A New MAP Based Channel Estimation Technique for Multiple-Input Multiple-Output (MIMO) Systems

Wajih Hoteit

A Thesis

in

The Department

of

Electrical and Computer Engineering

Presented in Partial a Fulfillment of Requirements

for the Degree of Doctor of Philosophy at

Concordia University

Montreal, Quebec, Canada

August 2006

Your file Votre référence
ISBN: 978-0-494-37754-3
Our file Notre référence
ISBN: 978-0-494-37754-3

④

ABSTRACT

A New MAP Based Channel Estimation Technique for Multiple-Input Multiple-Output (MIMO) Systems

Wajih Hoteit, Ph. D.

Concordia University, 2006

Multiple-Input Multiple-Output (MIMO) systems that provide significant increase in channel capacity is rapidly emerging as the new frontier of wireless industry. MIMO systems require the simultaneous use of multiple transmit and receive antennas to dramatically increase data rates and to improve performance reliability. An effective and practical way to approach the capacity promised by MIMO systems is to employ space-time coding (STC). It elegantly combines temporal and spatial correlation into the transmitted symbols to realize diversity and coding gains. Most STC schemes are designed for known quasi-static channels however this assumption is not always justified. MIMO channels often undergo frequency selective fading that leads to intersymbol interference (ISI), which limits the performance of MIMO systems.

The effect of imperfect channel estimation on the bit error rate (BER) of MIMO systems utilizing STC is investigated. An analysis and comparison into the BER degradations of simple transmit diversity (STD) and maximal ratio combining (MRC) schemes due to multipath channel estimation errors are presented. Closed form

expressions are derived for the BER performances of the schemes that employ an equalization process to mitigate the ISI caused by the multipath in frequency selective channel. BER curves show that the performance deterioration in the MIMO scheme outweighs the benefits achieved over the single antenna case when the channel estimation errors are large. Results expose the deleterious effects of inaccurate channel estimation on the performance of MIMO systems. Hence, the development of practical and novel channel estimation approaches are desired for MIMO systems using STC.

This dissertation introduces a new MAP based channel estimation technique that is amenable to STC scheme employing two transmit antennas and operating in multipath bandlimited channel. The complex channel parameters are treated as two real-valued tap coefficients; each taking one of M possible amplitude levels with equal probability. The proposed estimation technique is based on an iterative procedure derived through the maximum a posteriori (MAP) probability approach. Unlike classic estimation techniques, we iterate on the probabilities of the different coefficients rather than on the values of the coefficients.

Two low complexity algorithms based on the developed channel estimation technique and simple to implement in practical MIMO systems are also introduced. The performances of the two algorithms are assessed by combined analysis and simulation. Results are presented and compared against the performance of conventional channel estimation techniques. Results show that the required performance can be achieved with less number of iterations using the proposed algorithms compared to conventional techniques.

This dissertation is dedicated

to my parents,

my brother and my sisters.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude and appreciation towards my supervisor, Dr. Ahmed K. Elhakeem for his great guidance, advise, and support and above all for always believing in me. I consider myself extremely fortunate for having him as my advisors. During the years of my study period, he always followed up my research and encouraged me for more progress and contribution.

I would like to thank the examining committee of this dissertation for their invaluable comments and suggestions. I am specially grateful to Dr. Yousef R. Shayan for his valuable comments, suggestions, and encouragement.

I wish to express my appreciation to my parents, brother and sisters for their continuous help, support and encouragement.

Finally, I would like to thank my friends for their encouragement.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF ACRONYMS

APP	a posteriori probability
AWGN	additive white Gaussian noise
BER	bit error rate
BPSK	binary phase-shift keying
CNR	carrier-to-noise ratio
CSI	channel state information
EGC	equal gain combining
EM	expectation-maximization
ESR	error-to-signal ratio
FIR	finite impulse response
GCV	global communication village
ISI	intersymbol interference
LMS	least mean square
LS	least-squares
MAP	maximum a posteriori
MIMO	Multiple-Input Multiple-Output
ML	maximum likelihood
MLSE	maximum likelihood sequence estimation
MMSE	minimum mean square error

MRC	maximal ratio combining
MRRC	maximal ratio receiver combining
MSE	mean square error
ND	no diversity
OFDM	orthogonal frequency division multiplexing
PAM	pulse amplitude modulated
QPSK	quadrature phase-shift keying
pdf	probability density function
RF	radio frequency
RLS	recursive least square
SC	selection combining
SNR	signal-to-noise ratio
STBC	space-time block codes
STC	space-time coding
STD	simple transmit diversity
STTC	space-time trellis codes
ZF	zero forcing

Chapter 1

Introduction

1.1 Literature Review

More than one hundred years ago, the notion of transmitting information without the use of wires must have seemed like magic. In 1897, Marconi made it possible by demonstrating the first wireless communication system between a land-based station and a tugboat [1]. Since then, unbelievable, extraordinary, and rapid developments in the field of wireless communication have been taking place, which will shrink the world into a global communication village (GCV) by 2010 [2]. Today's wireless communication systems are limited to voice and low-speed data transmissions but multimedia services are envisaged for 21^{st} century applications [3].

Many industry practitioners believe that wireless subscribers worldwide will exceed two billion by the end of 2007. Such rapid consumer growth brings with it radio frequency (RF) interference and spectral crowding, as well as an urgent need to develop wireless systems with sufficient capacity, that would let subscribers make phone calls, surf the Web, exchange e-mail and conduct video conferences simultaneously. The most important issues in wireless multiple access techniques are the reliable and flexible transmissions of various data rates, and the efficient utilization of the limited frequency resources by as many users as possible.

The new frontier of wireless communications is rapidly becoming the introduction of Multiple-Input Multiple-Output (MIMO) diversity systems that require simultaneous use of multiple transmit and receive antennas. Deploying multiple antennas at both the transmitter and receiver achieves high date rate and provides significant increase in capacity without increasing the total transmission power or bandwidth [4], [5]. The capacity grows linearly with the smaller of the number of transmit or receive antennