite often, changes in the number of pulse as shown in Figure (2.3). Moreover, the time variations appear to be unpredictable to the user of the channel. Therefore it is reasonable to characterize the time-variant multipath channel statistically (see section 2.1.3).

2.1.2 Doppler Shift

Due to the relative motion between the transmitter and the receiver, each multipath wave is subjected to a shift in frequency. The frequency shift of the received signal caused by the relative motion is called the *Doppler shift*. It is proportional to the speed of the mobile unit. Consider a situation when only a single tone of frequency f_c is transmitted and received signal consists of only one wave coming at an incident angle θ with respect to the direction of the vehicle motion as in Figure (2.4). The Doppler shift of the received signal, denoted by f_d, is given by

$$f_d = \frac{vf_c}{c}\cos\theta \quad \text{.....2.1}$$

where v is the vehicle speed and c is the speed of the light and f_c is the carrier frequency. The Doppler shift in a multipath propagation environment spreads the bandwidth of the multipath waves within the range of $f_c \pm f_{dmax}$ where f_{dmax} is the maximum Doppler shift, given by

$$f_{dmax} = \frac{vf_c}{c} \quad \text{.....2.2}$$

The maximum Doppler shift is also referred as the maximum *fade rate.* As a result, a single tone transmitted gives rise to a received signal with a spectrum of nonzero width. This phenomenon is called *frequency dispersion* of the channel [6].

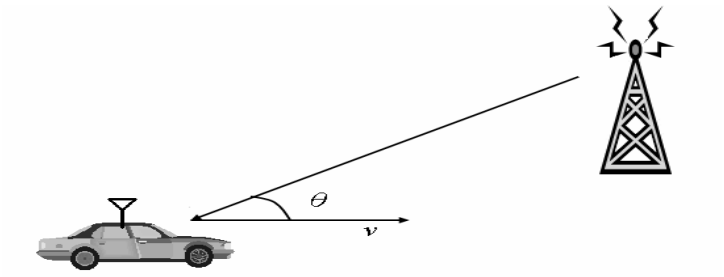


Figure 2.4: A receiver moving at a velocity v, thus causing a Doppler shift

2.1.3 Statistical Models for Fading Channels

Because of the multiplicity of factors involved in propagation in mobile transceiver environment, it is convenient to apply statistical techniques to describe signal variations.

In a narrow band system, the transmitted signal usually occupy a bandwidth smaller than the channel's *coherence bandwidth*, which is defined as the frequency range over which the channel fading process is correlated. That is, all spectral components of the transmitted signal are subject to the same fading attenuation. This type of fading is referred to as *frequency nonselective* or *frequency flat*. On the other hand, if the transmitted signal bandwidth is greater than the channel coherence bandwidth, the spectral components of the transmitted signal with a frequency separation larger than the coherence bandwidth are faded independently. The received signal spectrum becomes distorted, since the relationships between various spectral components are not the same as in the transmitted signal. This phenomenon is known as *frequency selective* fading. In wideband systems, the transmitted signals usually undergo frequency selective fading. In the next section we will see some points about Rayleigh Fading channel, because among the different models that are used in wireless fading channels, Rayleigh fading is going to be used by our system, for the reasons that will be discussed shortly.

2.1.3.1 Rayleigh Fading

Consider the transmission of a single tone with constant amplitude. In a typical land mobile radio channel for example, we may assume that the direct wave is obstructed and the mobile unit receives only reflected waves. When the number of reflected waves is large, according to the central limit theorem, two quadrature components of the received signal are uncorrelated Gaussian random process with zero mean and variance σ_s^2. As a result, the envelope of the received signal at any time instant undergoes a Rayleigh probability distribution and its phase obeys a uniform distribution between $-\pi$ and π. The probability density function of the Rayleigh distribution is given by

$$p(a) = \begin{cases} 0 & \\ \dfrac{a}{\sigma_s^2} e^{-a^2/2\sigma_s^2} & \end{cases} \qquad a<0 \text{.....} 2.3$$

The mean value, denoted by m_a, and the variance, denoted by σ_a^2, of Rayleigh distributed random variable are given by

$$m_a = \sqrt{\frac{\pi}{2}}\sigma_s = 1.2533\sigma_s$$
$$\sigma_a^2 = \left(2 - \frac{\pi}{2}\right)\sigma_s^2 = 0.4292\sigma_s^2 \quad \text{.....2.4}$$

If the probability density function in Equation (2.3) is normalized so that the average signal power ($E[a^2]$) is unity, then the normalized Rayleigh distribution becomes

$$p(a) = \begin{cases} 2ae^{-a^2}, a < 0 \\ 0, a \geq 0 \end{cases} \quad \text{.....2.5}$$

The mean and the variance are

$$m_a = 0.8862$$
$$\sigma_s^2 = 0.2146 \quad \text{.....2.6}$$

In fading channels with a maximum Doppler shift of f_{dmin}, the received signal experiences a form of frequency spreading and is band-limited between $f_c \pm f_{dmin}$. Assuming an omni-directional antenna with waves arriving in the horizontal plane, a large number of reflected waves and a uniform received power over incident angles, the power spectral density of the faded amplitude, denoted by $|P(f)|$, is given by

$$|P(f)| = \begin{cases} \dfrac{1}{2\pi\sqrt{f_{d_{\max}}^2 - f^2}} & if\,|f| \leq \left|f_{d_{\max}}^2\right| \\ 0, otherwise \end{cases} \quad \text{.....2.7}$$

Where f is the frequency and f_{dmax} is the maximum fade rate. The value of $f_{\mathrm{dmax}}\ T_{\mathrm{s}}$ is the maximum fade rate normalized by symbol rate. It serves as a measure of channel memory. For correlated fading channels this parameter is in the range $0 < f_{dmax} < T_s < 1$, indicating a finite channel memory. And lastly the autocorrelation function of the fading process is given by

$$R(\tau) = J_0(2\pi f_{d_{\max}}\tau) \quad \text{.....2.8}$$

where $J_0(\cdot)$ is the zero-order Bessel function of the first kind [6].

2.2 Introduction to MIMO

A MIMO communication system uses multiple antennas at the transmitter and receiver to achieve various advantages. Traditionally, antenna arrays have been used at the transmitter and the receiver to achieve array gain, which increases the output SNR of the system. In the mid-1990s, adaptive antennas and smart antennas were introduced to describe antenna arrays that are made adaptive in a manner that it changes its transmission and reception characteristics when the radio environment changes. Array antennas have been implemented in GSM networks, fixed broadband wireless access (BWA) networks, and third generation (3G) CDMA networks. More recently, a way of using multiple antennas has been discovered to achieve diversity and multiplexing gain by exploiting the once negative effect of multipath. Under suitable conditions, i.e. a scatter rich environment, the channel paths between the different transmit and receive antennas can be treated as independent channels due to the multipath effects caused by the scatterers. Initial works in this research area, suggests that MIMO effectively takes advantage of the random fading and multipath delay spread to increase the transfer rate of the system. The exploitation of this additional 'spatial' degree of freedom can increase the throughput and improve the performance of the system. In summary, the main advantages of MIMO system can be categorized as array gain, diversity gain, and spatial multiplexing gain.

2.2.1 Array Gain

Array gain is achieved by coherently combining the signals from the multiple antennas to increase the average output to signal ratio (SNR), which will improve the range and coverage of the system. Figure (2.5) illustrates a simple case of a system consisting of one transmit antenna and a set of receive antenna array. Assume the distance between the transmitter and the receiver is significantly larger than the antenna separation of the array at the receiver, and then the received signal at each antenna will differ in phase due to the relative delay caused by the antenna separation. To maximize the received signal energy, an optimum receiver will be beamforming techniques to adjust for different delays of the multiple antennas so that the received signal can be constructively combined. This will yield M_r-fold power gain, where M_r is the number of

receiver's antennas. In a MIMO case where antenna arrays are used at the transmitter and receiver, then an $M_t M_r$ –fold power gain is available; where M_t is the number of transmit antennas.

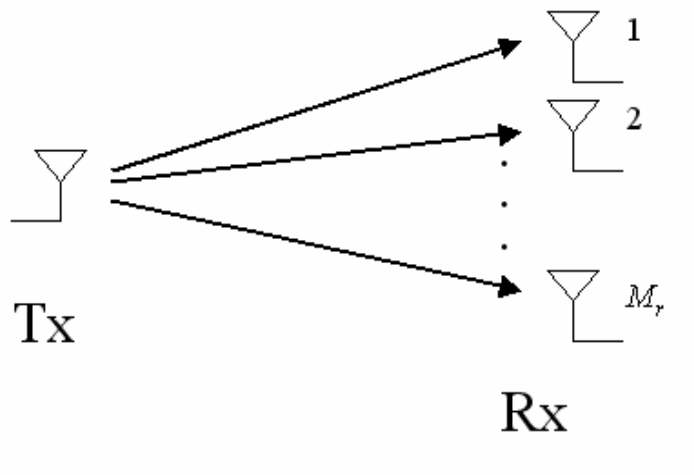


Figure 2.5: Single transmit and multiple receive antenna system.

2.2.2 Diversity Gain

Diversity is a technique to mitigate fading in wireless links by transmitting a signal over multiple independently fading paths. The main idea behind diversity is that by transmitting multiple copies of the signal increases the probability that at least one copy is recovered correctly at the receiver. Diversity can be obtained through time, frequency, or space. Time diversity assumes that a signal will experience independent fading at different times due to changes in the channel. The benefit of time diversity is that it does not require additional hardware. However it requires memory storage of the repeated signals for processing. Frequency diversity is achieved when the carrier frequencies are sufficiently separated such that each carrier frequency will experience independent fading. In frequency diversity, the same signal is transmitted on various independent carrier frequencies. Multiple receivers are used to detect the multiple signals at different frequencies, and the one with the highest energy will be selected. Alternatively, a multi-antenna system can exploit the independent multipath channels to achieve spatial diversity. The diversity order is the total number of the independent fading signal paths between the transmitter and the receiver, and depends on the spatial separation of the antennas and the scatterers of the environment. The maximum spatial diversity gain of a

MIMO system is M_tM_r. Joint diversity schemes such as space-time and space-frequency coding at the transmitter and the receiver has been developed to increase the diversity order of the system. In 1998, independent pioneer work by Alamouti and Tarokh et al. developed a breakthrough space-time diversity system that provides the diversity gain without sacrificing the bandwidth.

2.2.3 Spatial Multiplexing Gain

Multiplexing gain is achieved through transmitting different signals on independent channels in a MIMO system. The multiplexing gain order is the number of parallel independent spatial data pipes in the same frequency band between the transmitter and receiver. As shown in Figure (2.6), a signal is split into two parts and transmitted on two separate antennas. At each receive antenna it will detect a signal from a specific transmit antenna and the signals from other antennas will be seen as interference. Combining techniques are required at the receiver to eliminate the interference and to multiplex the signal back together. As demonstrated, capacity gain is achieved by reducing the transmission time without using additional bandwidth [1].

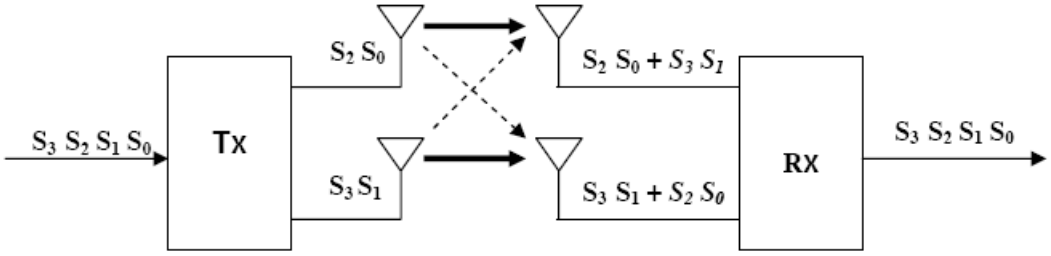


Figure 2.6: Multiplexing gain 2X2 MIMO system [1].

In summary, the use of MIMO has many benefits. However, it is not possible to achieve all the above benefits of MIMO techniques in one system as some of them are mutually conflicting goals. In general, a MIMO system improves

- Spectral efficiency: multiplexing gain
- Link reliability: diversity gain
- Coverage: Diversity gain and array gain
- Capacity: Multiplexing gain

2.3 Introduction to OFDM

In recent years, OFDM has gained considerable attention as it shows to be a promising technique for high data rate transmission. In OFDM, a broadband signal is split into multiple parallel narrowband signals, and then modulated onto orthogonal subcarriers for transmission. One of the most attractive features of OFDM is its robustness against frequency selective channels. The OFDM operation converts a frequency selective channel into multiple parallel flat fading channels, which greatly simplifies the channel estimation and equalization tasks of the receiver. When a wideband signal passes through a frequency selective channel as shown in Figure (2.7a), a significant portion of the signal is lost due to the deep fades in the channel. However, when the wideband signal is OFDM modulated, the frequency spectrum will be a composition of overlapping narrowband signals as shown in the Figure (2.7b). Now, when the OFDM modulated signal passes through the frequency selective channel only the narrowband signal at the location of the deep fades will be affected. In addition, it can be observed that each of the narrow band signals experiences flat fading; therefore, the channel response can be obtained simply by dividing the output of the signal by the input signal. Moreover, OFDM is bandwidth efficient since the subchannels can overlap yet still be separated due to the use of orthogonal subcarriers. With the current advancements in digital signal processor (DSP) and integrated circuit (IC) technology OFDM can be efficiently implemented by using the inverse fast Fourier transform (IFFT) and fast Fourier transform (FFT) for modulation and demodulation, respectively.

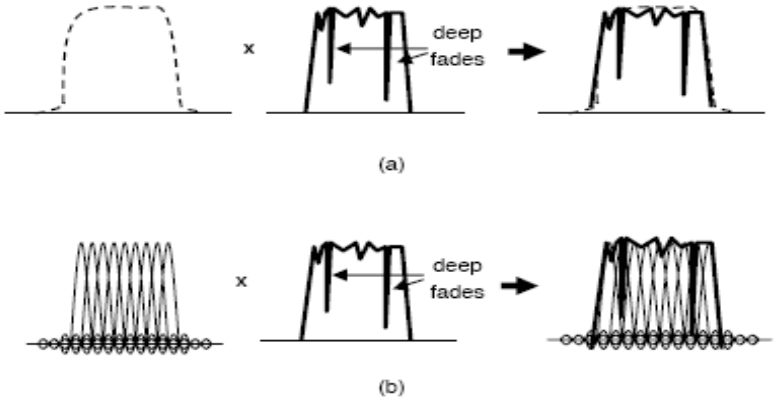


Figure 2.7: (a) a wide band signal multiplied with frequency selective channel. (b) An OFDM signal with a frequency selective channel.

A schematic diagram of the complete structure of an OFDM system is shown in Figure (2.8). The blocks on the top row correspond to components in the transmitter and the bottom row to the receiver. The input data stream is modulated using regular modulation techniques such as phase shift keying (PSK) of quadrature amplitude modulation (QAM). The modulated signal $X(n)$ ($n = 0, 1, \ldots, N-1$, where N is the number of subcarriers) is converted into parallel signals and passed to IFFT block. The IFFT operation modulates the parallel signals onto orthogonal subcarriers as a group. The narrow band signals outputted are $x(k)$ ($k = 0,1, \ldots, N-1$), where

$$x(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} X(n) e^{j2\pi kn/N} \quad , 0 \le k \le N-1 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 2.9$$

When an OFDM signal passes through the channel it will experience intersymbol interference ISI and interchannel interference (ICI). The ISI arises from channel delay spread and ICI is caused by the loss of orthogonality of the subcarriers due to the frequency response of the channel. In order to eliminate the effects of ISI a cyclic prefix (CP) of length N_{CP}, where N_{CP} is greater than the channel order, is appended to the beginning of the signal. The total length of the signal becomes $N+N_{CP}$. Repeating the last elements at the beginning converts the linear convolution of the channel into circular convolution thereby preserving the orthogonality of the subcarriers. This fact will be used in chapter five as an alternative representation of cyclic prefix insertion and removal in the system model equation. At the receiver the inverse operations are performed to recover the transmitted bits.

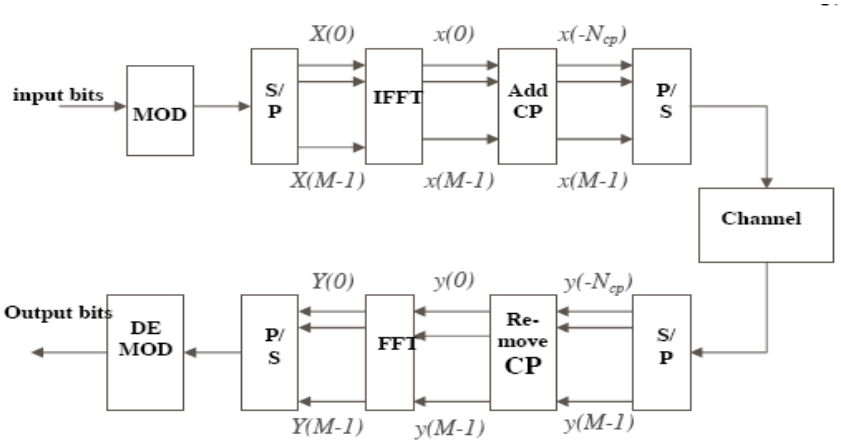


Figure 2.8: Schematic diagram of OFDM system.

2.4 MIMO-OFDM System Model

In the last two sections we have seen the two most important elements of our system. It is said that OFDM is a promising physical layer technology for high data rate wireless communication due to its effectiveness in frequency selective fading, high spectral efficiency, and low computational complexity. MIMO systems have the ability to improve spectral efficiency, link reliability, coverage or capacity depending on how the system is implemented. Clearly, OFDM integrated with MIMO transceivers will further enhance the performance and throughput of the system. Now we shall see these two naturally combined systems and investigate their input-output relationship based on their system model.

A typical MIMO-OFDM system is shown in Figure (2.9). The system consists of M_t transmit antennas and M_r receive antennas. At time t, a block of binary input data stream is modulated, and then passed through the MIMO encoder to produce M_t data streams for transmission over the multiple antennas. The data can be space-time coded for diversity gain, or de-multiplexed for spatial multiplexing gain. Each of the M_t data streams is grouped into blocks of N symbols, and then OFDM modulated for transmission across the MIMO channels. The received signal at each antenna will be a summation of all the signals from the multiple paths plus the noise. The noise process is additive white Gaussian noise (AWGN) with zero mean and variance σ_n^2. It is assumed that the signal and noise are independent of each other, which is a common assumptions made in literature.

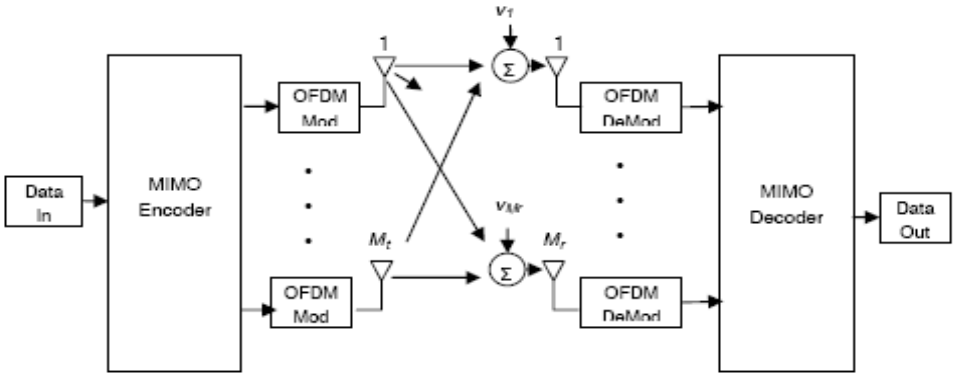


Figure 2.9: MIMO-OFDM system

2.4.1 Channel Model

The channel considered in this thesis is a time domain Rayleigh fading channel. The radio channel can be modeled as a linear filter with a time varying impulse response. Characterizing the channel as an impulse response will provide the necessary information and analysis of the signal transmission through the channel. The complex baseband representation of the mobile wireless channel impulse response for a fading multipath channel can be modeled as

$$h(t,\tau) = \sum_{k} \gamma_k(t) c(\tau - \tau_k) \quad \text{.....2.10}$$

Where τ_k is the delay of the k^{th} path, $\gamma_k(t)$ is the corresponding complex amplitude, and $c(t)$is the shaping pulse whose frequency response is usually a square-root raised-cosine Nyquist filter. The value of k ranges for the length of the channel i.e. $0 \leq k \leq L-1$, where L is the length of the channel. Due to the motion of the vehicle, $\gamma_k(t)$'s are wide-sense stationary (WSS) narrow-band complex Gaussian processes, which are independent for different paths. From Equation (2.10), the frequency response at time t is

$$H(t,f) = \int_{-\infty}^{\infty} h(t,\tau) e^{-j2\pi f\tau} d\tau \quad \text{.....2.11}$$

In an OFDM system, the channel frequency response can be expressed as

$$H[n,k] = H(nT_f, k\Delta f) = \sum_{l=0}^{K-1} h[n,l] W_K^{kl} \quad \text{.....2.12}$$

Where $W_K = \exp(-j2\pi / K)$, $h[n,l] = h[nT_f, k(Ts/K))$, and K is the number of tones of the OFDM block. T_f, T_s, and Δf in the above equations are block length, symbol duration and tone spacing respectively. They are related by $T_s = 1/\Delta f$ and $T_f = T_s + T_g$, where T_g is the duration of the cyclic extension. In general wireless communication channels are time-varying either due to the motion of the transmitter or receiver, or the wireless environment. However, in indoor environments under lower mobility scenarios the

channel variations are slow and can be ignored for the duration of the packet. Then the above equation becomes

$$H[k] = \sum_{l=0}^{L-1} h[l] e^{-j2\pi kl/K} \quad \text{.....2.13}$$

From here onwards vectors in frequency domain are represented by capital letters and their matrix bold and capital. The time domain representation will be done in small letters accordingly.

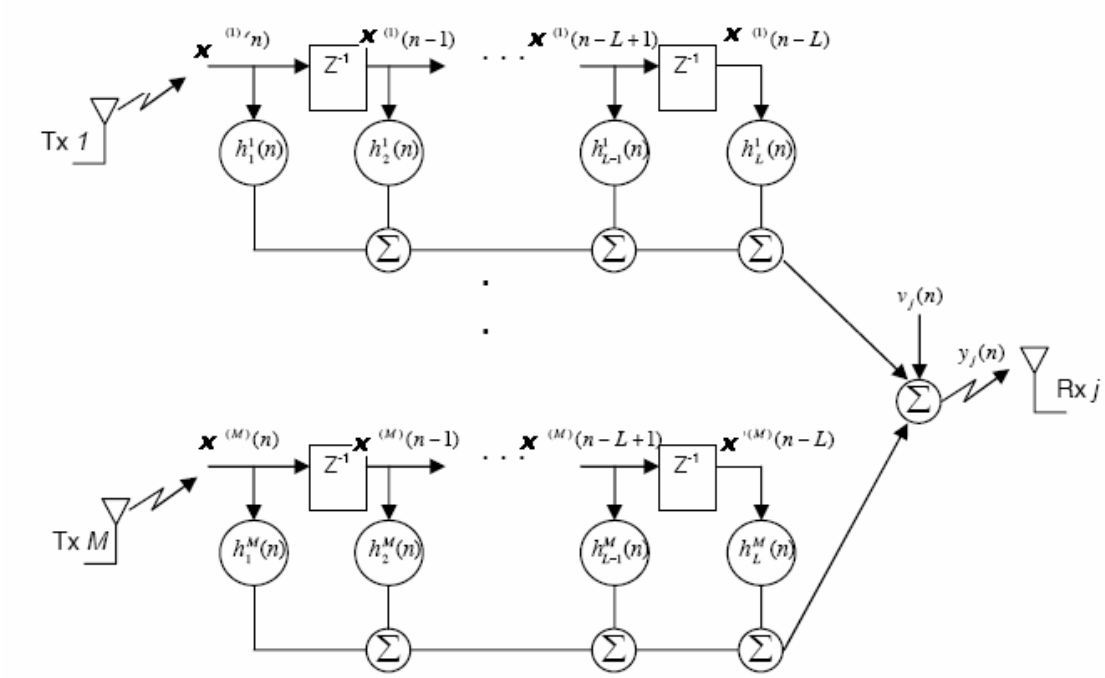


Figure 2.10: Tap delay line model of MIMO channel

2.4.2 Mathematical Model of the system

Consider the MIMO-OFDM systems, diagrammed in Figure (2.9), consisting of M_t transmit and M_r receive antennas operating in a Rayleigh fading environment. The tap-delay channel model of the multiple channels is shown in Figure (2.10). Let $\mathbf{X}_t^{(i)}$ be the frequency domain input block of N complex-valued symbol prior to OFDM processing at time t at the i^{th} transmit antenna, where $\mathbf{X}_t^{(i)} = [\mathbf{x}_t^{(i)}(0), \ \mathbf{x}_t^{(i)}(1), \ldots, \mathbf{x}_t^{(i)}(N+N_{CP}-1]^{\mathrm{T}}$, i =

1,2, … M_t and $\mathbf{x}_t^{(i)}$ is the time domain representation after the IFFT operation. With N_{CP} CP symbols inserted, where $\mathbf{x}_t^{(i)} = [\mathbf{x}_t^{(i)}(0), \mathbf{x}_t^{(i)}(1), \ldots, \mathbf{x}_t^{(i)}(N+N_{CP}-1)]^{\mathrm{T}}$, $i = 1,2, \ldots M_t$.

From the model and linear input output relationships we can have the following equations for the MIMO channel.

The time domain received signal at the j^{th} antenna before OFDM demodulation as shown in Figure (2.7) is represented as

$$y_t^{(j)}(n) = \sum_{i=1}^{M_t} \sum_{l=0}^{L_{ji}} h^{(j,i)}(l) x_t^{(i)}(n-l) + v_t^j(n) \text{ , } n = 0, 1, \ldots, (N+N_{CP})-1 \quad \ldots\ldots 2.14$$

Where $h^{(j,i)}(l)$ $(l = 0,1, \ldots, L_{ji})$ is the channel response between the i^{th} transmitter and the j^{th} receiver, and L_{ji} is the channel order and $v_t^{(j)}(n)$ is the AWGN. By changing the summation order of equation (2.14) and taking its vector form we have

$$y_t(n) = \sum_{l=0}^{L-1} h^{(i)}(l) x_t(n-l) + v_t(n) \text{ , } n = 0, 1, \ldots, (N+N_{CP})-1 \quad \ldots\ldots\ldots 2.15$$

Where $h(l) = [h^{(1,i)}(l), h^{(2,i)}(l), \ldots, h^{(M_t,i)}(n-l)]^T$ is a matrix of size $(M_r \; X \; M_t)$ and $x_t(n-l) = [x^{(1)}(n-l), x^{(2)}(n-l), \ldots, x^{(M_t)}(n-l)]^H$ is $(M_t \; X \; 1)$ vector. Finally, in matrix form Equation (2.15) can be expressed as

$$y_t = Gx_t + v_t, \quad t = 0,1,2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 2.16$$

Where $y_t = [y_t(0), y_t(1), \ldots, y_t(N+N_{CP}-1)]^T$ is a vector of size $(M_r \; (N+N_{CP}) \; X \; 1)$, G is a block toeplitz matrix constructed by h.

After y_t is OFDM demodulated by the IFFT transformation and the CP is removed, the signal at the receive antenna j^{th} in the frequency domain representation is

$$Y_k^{(j)} = \sum_{i=1}^{M_t} H_k^{(j,i)} X_k^{(i)} + V_k^{(j)}, \; k = 0, 1, \ldots, N-1 \ldots\ldots\ldots\ldots\ldots 2.17$$

where $H_k^{(j,i)}$ the frequency response at the k^{th} subcarrier, $X_k^{(i)}$ is the transmitted signal in the frequency domain. The frequency response of the channel $H_k^{(j,i)}$ is obtained as described in Equation (2.13). The $v_k^{(j)}$ is defined as

$$V_k^{(j)} = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} v^{(j)}(n) e^{-j2\pi kn/K} \quad \text{.....2.18}$$

The vector form equation for all *N* subcarriers can be expressed as

$$Y^{(j)} = \sum_{i=1}^{M_t} X^{(j,i)} H^{(j,i)} + V^{(j)}, \quad \text{.....2.19}$$

Where $Y^{(j)} = [Y_0^{(j)}, Y_1^{j}, ..., Y_{N-1}^{(j)}]^T$ is a vector of size *(N X 1),* $X^{(j,i)} = diag\{X_0^{(j,i)}, X_1^{(j,i)}, ..., X_{N-1}^{(j,i)}\}$ is a diagonal matrix with dimensions *(N X N),* $V^{(j)} = [V_0^{(j)}, V_1^{(j)}, ..., V_{N-1}^{(j)}]^T$ and $H^{(j,i)} = [H_0^{(j,i)}, H_1^{(j,i)}, ..., H_{N-1}^{(j,i)}]$ is *(N X 1)* vector. The received signal vector for all received antennas can be expressed in the following form

$$Y = \boldsymbol{X}H + V \quad \text{.....2.20}$$

CHAPTER THREE
CHANNEL ESTIMATION

In this chapter, a general overview of estimation theory is provided. The problem of channel estimation in communication system, while in attempting it to apply to real systems, is also investigated. Training based least square (LS) estimator and Bayesian minimum mean square (BMMSE) estimator are derived. Finally simulation results for LS and BMMSE estimation for a MIMO-OFDM system together with computational complexity analysis are shown.

In any communication system, the primary goal is to recover the signal transmitted at the transmitter side, at the receiver. In doing this there are a number of equalization /detection methods depending on the diversity or the spatial multiplexing system applied to the MIMO decoding techniques. Regardless of the type of the MIMO system, almost all of the equalization or detection methods require knowledge of the channel information in order to recover the transmitted signal. For an OFDM system with multiple transmit antennas, every tone at each receive antenna is associated with multiple channel parameters, which makes the channel estimation difficult**.** Hence, developing an efficient method of approximating the transmission channel between the transmitter and the receiver is an essential component of the receiver design. There are some studies on joint channel estimation and detection, and blind detection where the channel estimates are not required. Generally, these schemes are higher in complexity or lead to performance loss. In this thesis, we shall only focus on the channel estimation of the receiver design. While doing so we will use MMSE equalizers for some feedbacks in our processes.

3.1 General Channel Estimation Overview

Channel estimation can be categorized into two classes, training-based or blind channel estimation, each with its benefits and disadvantages. In training-based estimation a pilot sequence known to the receiver is embedded into the signal block and transmitted over the channel. At the receiver, the channel is estimated using the received signal and the

known training sequence. Some advantages of training based estimation are high accuracy, relatively lower complexity, and many existing standards such as GSM and IEEE 802.11a have allocated time slots for training sequence transmission. A drawback of training-based estimation is reduced bandwidth efficiency due to the wasteful transmission of training sequence. Blind estimation on the other hand does not require a training sequence; instead it estimates the channel based on solely the received signal. Through the exploitation of the statistical properties of the received signal and channel structure an estimate of the channel is generated. A widely studied blind estimation technique is the subspace method using second order statistics (SOS). In the subspace method, the autocorrelation matrix of the received signal is decomposed into the signal and noise subspace. Due to the orthogonality of the noise and signal subspace, the channel estimates can be calculated based on the noise subspace. There are several issues to be noted when using this blind technique. Firstly, the subspace method requires knowledge of the channel order. Some subspace methods can fail if the channel is over estimated. More importantly, the decomposition of the autocorrelation function via eigenvalue decomposition (EVD) or singular value decomposition (SVD) is highly complex. In real-time systems, these blind estimation techniques cannot be practically implemented due to their intense computational complexity. In addition, blind methods rely on time averaging over a constant period, thus requiring the channel to be slow varying over a large packet of symbols. This makes blind algorithm not suitable for fast varying channels. In general, blind channel estimation techniques are restrictive since it relies on some data or channel assumptions, and it is high in computational complexity. Due to these impracticalities of blind estimation, we focus our research on the more practical technique of training-based channel estimation.

3.2 Estimation Theory

In this section some basic background knowledge is provided on estimation theory. There are two approaches in the theory based on the use of random theories: 1) classical estimation, and 2) Bayesian estimation. In the first the parameter to be estimated is deterministic unknown vector. The estimation is done based on the probability density function (PDF). In the Bayesian estimation we use random theories of the unknown

vector and prior information like mean, variance, and apriori PDF are used to determine the estimate. Estimation problems by their nature, specially these days, are based on given finite data set $x[0]$, $x[1]$, … $x[N-1]$ and then find estimator function that maps data into estimates. Because the estimates are random variables (RV's) they are described by PDF. Hence the general mathematical statement for estimation problem is that for any measured data $\mathbf{x}= [x[0]\ x[1]\ \ldots x[N-1]\ ]$, we have the unknown parameter $\theta = [\theta_1, \theta_2, \ldots, \theta_p]$. Here x is an N-dimensional random data and θ is not random. So what captures all the statistical information to the estimation problem is the N-dimensional PDF of the data parameterized by θ, which is $p(x; \theta)$.

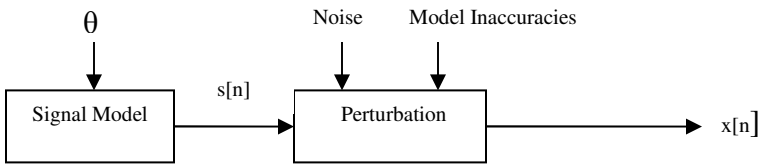
3.2.1 Classical Estimation

In classical estimation, the optimal estimator is one that is unbiased with minimum variance. It provides no way to include *apriori* information about θ. Classical estimators include Minimum Variance (MV), Maximum Likelihood (ML), and Least Squares (LS). As mentioned in [7], in general the minimum variance unbiased estimator (MVUE) does not always exist. A popular suboptimal estimator is the maximum likelihood (ML) estimator which is posed as the maximization of the likelihood function parameterized by the unknown CIR vector. However, a closed-form solution for this problem might not always exit. Even though there are proposed ways to solve this problem the initial estimates largely affects the performance. Hence one of the drawbacks of the ML estimator is that the PDF of the unknown vector must have maximum or else the estimate might not converge. Also, this method only approaches optimally when a large set of the data is available.

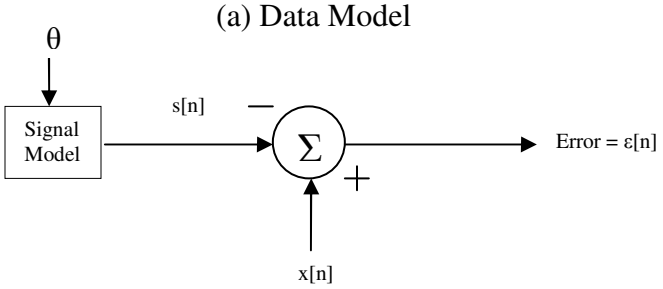
3.2.1.1 Least Square Estimation

In the LS approach we attempt to minimize the squared difference between the given data $x[n]$ and the input signal or noiseless data. In the LS estimate the focus is determining a good estimator that is unbiased and has minimum variance. These are the two benefits of using the LS estimate. In choosing the variance as our measure of goodness we implicitly sought to minimize the discrepancy (on the average) between our estimate and the true

parameter value. This is illustrated in Figure (3.1). The signal is generated by some model which in turn depends upon our unknown parameter θ.



(a) Data Model



(b) Least square error

Figure 3.1: Least square approach.

The signal $s[n]$ is purely deterministic as we have discussed in the previous section. The LS estimator of θ chooses the value that makes $s[n]$ closest to the observed data $x[n]$. Closeness is measured by the LS error criterion

$$J(\theta) = \sum_{n=0}^{N-1} (x[n] - s[n])^2 \quad \text{.....3.1}$$

where the dependence of J on θ is via $s[n]$. The value of θ that minimizes $J(\theta)$ is LSE.
The linear model that we are going to deal like the channel model for the input output relationship is as follows

$$Y = H\theta + W \quad \text{.....3.2}$$

where H is a known matrix of dimension NXP and is referred to as the observation matrix, the data Y are observed after θ is operated upon by H**.** Note also that, the noise

vector has the statistical characterization $\mathbf{W} \sim N(0, \sigma^2\mathbf{I})$. This model in equation (3.2) is termed the linear model. So the above cost function (Equation 3.1) for our linear model can be rewritten as follows:

$$J(\theta) = (Y - H\theta)^H (Y - H\theta) \quad \text{.....3.3}$$

Where $(.)^H$ is a Hermetian of a given complex valued matrix. Finding the gradient of the above equation and equating it to zero results the LS estimate for the linear model in Equation (3.2) which is written in Equation (3.4). Note in the linear model Y can be assumed the received signal at the output antennas, θ the input to the transmit antennas and H is the channel through which the input signal transmits through. (See ref. [1] and [7] for detailed analysis of the LS and MMSE estimators)

$$\theta_{est} = (H^H H)^{-1} (H)^H Y \quad \text{.....3.4}$$

The LS estimation MSE bound is given by

$$MSE_{\theta_{est}} = \frac{1}{P}\{\left\|\theta_{est} - \theta\right\|^2\} \text{……..……..3.5}$$

Substitute the LS estimate form Equation (3.4) into Equation (3.5) and manipulating the result, the MSE of the LS estimator becomes

$$MSE_{LS} = \frac{\sigma_w^2}{P} trace\{(H^H H)\} \text{……..……3.6}$$

3.2.2 Bayesian Estimation

When some prior knowledge about the unknown vector of the system is available it can be incorporated into the estimator to enhance the performance. If the prior knowledge is inaccurate, a Bayesian estimator will outperform the optimal classical estimator. This comes at the price of added computational complexity and dependence on additional information of the unknown vector. Bayesian estimators include MMSE, MAP, Wiener filter, and Kalman Filter.

3.2.2.1 Minimum Mean Square Error Estimation

Again we begin our discussion by assuming a vector parameter $\boldsymbol{\theta}$ is to be estimated based on the data set $[x[0], x[1], \ldots, x[N-1]]^T$, the unknown parameter is modeled as the realization of a random variable. We do not assume any specific form for the joint PDF $p(x, \theta)$, but only a knowledge of the first two moments. The θ may be estimated from x is due to the assumed statistical dependence of θ on x as summarized by the joint PDF $p(x, \theta)$, and in particular we rely on the correlation between θ and x. For the linear model in Equation (3.2) we have the Bayesian MSE given as

$$B_{mse}(\theta_{est}) = E[(\theta - \theta_{est})^2] \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.7$$

In Bayesian estimation, θ is a random variable, therefore the expectation operation is with respect to the joint PDF $p(x, \theta)$. So the estimator can be rewritten as

$$B_{mse}(\theta_{est}) = \iint (\theta - \theta_{est})^2 P(Y, \theta) dY d\theta. \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.8$$

Using the Bayes' theorem

$$P(Y, \theta) = P(\theta / Y) P(Y) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.9$$

The estimate becomes

$$B_{mse}(\theta_{est}) = \int [\int (\theta - \theta_{est})^2 P(\theta / Y) d\theta] p(Y) dY \quad \ldots\ldots\ldots 3.10$$

Taking the gradient of the above equation and equating it to zero gives the Bayesian MMSE estimate of

$$\theta_{est} = \int \theta P(\theta / Y) d\theta \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.11$$

or

$$\theta_{est} = E(\theta / Y), \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.12$$

This is simply the mean of the posterior PDF p(x, θ). If the unknown vector and the noise vector are Gaussian with zero mean and are uncorrelated, then the posterior PDF $p(x, \theta)$ is also Gaussian with mean defined as

$$E(\theta / Y) = E(Y) + C_\theta H^H (H C_\theta H^H + C_w)^{-1} (Y - E(Y)) \quad \ldots\ldots\ldots\ldots\ldots 3.13$$

$$= C_\theta H^H (H C_\theta H^H + C_w)^{-1} Y \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.14$$

Where C_θ and C_w are covariance matrices for the parameter vector θ and the Gaussian vector noise W. Equation (3.14) is the solution of MMSE estimate of θ.

3.3 Channel Estimation

In the previous two sections we have seen LS and MMSE estimators for a general linear model. Now, we will discuss their application to channel estimation problems in MIMO-OFDM systems. A literature review will take the first part of this section. Next channel estimation structure for OFDM systems will be developed. And finally the results will be extended to MIMO-OFDM systems. But before that it is imperative to brief about training-based channel estimation and their pilot signal structure.

In training-based channel estimation the pilot symbols can be placed in block-typed structures or comb-typed structures. As shown in Figure (3.2a), for block-type arrangement the entire OFDM symbol is dedicated to carry pilot symbols on all the subcarriers. The estimate obtained with the training symbol will be used to detect the data symbols within the OFDM packet. This arrangement is most suitable for static or slow varying channels. However, in a time varying channel the comb-type structure as depicted in Figure (3.2b) is more suitable. In the comb-type arrangement, pilot symbols are sparsely spread on selected subcarriers and repeated over multiple symbols. Channel estimation is performed at each symbol and interpolation is required to infer the channel frequency values of the non-pilot subcarriers. The choice of pilot arrangement depends on the channel environment. In this thesis a Block-type pilot arrangement is used since no special arrangement for the pilot symbols is required.

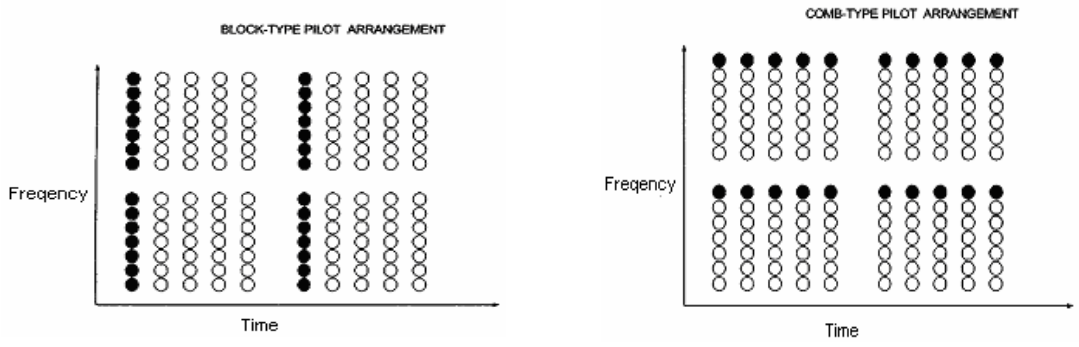


Figure 3.2: Pilot signal arrangement (the solid blacks are pilot signals)

3.3.1 Literature Review

Channel estimation for OFDM received significant amount of attention since the mid 1990's. Van de Beek et al. were one of the first to look at the basic solutions of LS and MMSE estimation for OFDM in [8]. In his work, van de Beek used the property that the channel is a finite-length impulse response to develop a system model so that the LS and MMSE estimator can be applied. He showed the CIR can be estimated from the frequency domain by structuring the CFR as a linear transformation of the CIR through the IFFT operation. This type of estimators became known as transform-domain estimators. Different researchers also try to improve the works of van de Beek by proposing different methods. SVD is one of the methods used, but these methods are of high complexity [9]. In [10] channel estimation for OFDM was dealt with some pilot arrangements. The estimation of channel at pilot frequencies is based on LS and LMS while the channel interpolation is done using linear interpolation, second order interpolation, low-pass interpolation, spline cubic interpolation, and time domain interpolation.

The works mentioned so far only take into consideration the frequency-domain correlations of the frequency response of the channel and were only for OFDM systems. These methods are all using either of the pilot arrangements shown above. It was also a great concern in the design of pilot signal arrangements in the development of efficient channel estimations for MIMO-OFDM systems. However use of well designed pilot signals makes the transmission system to be bandwidth inefficient, for the pilots needs additional bandwidth.

In summary, there has been a handful of research in the area of channel estimation for MIMO-OFDM. Earlier works were only applicable to OFDM systems, but in recent years, the works have been extended to MIMO-OFDM due to the promising potentials of combining MIMO-OFDM. In much of the works reviewed, performance of the channel estimation scheme is the primary concern. Channel estimation in MIMO-OFDM is a complicated problem because there are more unknowns to be determined. Usually the solution involves an inversion of a large matrix. The problem of complexity has to be addressed because ultimately the algorithm needs to be implemented in

hardware. In some recent works complexity problems have been addressed. There are different methods in complexity reductions of channel estimation. In [11] for example specially designed training sequences are used to avoid the large matrix inversions. However the need for specific training sequences reduces the robustness of the algorithm. The work in chapter five presents robust channel estimation with out setting limitations to design of training sequences.

3.3.2 OFDM channel estimation

One of the main attractions of OFDM for frequency selective channels is its ability to simplify the channel estimation process. It was shown in Figure 2.7(b) that in the frequency domain the transmitted signal on each subcarrier is multiplied by a small portion of the channel frequency response, where each subcarrier only experiences flat fading. Referring to the schematic diagram of an OFDM system found in Figure (2.8), the channel estimation can be performed before or after the FFT block. If the channel estimation is done before the FFT block, then the estimation is said to be undergone in the time domain. If the estimation is done after the FFT block then the estimation is said to be in the frequency domain. Each has its own merits and demerits, but tracking in time is more robust to errors because the frequency correlations of the taps can be efficiently exploited. In chapter five channel tracking is done in time domain. The received signal can be modeled with the following equation:

$$Y = XH + V \quad \text{.....3.15}$$

Where **Y** is the received signal vector, **X** is a diagonal matrix of the transmitted signal, **H** is the channel frequency response vector, and **V** is the noise vector in the frequency domain. The received signal in Equation (3.15) has the same structure as the general linear data model described by Equation (3.2). Using the results of the LS and MMSE estimator developed in Section 3.2, the LS estimator for an OFDM system is described as:

$$\hat{\mathbf{H}}_{LS} = (\mathbf{X}^H \mathbf{X})^{-1} \mathbf{X}^H \mathrm{Y} \quad \text{.....3.16}$$

Since **X** is a diagonal matrix, the estimation reduces to

$$\hat{H}_{LS} = \mathbf{X}^{-1} Y \quad \text{.....3.17}$$

This indicates that the LS estimate of the frequency response channel is simply the division of the received signal by the transmitted signal. Likewise applying Equation (3.14) for the model in Equation (3.15), the MMSE estimator for OFDM results as

$$H_{MMSE} = C_H \mathbf{X}^H (\mathbf{X} C_H \mathbf{X}^H + C_V)^{-1} Y \quad \text{.....3.18}$$

where C_H and C_V, are the covariance matrices of, the channel's frequency response and noise, respectively.

3.3.3 MIMO-OFDM channel estimation

The problem of channel estimation for OFDM has been well researched; however, the results are not directly applicable to MIMO-OFDM systems. In MIMO systems, the number of channels increases by $M_t M_r$-folds, where M_t and M_r are the number of transmit and receive antenna, respectively. This significantly increases the number of unknowns to be solved. Conventional estimation techniques for single input single output (SISO) systems have to be modified to be applicable in MIMO systems.

Using the MIMO-OFDM system model described in Chapter 2, we will develop the channel estimator for MIMO-OFDM. We will assume a *2-by-2* MIIMO system to illustrate the added complexity of MIMO-OFDM channel estimation. The received signal at the j^{th} antenna for the k^{th} subcarrier in expanded form is defined as:

$$\boldsymbol{Y}_k^{(j)}[n] = \boldsymbol{H}_k^{(j,1)} S_k^{(1)}[n] + \boldsymbol{H}_k^{(j,2)} S_k^{(2)}[n] \text{ , k= 0,1 … , } N\text{-}1 \quad \text{.....3.19}$$

The above equation is under determined. There are two unknown elements $\boldsymbol{H}_k^{(j,1)}$ and $\boldsymbol{H}_k^{(j,2)}$ from different channels, however the unknowns cannot be solved with just one equation. In a 2-by-2 system, two samples of the received signal, $\boldsymbol{Y}_k^{(j)}[n]$ and $\boldsymbol{Y}_k^{(j)}[n+1]$, are required to estimate $\boldsymbol{H}_k^{(j,1)}$ and $\boldsymbol{H}_k^{(j,2)}$. For our system of 2-by–2 MIMO system for example the received signal at the j^{th} antenna j = 1, 2 is represented by the following equation

$$\begin{bmatrix} Y_1^j(n) \\ Y_2^j(n) \\ . \\ . \\ . \\ Y_N^j(n) \end{bmatrix} = \begin{bmatrix} X_1^1(n) & 0 & . & . & . & 0 & X_1^2(n) & 0 & . & . & . & 0 \\ 0 & X_2^1(n) & & & & . & & X_2^2(n) & & & & \\ . & & & & & . & & & & & & \\ . & & & & & . & & & & & & \\ & & & & & 0 & & & & & & \\ 0 & . & . & . & 0 & X_N^1(n) & 0 & . & . & . & 0 & X_N^2(n) \end{bmatrix} \begin{bmatrix} H_1^{(1,j)} \\ H_2^{(1,j)} \\ . \\ . \\ . \\ H_N^{(1,j)} \\ H_1^{(2,j)} \\ H_2^{(2,j)} \\ . \\ . \\ . \\ H_N^{(2,j)} \end{bmatrix} + \begin{bmatrix} V_1^{(j)}(n) \\ V_2^{(j)}(n) \\ . \\ . \\ . \\ V_N^{(j)}(n) \end{bmatrix}$$

$$Y^{(j)} = \boldsymbol{X}H^{(j)} + V^{(j)} \quad \text{.....3.20}$$

Note, the $H^{(j)}$ vector contains NM_t unknown elements but there are only N equations available in a block. In general, for M_t transmit antennas we need to collect a minimum of M_t blocks to solve for the channel unknowns. The complexity of the estimation problem increases significantly since the matrix size is increased by M_t -folds. This problem can be reduced by looking at an alternate representation of the received signal, called the transform-domain estimator that was first proposed by van de Beek in [8] for OFDM systems. Basically we know that the CFR is a Fourier transform of the CIR, which is a linear transformation through IFFT operation. In other words, the CFR can be expressed in terms of the CIR through the Fourier transformation. Hence, the received signal model in Equation (3.20) can be expressed in terms of the CIR. The benefit of this representation is that usually the length of the CIR is much less then the number of subcarriers of the system. In order to model the received signal in terms of the CIR we first need to express the CFR as a function of the CIR. In Equation (2.12), the Fourier transform of a single CIR is defined. The equation can be rewritten in vector form as

$$H^{(j,i)} = \boldsymbol{F}h^{(j,i)} \quad \text{.....3.21}$$

Where F is called the Fourier transform matrix of size *(N X L)*, and *h* as given by Equation (2.10) is the *(L X 1)* channel impulse response vector. To extend the Fourier transformation to multiple channels we need to define the following

$$\Phi = \begin{bmatrix} F & 0 & . & . & . & 0 \\ 0 & F & & & & . \\ . & & . & & & . \\ . & & & . & & . \\ . & & & & . & 0 \\ 0 & . & . & . & 0 & F \end{bmatrix} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.22$$

which is a block diagonal matrix of M_t *F's*. Then the alternate received signal of the j^{th} antenna in terms of the CIR can be expressed as

$$\begin{aligned} Y^{(j)} &= \boldsymbol{X\Phi} h^{(j)} + V^{(j)} \\ &= \boldsymbol{W} h^{(j)} + V^{(j)} \end{aligned} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.23$$

In this representation, the number of elements in $h^{(j)}$ to be solved are M_tL. If we assume that the number of subcarriers, N, in an OFDM block is greater than M_tL, then only one OFDM is required to solve for $h^{(j)}$.

Using the received signal model in Equation (3.23) the LS estimate is expressed as:

$$\hat{\boldsymbol{h}}_{\boldsymbol{LS}}^{(j)} = (\boldsymbol{W}^H\boldsymbol{W})^{-1}\boldsymbol{W}^H(Y)^{(j)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.24$$

Using the same model the MMSE estimate becomes, with the assumption that the noise is AWGN ($C_V = \sigma^2_V\mathbf{I}$) and using the matrix inversion theorem,

$$\hat{\boldsymbol{h}}_{\boldsymbol{MMSE}}^{(j)} = (\sigma_V^2\, C_h^{-1} + \boldsymbol{W}^H\boldsymbol{W})^{-1}\boldsymbol{W}^H(Y)^{(j)} \quad \ldots\ldots\ldots\ldots\ldots\ldots 3.25$$

3.4 Simulation Results

The objective of this section is to show how traditionally channel estimation is performed and to lay a ground work for our work in chapter five. A simulation of an MIMO-OFDM system was developed to compare the performance of the LS and MMSE estimator under different circumstances. The following parameters are for typical urban type of scenario which we have attempted to use a realistic environment for the transmission and receiving process. These results are to show how channel

estimation performances are evaluated and in addition to compare the performances of LS and MMSE estimators. These two will in turn be compared with our system in chapter five in the sense of MSE errors.

3.4.1 System Parameters and Assumptions

I. Number of transmitter antennas, $M_t = 2$
II. Number of receiver antennas, $M_r = 2$
III. BPSK symbol modulation
IV. Number of subcarriers, $N = 256$
V. Cyclic Prefix, $N_{CP} = N/4$
VI. OFDM block size, $N + N_{CP}$
VII. Bandwidth = 1.2 MHz.
VIII. Carrie frequency is $f_c = 2.4$ GHz.
IX. Wireless channels from each transmit to each receive antenna experience Rayleigh fading with independent propagation paths.
X. The receiver speed, the speed for which the mobile user travels, is of 36 Km/h.
XI. The Doppler spectrum is Jake's and the channel length is 16.
XII. No correlations is considered neither at the transmitter and nor at the receiver side, which means the four channel impulse responses, h_{11},h_{12},h_{21},and h_{22} fade independently.
XIII. Channels are considered to remain stationary during the OFDM block time.

Some of the parameters are taken from [2] where the Author used some realistic data. These parameters are to be used for the work in chapter five.

3.4.2 Performance Analysis

A training symbol is placed at the beginning of the OFDM block. The channel estimate obtained from the training symbol will be used to estimate the channel for the remaining data symbols within the block. The performance of the two estimation techniques is measured in terms of the MSE of the channel estimates.

Figure (3.3) shows the normalized MSE curve of the simulated results of the LS and MMSE estimators as derived in Section 3.3.3. From the figure we can point out that at lower SNR values the MMSE outperforms than the LS estimator. But at higher values of SNR the two performs almost the same. This can be also explained from their equations derived in section 3.2. When the SNR is high, the noise power (σ_V^2) is low; and it will reduce the effect of C_h^{-1}. Therefore, when the noise power approaches zero, the MMSE estimator will approach the LS estimator. If we talk of complexity of the estimators, we can see from Equations (3.24) and Equation (3.25) that the MMSE estimator is more complex than the LS estimators. But in applying the two estimators there should always be a compromise with respect to signal power issues. The added complexity in MMSE is acceptable only at low SNR values (0-15dB). As we can see from Figure (3.3) at higher SNR values (15-50dB) both estimators perform equivalently.

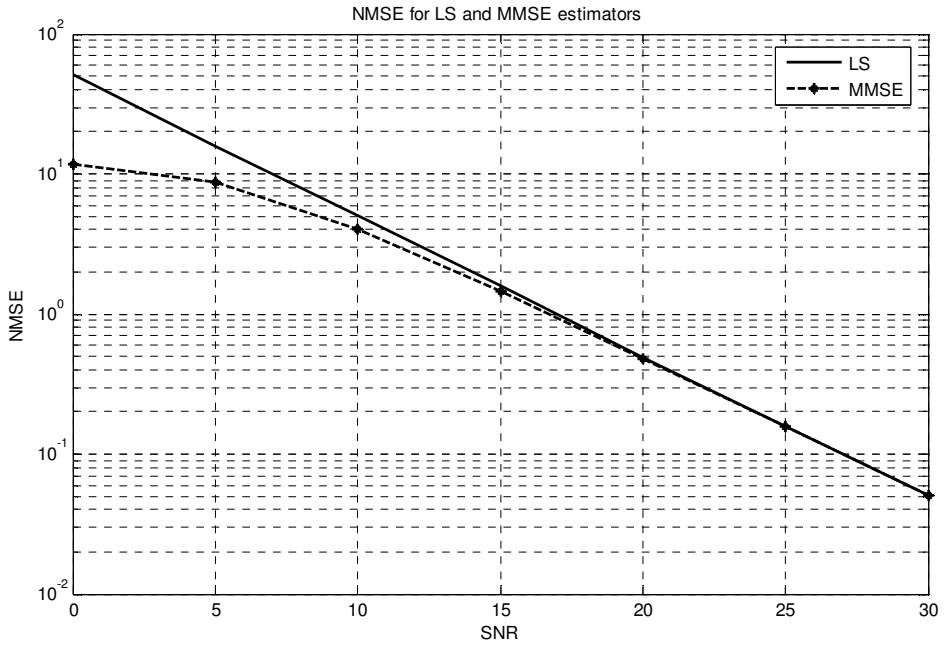


Figure 3.3: comparison of MMSE and LS estimators for a MIMO-OFDM system.

3.4.3 Computational Complexity

The main advantage of analyzing computational complexity of the channel estimation helps to visualize how much it is complex to apply it in real systems. In this section, we measure the computational complexity in terms of number of operations for each method. Our derivations are based on an M_t-by-M_r MIMO-OFDM system with N subcarriers and a channel length of L. The known matrix $\mathbf{W=X\Phi}$ has dimensions (N x M_t). For simplicity in notation we denote NM_t with just M. For a consistent comparison, the complex operations are converted to real operation equivalents. In addition, each type of real operations has different levels of complexity when implemented in hardware. It should be emphasized that counting of the number of operations is only an estimate of the computational complexity of the algorithms. A more exact measure would be to implement the algorithm in hardware and count the number of instructions and processing time required. However, in computer simulations, operation counts can give a good indication of the relative complexity of different algorithms.

Operation	No. of Complex multiplications	No. of Complex Add/Sub	No. of Complex Division	No. of square root
$A=(W^HW)$	NM^2	NM^2-M^2	*0*	*0*
A^{-1}	M^3	M^3	$3M^2/2+M/2$	*0*
$B=A^{-1}W^H$	NM^2	NM^2-NM	*0*	*0*
$\hat{h} = BY$	NM	$NM-M$	*0*	*0*

Table 3.1 Channel Estimation Operation Count for LS

Table 3.1 summarizes operation counts for LS channel estimator which is represented in Equation (3.24). Table 3.2 summarizes the same analysis for MMSE estimator represented in Equation (3.25). Let us represent $B = C_h$ and the others remain as it was for the LS estimator, then we have the counts summarized as follows.

Operation	No. of Complex multiplications	No. of Complex Add/Sub	No. of Complex Division	No. of square root
$B = C_h$	*0*	*0*	*0*	*0*
B^{-1}	M^3	M^3	$3M^2/2 + M/2$	*0*
$\sigma_v^2 B^{-1}$	M^2	*0*	*0*	*0*
$C = \sigma_v^2 B^{-1} + A$	NM^2	M^2	*0*	*0*
C^{-1}	M^3	M^3	$3M^2/2 + M/2$	*0*
$D = C^{-1} W^H$	NM^2	$NM^2 - NM$	*0*	*0*
$\hat{h} = DY$	*NM*	*NM-M*	*0*	*0*

Table 3.2 Channel Estimation Operation Count for MMSE (Equation 3.25)

To have a more visualization on the computational complexity for the two estimators let us compare them for a channel length L=5 and 2 transmit antennas with 256 subcarriers.

Method	No. of Complex multiplications	No. of Complex Add/Sub	No. of Complex Division	No. of square root
LS	54,760	52100	155	0
MMSE	55,860	53,190	310	0

Table 3.3 Comparison of LS and MMSE for a two transmit antenna and for L=5

CHAPTER FOUR
KALMAN FILTERS

In this chapter we discuss the significance of using Kalman filters which accommodate vector signals and noises which additionally may be nonstationary. A thorough discussion of sequential MMSE, dynamical signal models, scalar and vector Kalman filters and their properties is presented in the subsequent sections. The applications of the Kalman filters to our work are also addressed. In section 4.1dynamical signal model is reviewed. Scalar Kalman filter is discussed to elaborate the principles of the filter with its equations and its block diagrams in section 4.2. The properties of Kalman filters which make them essential in this work is also briefed. Section 4.3 discusses the vector Kalman filter and some properties that make it different from the scalar Kalman filter.

4.1 Dynamical Signal Model

Suppose that we have a Bayesian linear model as in Equation (3.2) which we repeat here with some convenient changes in variable as follows:

$$x = H\theta + w \quad \text{.....4.1}$$

Where each vector observation form has this form. Indexing the quantities by n and replacing θ by s, the observation model will be

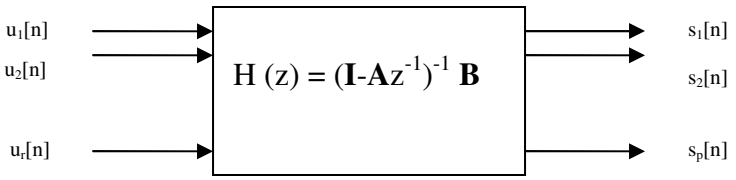
$$x[n] = H[n]s[n] + w[n] \quad \text{.....4.2}$$

Where $H[n]$ is a known $M\ X\ P$ matrix, $x[n]$ is an $M\ X\ 1$ observation vector, and $w[n]$ is a $MX\ 1$ observation noise sequence. The $w[n]$'s are independent of each other and $w[n] \sim \mathcal{N}(\mathbf{0}, \mathbf{C}[n])$. Except for the dependence of the covariance matrix on n, $w[n]$ can be thought of as vector AWGN. For an array processing problems for example $s[n]$ represents a vector of p transmitted signals, which are modeled as random, and $H[n]$ models the linear transformation due to the medium.

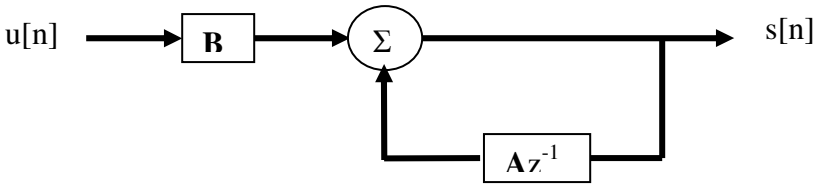
For such a process a simple model for s[n] which allows us to specify the correlation between samples is the *first-order Gauss-Markov* process (see Figure 4.1).

$$s[n] = As[n-1] + Bu[n] \quad \text{.....4.3}$$

Where *A* is a *p x p* nonsingular matrix (termed the *state transition matrix*) and *B* is *p x 1* vector, and s[n] is a *p x 1* signal vector. Equation (4.3) is referred to as the *state model.* In the vector Gauss-Markov model of Equation (4.3) the matrices A and B are assumed to be known and the eigenvalues of A are less than 1 in magnitude for the reason of stability of the system. The deriving noise vector u[n] has dimension *r x 1* and is vector AWGN or $u[n] \sim \mathcal{N}(0, Q)$ with $u[n]$'s independent.



(a) Multi-input---multi-output model



(b) Equivalent vector model

Figure 4.1 Vector Gauss-Markov signal system model

4.2 Scalar Kalman filters

The scalar Gauss-Markov signal model form for Equation (4.3) discussed in the previous section is

$$s[n] = as[n-1] + u[n], \qquad n \geq 0 \quad \text{.....4.4}$$

Suppose that we have a sequential MMSE estimator which will allow us to estimate $s[n]$ based on the data $\{x[0], x[1], \ldots, x[n]\}$ as n increases. Such an operation is referred to as filtering. The approach computes the estimator $\hat{s}[n]$ based on the estimator for the previous time sample $\hat{s}[n-1]$ and so is recursive in nature. This is the so called *Kalman filter*. The versatility of the Kalman filter accounts for its widespread use. It can be applied to estimation of the scalar Gauss-Markov signal as well as to its vector extension. Furthermore, the data, which previously in our discussions consists of a scalar sequence such as $\{x[0], x[1], \ldots, x[n]\}$, can be extended to vector observations or $\{\boldsymbol{x}[0], \boldsymbol{x}[1], \ldots, \boldsymbol{x}[n]\}$. There different levels of generality are there for Kalman filter based on the state, s[n], and the data observations ,$\{x[0], x[1], \ldots, x[n]\}$, are either scalar or vector

1. Scalar state – scalar observation (s[n-1] ,x[n])
2. Vector state – scalar observation (**s**[n-1], x[n])
3. Vector state – vector observation (**s**[n-1], **x**[n])

The first one is treated in this section, and the third one in the next section. Refer [7] for the complete discussion on the three.

Consider the scalar state equation and scalar observation equation

$$\begin{aligned} s[n] &= as[n-1] + u[n] \\ x[n] &= s[n] + w[n] \end{aligned} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 4.5$$

Where $u[n]$ is zero mean Gaussian noise with independent samples and $E(u^2[n]) = \sigma_u^2$, $w[n]$ is zero mean Gaussian noise with independent samples and $E(w^2[n]) = \sigma^2$. Finally it is assumed that $s[-1]$, $u[n]$, and $w[n]$ are all independent with $s[-1] \sim \mathcal{N}(\mu_s, \sigma_s^2)$. For the purpose of derivation we can assume that $\mu_s = 0$ so that $E(s[n]) = 0$ for $n<0$. The purpose here is to estimate s[n] based on the observations {x[0], x[1], …, x[n]} or to filter x[n] to produce ŝ[n]. More generally, the estimator of s[n] based on the observations {x[0], x[1], …, x[n]} will be denoted by $\hat{s}[n/m]$. The criterion for optimality will be the minimum Bayesian MSE or

$$E[(s[n] - \hat{s}[n/n])^2]$$

Where the expectation is with respect to $p(x[0], x[1], \ldots, x[n], s[n])$. But the MMSE estimator is just the mean of the posterior PDF or

$$\hat{s}[n/n] = E(\text{s[n] / x[0], x[1], …, x[n]}) \quad \text{……..4.6}$$

Using zero means this becomes

$$\hat{s}[n/n] = C_{\theta x} C_{xx}^{-1} x \quad \text{…….4.7}$$

Where C_{xx} is the covariance matrix between the observation vector and $C_{\theta x}$ is the covariance matrix between the observation {x[0], x[1], …, x[n]} and the data s[n]. Equation (4.7) is true since $\theta = s[n]$ and $x[n] = [x[0], x[1], \ldots, x[n]]^T$ are jointly Gaussian. Because we are assuming Gaussian statistics for the signal and noise, the MMSE estimator is *linear* and is *identical* in algebraic form to the LMMSE estimator. The algebraic properties allow us to utilize the vector space approach to find the estimator. The implicit linear constraint does not detract from the generality since it is already known that the optimal estimator is linear. Furthermore, if the Gaussian assumption is not valid, then the resulting estimator is still valid but can only said to be the optimal LMMSE estimator. Returning to the sequential computation of (4.7) we note that if x[n] is uncorrelated with $\{x[0], x[1], \ldots, x[n]\}$, then form Equation (4.6) and the orthogonality principle we will have

$$\begin{aligned}\hat{s}[n/n] &= E(\text{s[n] / x[0], x[1], …, x[n-1]}) + E(\text{s[n] / x[n]}) \\ &= \hat{s}[n/n-1] + E(\text{s[n] / x[n]})\end{aligned}$$

which has the desired form. Unfortunately, the x[n]'s are correlated due to their dependence on s[n], which is correlated from sample to sample. From the LMMSE estimator, the vector space interpretation can be used to determine the correlation of the old estimator $\hat{s}[n/n-1]$ due to the observation of x[n] [7]. Summarizing some of the properties of the MMSE estimator that will be used in the derivation of Kalman filter

1. The MMSE estimator of θ based on two uncorrelated data vectors, assuming jointly Gaussian statistics, is

$$\hat{\theta} = E(\theta / x_1, x_2)$$
$$= E(\theta / x_1) + E(\theta / x_2)$$

If θ is zero mean.

2. The MMSE estimator is additive in that if $\theta = \theta_1 + \theta_2$, then

$$\hat{\theta} = E(\theta / x)$$
$$= E(\theta_1 + \theta_2 / x)$$
$$= E(\theta_1 / x) + E(\theta_2 / x)$$

With these properties the scalar state – scalar observation Kalman filter equations can be summarized as follows (detail derivation is shown in [7]). Although the derivation is tedious, the final equations are actually quite simple and intuitive. For n≥0

Prediction:

$$\hat{s}[n/n] = a\,\hat{s}[n\text{-}1/n\text{-}1] \quad \text{.....4.8}$$

Minimum Prediction MSE:

$$M[n/n\text{-}1] = a^2\, M[n\text{-}1/n\text{-}1] + \sigma^2_u \quad \text{.....4.9}$$

Kalman Gain:

$$K[n] = \frac{M[n/n-1]}{\sigma_u^2 + M[n/n-1]} \quad \text{.....4.10}$$

Correction:

$$\hat{s}[n/n] = \hat{s}[n/n\text{-}1] + K[n](x[n] - \hat{s}[n/n\text{-}1] \quad \text{.....4.11}$$

Minimum MSE:

$$M[n/n] = (1 - K[n])\, M[n/n\text{-}1] \quad \text{.....4.12}$$

Although derived for $\mu_s = 0$ so that E(s[n]) = 0, the same equation apply for $\mu_s \neq 0$. Hence, to initialize the equations we use $\hat{s}[\text{-}1/\text{-}1] = E$(s[-1]) = μ_s and M[-1/-1] = σ^2_s since this amounts to the estimation of s[-1] without any data. A block diagram of the Kalman filter is shown in Figure 4.2. It is interesting to note that the dynamical model for the signal is an integral part of the estimator. Furthermore, we may view the output of the gain block as an estimator ŭ[n] of u[n]. From (4.8) and (4.11) the signal estimate is

$$\hat{s}[n/n] = a\,\hat{s}[n\text{-}1/n\text{-}1] + \check{u}[n]$$

Where $\check{u}[n] = K[n](x[n] - \hat{s}[n/n\text{-}1]$. To the extent that this estimate is approximately u[n], we will have $\hat{s}[n/n] \approx s[n]$, as desired.

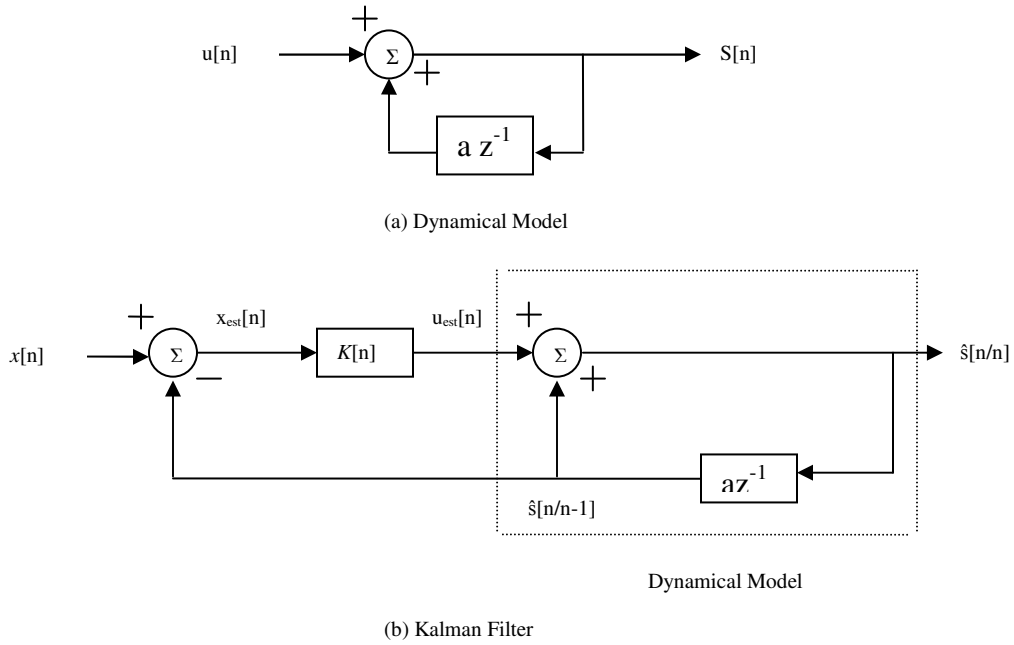


Figure 4.2 Scalar state – scalar observation Kalman filter and relationship to dynamic model

4.3 Important properties of Kalman Filters

In this section some of the important properties of Kalman filters that make them a candidate for different applications are discussed. Because of these properties Kalman filters find their applications in navigation systems, rocket tracking and different other areas of estimations. The following are some of the important properties of the Kalman filters

1. The Kalman filter extends the sequential MMSE estimator to the case where the unknown parameter evolves in time according to the dynamical model. Note that the Kalman filter will reduce to sequential MMSE estimator equations when the parameter to be estimated does not evolve in time.
2. No matrix inversions are required. This should be compared to Equation (3.25).

3. The Kalman filter is a time varying linear filter. Note from Equation (4.8) and (4.11) that it is a first-order recursive filter with time varying coefficients as shown in Figure (4.2).
4. The Kalman filter provides its own performance measure. From Equation (4.12) the minimum Bayesian MSE is computed as an integral part of the estimator. Also, the error measure may be computed off-line, i.e., before any data are collected. This is because $M[n/n]$ depends only on Equation (4.9) and (4.10), which are independent of the data.
5. The prediction stage increases the error, while the correction stage decreases it. There is an interesting interplay between the minimum prediction MSE and the minimum MSE. If $n \rightarrow \infty$ a steady state condition is achieved, then $M[n/n]$ and $M[n/n-1]$ each become constant with $M[n/n-1] > M[n-1/n-1]$. Hence, the error will increase after the prediction stage. When the new data sample is obtained, the estimate is corrected, which decreases the MSE according to Equation (4.12) (since $K[n] < 1$).
6. Prediction is the integral part of the Kalman filter. It is seen from (4.8) that to determine the best filtered estimate of s[n] prediction is employed. So it is possible to find the best one-step prediction of s[n] based on {x[0], x[1], … , x[n-1] }from (4.8).
7. The Kalman filter is optimal in that it minimizes the Bayesian MSE for each estimator $\hat{s}[n]$. If the Gaussian assumption is not valid, then it is still the optimal linear MMSE estimator.

4.5 Vector Kalman filter

The scalar state – scalar observation Kalman filter is easily generalized. All the properties of the Kalman filter carry over to the vector state case except for property (2) , if the observations are vector, i.e., there is matrix inversion in vector Kalman filter. The two generalizations are to replace s[n] by **s**[n], where **s**[n] obeys the Gauss-Markov model, and to replace the scalar observation x[n] by the vector observation **x**[n]. The first generalization will produce the vector state – scalar observation Kalman filter,

while the second leads to the most general form, the vector state – vector observation Kalman filter. In either case the state model is

$$\mathbf{s}[n] = \mathbf{A}\mathbf{s}[n-1] + \mathbf{B}\mathbf{u}[n] \quad n \geq 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 4.13$$

Where **A**, **B** are known $p \times p$ and $p \times r$ matrices, **u**[n] is vector AWGN with $\mathbf{u}[n] \sim \mathcal{N}(\mathbf{0},\mathbf{Q})$, $\mathbf{s}[-1] \sim \mathcal{N}(\boldsymbol{\mu}_s, \boldsymbol{C}_s)$, and **s**[-1] is independent of the **u**[n]'s.

Let us generalize the Kalman filter for the case of *vector observation*, which is quite common in practice and the work in chapter five, is also based. Then the observations are modeled using the Bayesian linear model for vector observation.

$$\mathbf{x} = \mathbf{H}\boldsymbol{\theta} + \mathbf{w} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 4.14$$

where each vector observation has the form in equation (4.14). Indexing the quantities by n and replacing $\boldsymbol{\theta}$ by **s**, the observation model becomes

$$\mathbf{x}[n] = \mathbf{H}[n]\mathbf{s}[n] + \mathbf{w}[n] \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 4.15$$

where **H**[n] is a known $M \times P$ matrix, **x**[n] is $M \times 1$ observation vector, and **w**[n] is a $M \times 1$ observation noise sequence. The **w**[n]'s are independent of each other and of **u**[n] and **s**[-1], and $\mathbf{w}[n] \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{C}[n])$. Except the dependence of the covariance matrix on n, **w**[n] can be thought of as vector AWGN. For the problem in chapter five **s**[n] represents a vector of p transmitted signals, which are modeled as random, and **H**[n] models the linear transformation due to the medium. The medium may be modeled as time varying since **H**[n] depends on n. Hence, **H**[n] **s**[n] is the signal output at the M_t antennas. Also the antennas output are corrupted by noise **w**[n]. The statistical assumptions on **w**[n] indicate the noise is corrupted from antenna to antenna at the same time instant with covariance **C**[n], and this correlation varies with time. From time to time instant, however, the noise samples are independent since $E(\mathbf{w}[i]\ \mathbf{w}[j]) = 0$ for $i \neq j$. With this data model the vector state – vector observation Kalman filter can be summarized with the above assumptions for the data and the noise, together with Equations (4.13) and (4.15). This is the MMSE estimator of **s**[n] based on {**x**[0], **x**[1], … , **x**[n]} or

$$\hat{\mathbf{s}}[n/n] = E(\mathbf{s}[n]/\ \mathbf{x}[0], \mathbf{x}[1], \ldots, \mathbf{x}[n]\}$$

This can be computed sequentially in time using the following recursion.

Prediction:

$$\hat{\mathbf{s}}[n/n-1] = \mathbf{A}\,\hat{\mathbf{s}}[n-1/n-1] \quad \text{.....4.16}$$

Minimum Prediction MSE Matrix (p x p):

$$\mathbf{M}[n/n-1] = \mathbf{A}\mathbf{M}[n-1/n-1]\mathbf{A}^T + \mathbf{B}\mathbf{Q}\mathbf{B}^T \quad \text{.....4.17}$$

Kalman Gain Matrix (p x p):

$$\mathbf{K}[n] = \mathbf{M}[n/n-1]\mathbf{H}^T[n](\mathbf{C}[n] + \mathbf{H}[n]\mathbf{M}[n/n-1]\mathbf{H}^T[n])^{-1} \quad \text{.....4.18}$$

Correction:

$$\hat{\mathbf{s}}[n/n] = \hat{\mathbf{s}}[n/n-1] + \mathbf{K}[n](\mathbf{x}[n] - \mathbf{H}[\mathrm{n}]\,\hat{\mathbf{s}}[n/n-1]) \quad \text{.....4.19}$$

Minimum MSE Matrix (p x p):

$$\mathbf{M}[n/n] = (\mathbf{I} - \mathbf{K}[n]\mathbf{H}[n])\mathbf{M}[n/n-1] \quad \text{.....4.20}$$

The recursion is initialized by $\hat{\mathbf{s}}[-1/-1] = \mu_s$, and $\mathbf{M}[-1/-1] = \mathbf{C}_s$. All the comments of the scalar state – scalar observation Kalman filter apply here as well, with the exception of the need for matrix inversions. In here inversion of the *M x M* matrix is needed to find the Kalman gain.

CHAPTER FIVE
KALMAN FILTERING CHANNEL ESTIMATION

So far we have been laying the ground, for our final objective; familiarizing ourselves with channel estimation, MIMO-OFDM systems and its model, and Kalman filtering. The scalar and vector Kalman filtering and the Gauss-Markov model (dynamical model) of the system was also discussed in detail. It is time to combine all this materials to achieve our final goal, which is Kalman filtering channel estimation for MIMO-OFDM systems.

Section 5.1 discusses the alternative mathematical representation for our system model in chapter 2. In section 5.2 MMSE equalizer is introduced for this model. In 5.3 the working principle of our channel estimation algorithm is discussed. Section 5.4 displays the simulation result for channel tracking. The MSE simulation result for Kalman filtering is also displayed with comparison to the LS and MMSE estimators together with their computational analysis.

5.1 Mathematical Model of the System

In chapter two we have seen the system model for MIMO-OFDM systems and their mathematical representations. The vector received signal for all antennas is expressed in Equation (2.20). It is also discussed that the cyclic prefix (CP) is also inserted after the IFFT operation for mitigating ICI. In this section an alternative representation for these equations is discussed. In Equation (3.21) we have the Fourier transform matrix used for CIR. In this section the input signal is multiplied by this matrix and the cyclic prefix addition is done in a different way, which is revealed in the channel matrix H.

For the simulations and analysis in the following sections we now put the model equations considering 2-transmit and 2- receive antenna. This helps us to visualize the channels to be estimated, but the generalization to M_t transmit and M_r receive antenna is straight forward as discussed in chapter two.

Let the k^{th} modulated OFDM block at transmit antenna i is written as $x_i = \boldsymbol{F}\ a_i\ (k)$, where $\boldsymbol{F}$ is the $N\ X\ N$ Fourier transform matrix, as described in section 3.3.3, N being

the number of subcarriers, and $a_i(k)$ is the *N x 1* complex symbol vector sent from antenna *i*. The received *2N x 1* signal block after cyclic prefix removal is expressed as:

$$\begin{bmatrix} r_1(k) \\ r_2(k) \end{bmatrix} = \begin{bmatrix} H_{11} & H_{21} \\ H_{12} & H_{22} \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + \begin{bmatrix} w_1(k) \\ w_2(k) \end{bmatrix} \quad \text{.....5.1}$$

or in compact form :

$$r(k) = H(k)\, x(k) + w(k) \quad \text{.....5.2}$$

where $r_j(k)$ is the received block of size *N x 1* at antenna *j* and H_{ij} is the *N x N* channel convolution matrix, modeling the wireless environment between the i^{th} transmit and the j^{th} receive antenna. Due to the cyclic prefix insertion and removal operations, H_{ij} matrices are circulant. This is the alternate representation rather than insertion of the CP at the transmitter and removal of it at the receiver mentioned above. Equation 5.2 can also be rewritten as:

$$r(k) = \begin{bmatrix} X_1(k) & X_2(k) & 0_{NxL} & 0_{NxL} \\ 0_{NxL} & 0_{NxL} & X_1(k) & X_2(k) \end{bmatrix} \begin{bmatrix} h_{11}(k) \\ h_{21}(k) \\ h_{12}(k) \\ h_{22}(k) \end{bmatrix} + w(k)$$

$$= X(k)h(k) + w(k) \quad \text{.....5.3}$$

Where $X_i(k)$ is a *N x L* circulant matrix formed with the modulated block $x_i(k)$ and $h = [h_{11}, h_{21}, h_{12}, h_{22}]^T$ is the *4L x 1* state vector obtained by stacking the individual channel vectors $h_{ij}(k) = [h_{i,j,0}(k), h_{i,j,0}(k), \ldots, h_{i,j,L-1}]^T$, $i,j = 1,2$ and *L* is the channel length. The maximum channel length *L* is assumed to be the length of the cyclic prefix, in order to avoid inter-block interference. Equation (5.2) is suitable for the MMSE equalizer to be discussed in the next section, whereas Equation (5.3) is formed by making X_i's circulant rather than H_{ij}, which is again the effect of insertion and removal of cyclic prefix.

Equation (5.3) is one of our model equations, which is still in the Bayesian linear model. The other one is the time evolution of the channel taps which is the Gauss-Markov model stated in Equation (4.13).

$$\boldsymbol{h}(k) = \boldsymbol{A}h(k\text{-}1) + v(k) \quad \text{.....5.4}$$

where $\boldsymbol{A}$ is the state transition matrix and $\boldsymbol{v}$ is the state noise and they behave as described in chapter four.

5.2 MMSE Equalizer

In this section MMSE equalization is done for our system in section 5.1. In Equation (5.2) the H_{ij}'s are circulant matrices, hence they implement circular convolutions. Because of this they are diaganolized by DFT (FTM) and IDFT (IFTM) operations which are multiplication operations with F^H and F [1,11]. After having performed the discrete Fourier transform on $r_j(k)$ (after the FFT block on the OFDM block diagram) we obtain:

$$\begin{aligned} r_j(k) &= F^H H_{1j}(k) F a_1(k) + F^H H_{2j}(k) F a_2 + F^H w_j(k) \\ &= D_{1j}(k) a_1(k) + D_{2j}(k) a_2(k) + n_j(k) \quad \text{.....5.5} \end{aligned}$$

For $i, j = 1, 2$ and where $n_j(k) = F^H w_j(k)$. and the diagonal matrices

$$\begin{aligned} D_{ij}(k) &= F^H H_{ij}(k) F \\ &= diag\left\{ \sum_{l=0}^{N-1} h_{i,j,l}(k) \exp\left(-j \frac{2\pi nl}{N} \right) \right\}_{n=0,\ldots,N-1} \quad \text{.....5.6} \end{aligned}$$

Contain the frequency response of the channel $h_{ij}(k)$, evaluated at the subcarrier frequencies. Equation (5.5) for $j = 1, 2$ can be merged into a single one:

$$\begin{bmatrix} r_1(k) \\ r_2(k) \end{bmatrix} = \begin{bmatrix} D_{11}(k) & D_{21}(k) \\ D_{12}(k) & D_{22}(k) \end{bmatrix} \begin{bmatrix} a_1(k) \\ a_2(k) \end{bmatrix} + \begin{bmatrix} n_1(k) \\ n_2(k) \end{bmatrix} \quad \text{.....5.7}$$

This can be written in more compact form as:

$$r(k) = D(k)a(k) + n(k) \quad \text{.....5.8}$$

The equalization stage, when operating in frequency plane is

$$\begin{aligned} s(k) &= G(k) r(k) \\ &= G(k)(D(k)a(k) + n(k)) \quad \text{.....5.9} \end{aligned}$$

$G(k)$ is the *2N x 2 N* equalizer matrix, at OFDM block time *k*. with the linear model given in Equation (5.8), the Zero Forcing equalizer is easily derived as:

$$G_{ZF}(k) = D^{-1}(k) \quad \text{.....5.10}$$

whereas MMSE equalization is performed using:

$$G_{MMSE} = R_{a,r} R_{r,r}^{-1} \quad \text{.....5.11}$$

Where $R_{r,r} = E(rr^H) = \sigma_a^2 D(k) D^H(k) + \sigma^2 I$ and $R_{a,r} = E(ar^H) = \sigma_a^2 D^H$, σ_a^2 being the average symbol energy. A closer look at the G_{ZF} and G_{MMSE} matrices shows that a structure similar to

$$D = \begin{bmatrix} D_{11} & D_{21} \\ D_{12} & D_{22} \end{bmatrix},$$

defined in Equation (5.8), each of the blocks D_{11}, D_{21}, D_{12} and D_{22} being diagonal. Exploiting this special matrix structure, significantly lower complexity can be achieved in the equalization stage since direct matrix inversion in Equation (5.11) is avoided. Furthermore, this highlights an important property of OFDM types of transmission: the initial frequency selective channel has been turned into a set of N frequency flat channels. The equalizer is needed here to compensate the flat fading experienced on each subcarrier and also to de-multiplex the two transmitted streams a_1 and a_2.

5.3 Time Domain Channel Tracking

In mobile MIMO wireless communications, the channels are time-varying, and the channels needs to be tracked and equalizer coefficients updated periodically. With the two Equations (5.3) and (5.4), which form the state-space model for our transmission system, we now discuss the algorithm for the channel tracking. In Equation (5.4) there are $M_r M_t L$ channel taps in the state vector of dimension $M_r M_t L$ matrix. The state transition matrix is of size $M_r M_t L$ *x* $M_r M_t L$. In this thesis A is considered to have a value close to one ($A = 0.999 I_{MtMrL}$) for the reason discussed in chapter four. The state transition matrix describes the dynamics of the state vector, i.e., its time auto-

correlation properties. Low order auto-regressive (AR) are widely and practically used in literature [11, 13]. The covariance matrix for the channel taps is $Q_h = \sigma_h^2 I_{M_r M_t L}$ with σ_h^2 is the variance of the state noise associated with channel coefficients. Since now we have the linear model in Equations (5.3) and (5.4) and the noise is Gaussian, Kalman filtering discussed in the previous chapter can be applied to estimate the state vector *h*. We start with the covariance matrix of the measurement noise which has a value equal to $\boldsymbol{R} = \sigma^2 \boldsymbol{I}$, which is assumed to be uncorrelated. Let *M*(*k/k-1*) is the covariance matrix of the prediction error, then it is updated as

$$M(k/k-1) = AM(k-1/k-1)A^H + Q_h \quad \text{.....5.12}$$

Q_h being the covariance matrix of the state noise. *M (-1/-1)* is originally obtained from the covariance matrix of the channel vectors and is a matrix of size $M_t M_r L \; X \; M_t M_r L$ (see also section 4.5). The updated equation for *M* (*k/k*), the covariance matrix of the estimated error, is

$$M(k/k) = M(k/k-1) - K(k)X(k)M(k/k-1) \quad \text{.....5.13}$$

Where *K* (*k*) is known us the Kalman gain with matrix size of $M_r \; X \; M_r M$ where $M=M_t L$ and is given as

$$K(k) = M(k/k-1)X^H(k)\left[X(k)M(k/k-1)X^H(k) + R\right]^{-1} \quad \text{.....5.14}$$

The estimated channel at time *k* with information and predictions at time *k-1* and *k* respectively is given as

$$\hat{h}(k/k) = \hat{h}(k/k\text{-}1) + K(k)[\mathrm{r}(k) - X(k)\,\hat{h}(k/k\text{-}1)] \quad \text{.....5.15}$$

Where $\hat{h}(k/k\text{-}1) = A\,\hat{h}(k\text{-}1/k\text{-}1)$,and with r(k) being $M_r N \; X \; 1$ and $\hat{h}(k/k\text{-}1)$ being $M_r M \; X \; 1$ vectors. Equations (5.12) to (5.15) are the vector Kalman equations discussed in the previous chapter except for some changes in the terms used. The channel tracking algorithm is illustrated is Figure (5.1) and it works as follows:

1. The received vector at time *k*, *r*(*k*), is used to obtain the symbol estimate $\hat{a}(k)$. This is done using the MMSE equalizer and the channel estimate at time *k-1*; $\hat{h}(k\text{-}1/k\text{-}1)$.

Note that to start with the channel estimate at time k-1, $\hat{h}(k\text{-}1/k\text{-}1)$ is initialized with the facts discussed in chapter four, i.e., $\hat{h}(-1/-1) = \mu_h$, or even with reasonable guess.

2. The estimate at time k, $\hat{a}(k)$ is re-modulated with IFTM so that the estimate is in time domain for use in the Kalman filter, $\hat{x} = F\hat{a}(k)$.
3. Now $\hat{x}(k)$ is used to build the estimate $\hat{X}(k)$ as in Equation (5.3).
4. And finally the Kalman filter is run to obtain $\hat{h}(k/k)$ using Equation (5.3) and (5.4) as our transmission model.
5. Step 1 to 4 is repeated for the next OFDM block.

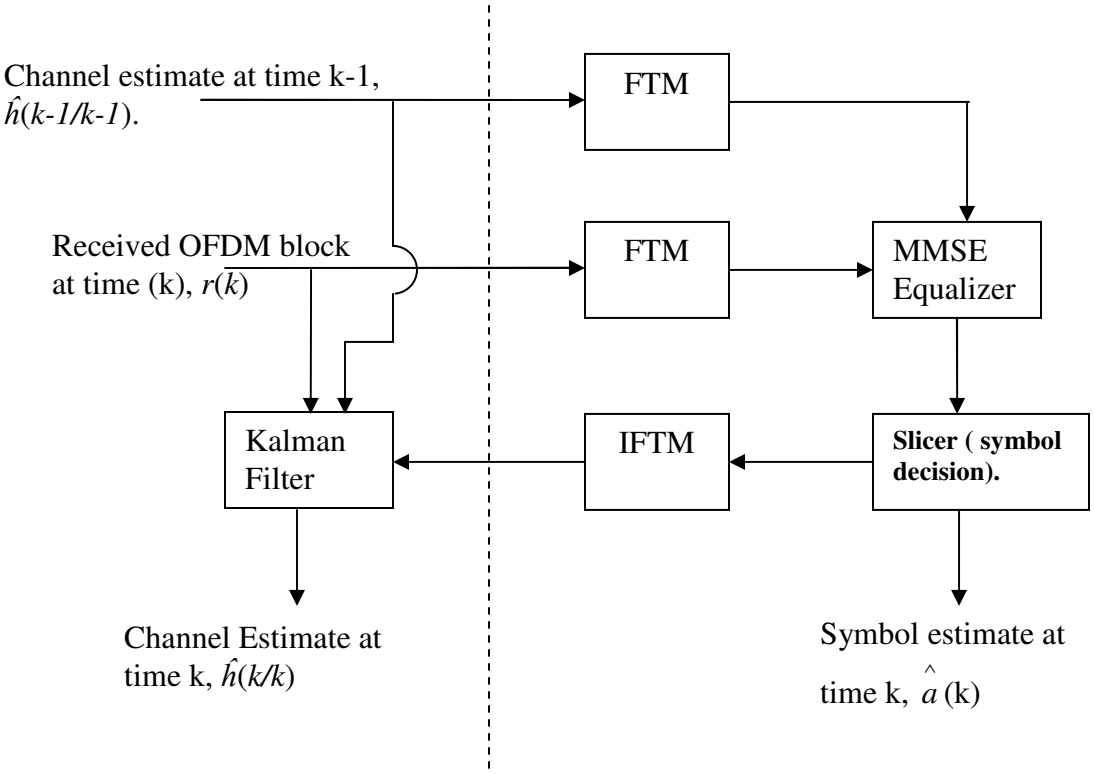


Figure 5.1 Channel estimation followed by frequency domain equalization.

At this point the estimates can be refined by re-decoding the symbol $\hat{a}(k)$ using the filtered estimate $\hat{h}(k/k)$. Substantial performance is obtained as is shown in the next section. Except for the first OFDM block, which will be used as a training signal, the

algorithm works in a decision directed mode. In the blocks in Figure (5.1) the Kalman filter equations and channel estimation procedures are integrated.

5.4 Simulation Results

In this section simulation results are displayed. The performance measure is performed in the normalized mean square error (NMSE). But, before that, channel tracking is done in time domain for MIMO-OFDM systems. The simulation parameters used in chapter three are also used in here. Once again they are displayed here with additional parameters because of the state model used in Kalman filtering.

5.4.1 Simulation parameters and assumptions

I. Number of transmitter, $M_t = 2$
II. Number of receiver, $M_r = 2$
III. BPSK symbol modulation
IV. Number of subcarriers, $N = 256$
V. Cyclic Prefix, $N_{CP} = N/4$
VI. OFDM block size, $N + N_{CP}$
VII. Bandwidth = 1.5MHz.
VIII. Carrie frequency is $f_c = 2.4$ GHz.
IX. Wireless channels from each transmit to each receive antenna experience Rayleigh fading with independent propagation paths.
X. The receiver speed is of 36 Km/h.
XI. The Doppler spectrum is Jake's and the channel length is 16.
XII. No correlations is considered neither at the transmitter and nor at the receiver side, which means the four channel impulse responses, h_{11},h_{12},h_{21},and h_{22} fade independently.
XIII. Channels are considered to remain stationary during the OFDM block time.

The MMSE equalizer discussed in section 5.2 is chosen, as it does not suffer from noise enhancements as the zero forcing does. Tracking capabilities of the method is illustrated with time. Figure (5.2) shows the true and estimated channels at a given OFDM block time. In the figure the real and imaginary part for each of the four paths

are shown. Note also that as many literatures claim tracking in time is turned out to be more robust to estimation errors because the frequency correlation of the taps can be efficiently exploited [1, 11, 12, and 13]. Frequency domain tracking, without any autoregressive (AR) modeling, suffers from decision errors: each subcarrier being tracked separately, a single decision error includes an erroneous feedback, provoking a permanent loss of both the channel track and the data stream on the corresponding subcarrier.

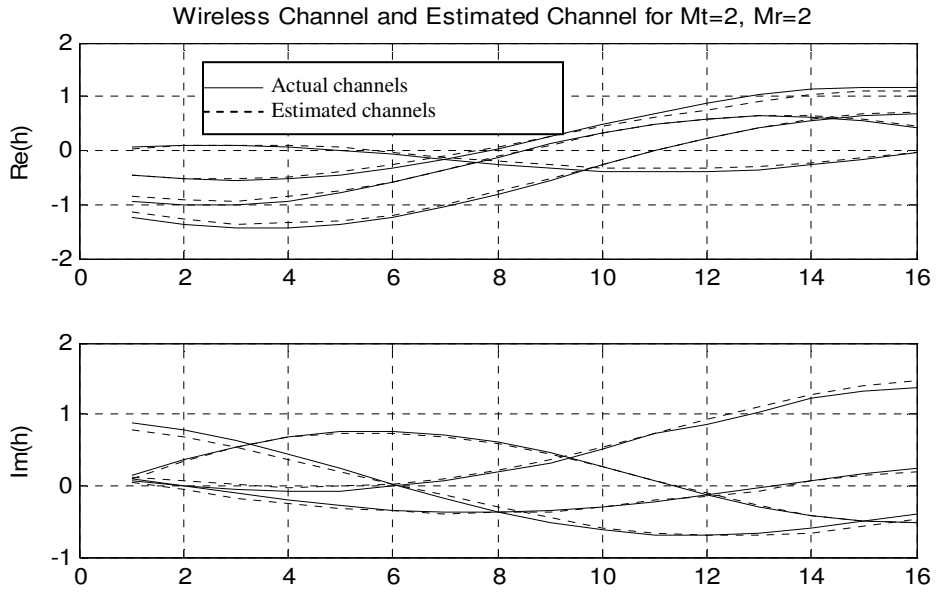


Figure 5.2 Channel tracking for the h_{11}, h_{21}, h_{12}, and h_{22} with equal delay time (L=16).

To have a clear observation for the tracking capability of our system the channel tracking is displayed for h11 in Figure 5.3. As it seen from these two figures superior tracking capability is achieved.

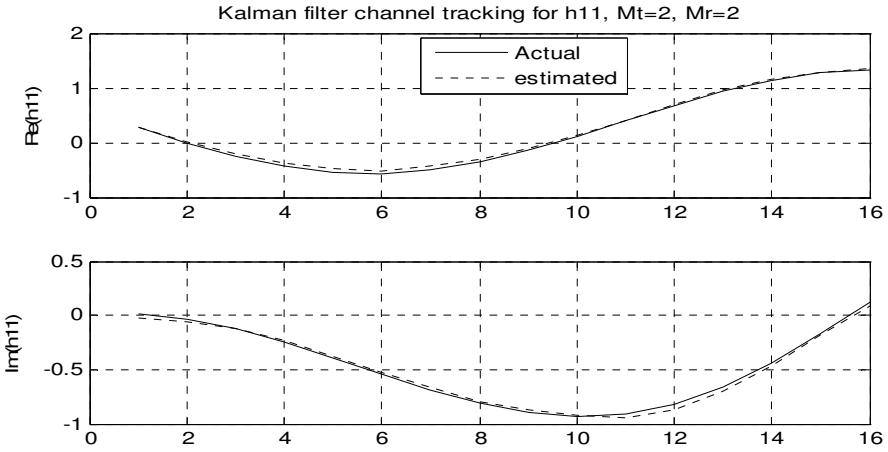


Figure 5.3 Real and Imaginary parts of h_{11}; actual and estimated value, L=16.

The NMSE for the Kalman filter is also displayed in Figure (5.4) for a channel length L = 16. As we can see from the figure the performance of our system is increased tremendously. At this point the simulation parameters, as described in chapter three, are for typical urban area scenario. Note also from Figure (5.4) that the NMSE at low SNR values are even smaller than that of the LS and MMSE estimators. This is because of the time domain tracking, appropriate model of the transmission system and proper initialization for the Kalman filtering equations.

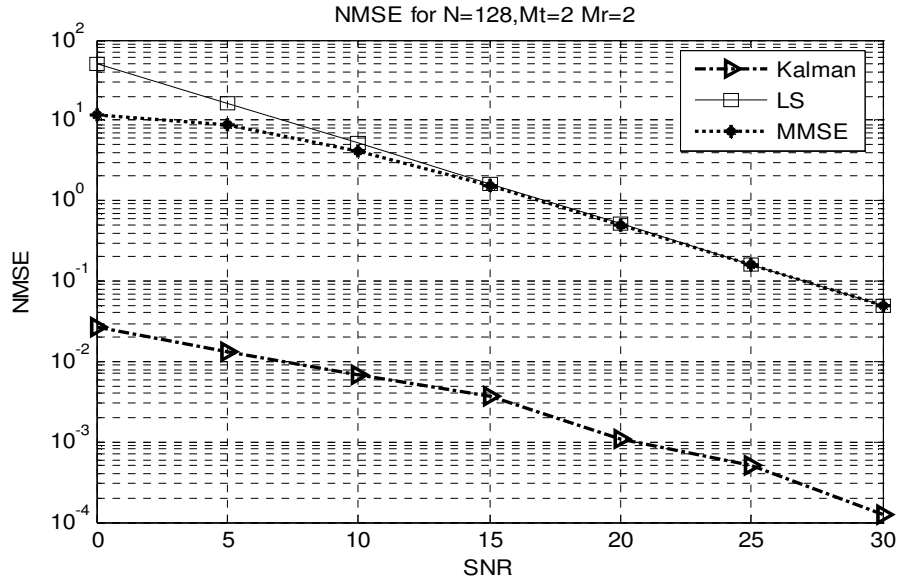


Figure 5.4 NMSE for Kalman filtering and LS, MMSE estimators

The performance of NMSE for different channel lengths are also investigated and displayed in Figure (5.5). The figure shows the results of Kalman filter for channel lengths of 21, 16, and 8. The accuracy of the estimates increased as the number of channel lengths decreased. Mathematically, a larger L implies that there are more unknown to be solved for a given set of equations, hence the accuracy will decrease.

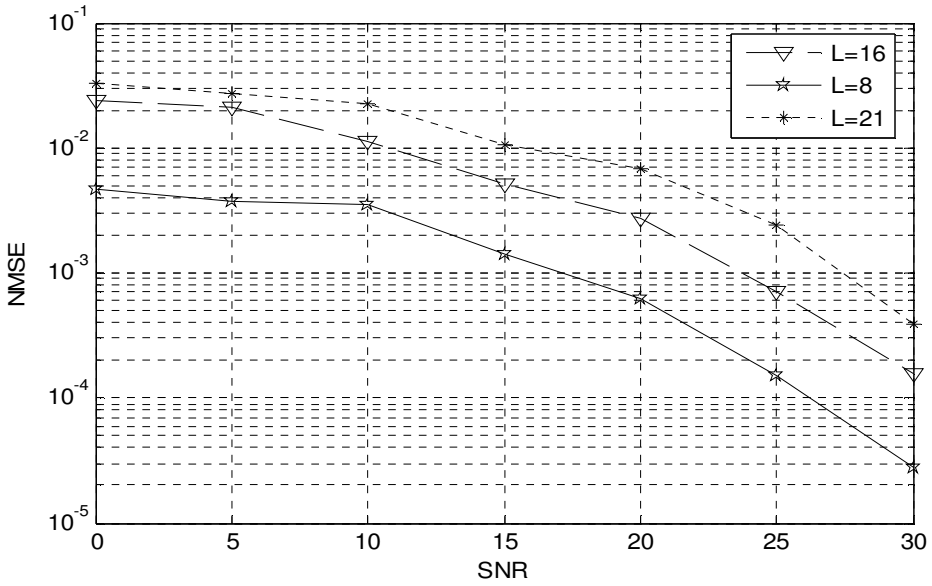


Figure 5.5 NMSE curves for Kalman filter estimator for different channel lengths.

There is also one thing remains to be discussed as long as we are dealing with mobile wireless systems. This is the performance of the filter at different mobile speed users. Figure (5.6) shows the MSE performance of Kalman filter for different mobile speed users. As expected for slower changing channels, such as the 5 km/hr case, the performance is better. Conversely, in a relatively fast changing channel of 36 km/hr the performance degrades.

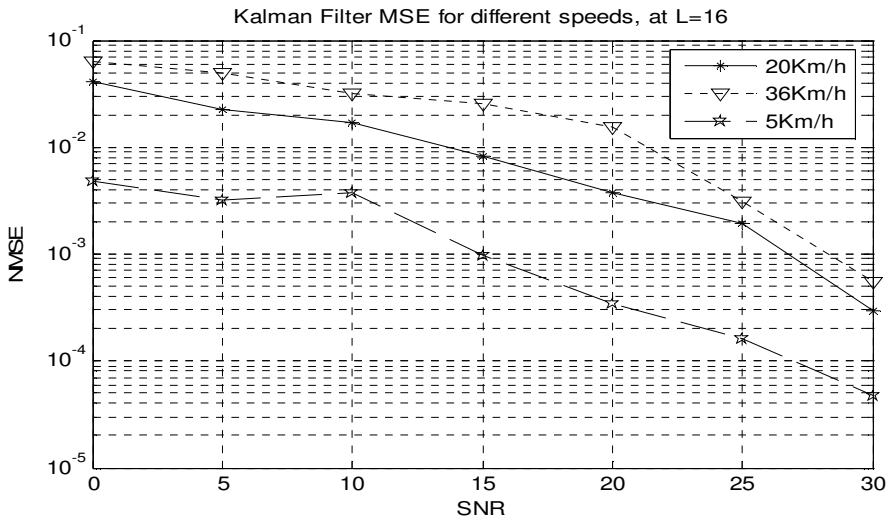


Figure 5.6 NMSE of Kalman Filter for different mobile speed users.

5.4.2 Computational complexity

The main advantage of using Kalman Filtering in the thesis is actually to improve the efficiency of channel tracking in mobile wireless MIMO-OFDM systems. In addition, the channel is time-varying and it is tracked with superior capability because of appropriate modeling of the system and the properties of Kalman filter. Still it is worth mentioning for the computational complexity of using Kalman filtering to clue researchers for reducing this complexity. As in the case of section (3.4.3) the derivations are based on an M_t - *by* - M_r MIMO-OFDM system with N subcarriers and a channel length of L. Equation (5.15) will be used for the analysis with the obvious introduction of Equation (5.14) which takes the major part of the complexity in vector Kalman filtering. It is a Kalman gain with its operation of matrix inversion. In Equation (5.3) the matrix $\boldsymbol{X}$ has a size of M_rNXM_rM, where $M = M_tL$. Table 5.1 summarizes the operational analysis in terms of operational counts for a receive antenna 1, that is $M_r = 1$. In the table the operational counts are done for each steps in the estimation for equation (5.15).

Operation	No. of Complex multiplications	No. of Complex Add/Sub	No. of Complex Division	No. of square root
A=XM(k/k-1)X^H	$MN^2 + M^2N$	NM^2-NM	*0*	*0*
B = A+R	*0*	N^2	*0*	*0*
B^{-1}	N^3	N^3	$3N^2/2+N/2$	*0*
C=M(k/K-1)X^HB	NM^2+MN^2	NM^2-NM	*0*	*0*
D=X $\hat{h}$(k/k-1)	NM^2	*NM-N*	*0*	*0*
E = r-D	*0*	*N*	*0*	*0*
F=K(k)E	N^2	N^2-N	*0*	*0*
G= $\hat{h}$(k/k-1)-F	*0*	*N*	*0*	*0*

Table 5.1 Channel Estimation Operation for Operation Count

Using the values L=5 and M_t=2 let us summarize the complexity analysis for the three channel estimators. Refer section 3.4.3 for the LS and MMSE estimators operation count.

Method	No. of Complex multiplications	No. of Complex Add/Sub	No. of Complex Division	No. of square root
LS	54,760	52100	155	0
MMSE	55,860	53,190	310	0
Kalman Filter	21,823,488	17,430,384	98,432	0

Table 5.2 Comparison for operation count in LS, MMSE and Kalman Filter

As can be seen from table (4.2) Kalman filtering has the highest computational complexity. If we see carefully on the operations all of operations contribute much of the computational complexity. However, the operations other than the division operations are computationally easy to implement in real time system. Hence as compared to the efficiency obtained in the channel tracking capability for wireless mobile users it can be said that we are still achieving a great deal of performance improvement in channel estimation for MIMO-OFDM systems and assuring major requirements of B3G wireless communication systems. The computational complexity can be a great deal to decrease, especially the inversion of the large matrix. QR decomposition is one possible solution to avoid so much complexity because of the inversions. In this thesis the performance is meant to be improved, which makes it complete if computational complexity reduction was included.

CHAPTER SIX

CONCLUSION AND FUTURE WORK

6.1 Thesis Summary and Conclusions

MIMO-OFDM has the potential to sufficiently increase the capacity and reliability of the system due to the added spatial degree of freedom from multiple independent paths. Moreover, the use of OFDM techniques makes the system robust to frequency selective channels. However, these promising features of MIMO-OFDM system have been difficult to deploy because of the complicated receiver structure due to the additional unknown parameters. This leads to the extremely complicated channel estimation schemes for MIMO-OFDM system. The main objective of this thesis was to develop an efficient channel estimation technique for these systems. For this we choose the Kalman filtering for many of its outstanding features in applications like tracking, navigation and estimations of time-varying systems.

It was claimed that training based channel estimation to be more applicable and less complex than the blind ones. To illustrate what channel estimation is meant in MIMO-OFDM systems, LS and MMSE channel estimators were discussed and their results were displayed for MSE operations. There performance was shown to be not satisfactory in the basis of some realistic parameters. Apart from their complexity in their applications they are found to be less effective in an urban type scenario. Also the analysis made showed that MMSE outperforms the LS estimation at the expense of increased complexity. However, the results also showed that at high SNR values the performance of MMSE estimator approaches that of the LS estimator.

Kalman filtering channel estimation was investigated for time domain channels. The vector Kalman filtering equations, discussed in chapter four were applied to the MIMO-OFDM systems. The models for the transmission systems were developed based on Gauss-Markov channel model and the Bayesian linear model. Incorporating these two concepts together with general channel estimation schemes the algorithm in Figure (5.2) was developed. MMSE equalizer as opposed to ZF equalizer contributes to the system efficiency, as the later suffers from noise enhancement. Due to these reasons, superior channel tracking capabilities was achieved. As per the results from the NMSE curves, it was able to observe that, for a typical urban type scenario, the

Kalman filtering performs very well for our system than the LS and MMSE estimators. In the case of computational complexity, it was shown that Kalman filtering receives the highest computational complexity which is not required in the real time applications.

6.2 Future Works

Due to limited time frame of this research, there are still some important issues that have not been dealt. The following includes list of suggested future works that can be investigated.

- The computational complexity in the Kalman filtering channel estimation can be improved by using QR decomposition in the Kalman gain equation, since matrix inversion is computationally inefficient to be implemented in hardware.
- The other one is there are still impediments in the channel estimations procedure among which effects of frequency offsets are the major one. If Extended Kalman Filtering can be used frequency offsets can be tracked along with channel estimation for MIMO-OFDM systems. In this thesis perfect synchronization of frequency is assumed.

References

1. Kathryn Kar Ying Lo, "Channel Estimation of Frequency Selective Channels for MIMO-OFDM", *dissertations*, Calgary, Alberta, *2005.*
2. Timo Roman, "Advanced Receiver Structure For Mobile MIMO Multicarrier Communication System", *Ph.D dissertations, Helsinki university of Technology,2006.*
3. Ilkka Harjula, "Channel Estimation Algorithm for Space-Time Block Coded OFDM Systems", *VTT Electronics, Finland, 2003.*
4. H. Minn, D. I. Kin, and V. K. Bhargava, " A Reduced Complexity Channel Estimation for OFDM Systems With Transmit Diversity in Mobile Wireless Channels", *IEEE transactions on Communications, vol.50, no. 5, pp. 799-807, May 2002*
5. Rupul Safaya, "A Multipath Channel Estimation Algorithm Using A Kalman Filter", *Dissertation*, *Illions Institute of Technology, 1997.*
6. Branka Vucetic, Jinhong Yuan, "Space-Time Coding", *John Wiley & Sons Ltd, 2003.*
7. M.S. Kay, " Statistical Signal Processing, Estimation theory", *Prince-Hall PTR, 1993*
8. Jan-Jaap van de Beek, Ove Edfords, Magnus Sandell, Sarah Kate Wilson and Per Ola Borjesson, " On Channel Estimation In OFDM Systems", *IEEE, In Proceedings of Vehicular Technology conference, vol. 2, PP. 815-819, Sept. 1995.*
9. Ove Edfords, " Channel Estimation Using Singular Value Decomposition",
10. Hlaing Minn, Naofal Al-Dhahir,"Optimal training signal for MIMO-OFDM Channel estimation", *IEEE, Glbecom 2004.*
11. Timo Roman, Mihai Enescu, and Visa koivunen, " Joint Time-Domain Tracking Of Channel And Frequency Offsets For MIMO-OFDM Systems", *wireless Personal Communications 31:pp 181-200, 2004.*
12. Ye (Geoffrey) Li, "Simplified Channel Estimation for OFDM Systems with Multiple Transmit Antenna", *IEEE Transactions on Wireless Communications, vol. 1, No.1, January 2002.*

13. Dieter Schafhuber, Gerald Matz, and Franz Hlawatsch, "Kalman Tracking of Time-Varying Channels in Wireless MIMO-OFDM Systems", *invited papers in Proc. 37th Asilomar conf. Signals, Systems, Computers, Nov. 2003, pp 1261-1265.*
14. Jaime Adeane, Miguel R.D. Rodrigues, Inaki Berenguer, and Ian J. Wassell, "Improved Detection Methods for MIMO-OFDM-CDM Communication Systems", *Paper, Laboratory for Communication Engineering, University of Cambridge, 2004*.
15. Enrique ulffe Whu, "MIMO-OFDM Systems for High Data Rate Wireless Networks", *Paper, Stanford University, CA 94305.*
16. Zhenlan Cheng and Dirk Dahlhaus "Time versus Frequency Domain Channel Estimation for OFDM Systems with Antenna Arrays", *Paper, Communications Technology Laboratory (ETH) Zurich.*
17. Bretrad Mugat, More de Couville, and Pierre Duhamel,"Subspace-Based Blind and Semi-Blind channel estimation for OFDM systems", *IEEE transactions on Signal Processing, Vol. 50, No 7. July 2002.*
18. Muhmet Kemal Oldemir, Husseyin Arslan and Ercument Arvas, " Adaptive Low-Rank MIMO-OFDM Channel Estimation", *Paper, Syracus University, Syracus, NY, USA, 2002.*
19. Robert Grover,Brown Patrick, Y.C. Hwang, " Introduction to Random Signals and Applied Kalman Filtering", *John Wiley & Sons Inc. 1997*.
20. Brown D.O. Anderson, John B.More, " Optimal Filtering", *Printce-Hall, Inc. 1979.*
21. Michael McGuire and Mihai Sima, " Low-Order Kalman Filters for Channel Estimation", *Dissertation, University of Victoria, Canada, 2002.*
22. Hou Xia-Yun, Zheng Bao-Yu, Xu You-Yun, Song wen- Tao, " An Improved Channel Estimation with Multipath Search For MIMO-OFDM Systems", *Journal of Zhejiang University ISSN 1009-309 pp 149-155, 2006.*
23. Ronald Raulefs, Armin Dommann, Stephen Sand, " Combining transmit Diversity Schemes for Multicarrier Systems", *Dissertation, German Aerospace Center(DLR), 2004.*

24. Song Bo-Wei, Guan Yun-Feng, Zhang wen-Yun," An Efficient training Sequence Strategy ffor Channel Estimation in OFDM Systems with Transmitter Diversity", *Journal of Zhejiang University ISSN 1009-3095 pp 613-618, 2005.*
25. Burrus N. Parks, "DFT/FFT and Convolution Algorithms, Theory and Implementations", *Texas Instruments, Inc. 1985.*
26. John G. Proakis, " Digital Communications ", *McGraw-Hill, Inc. 2001*.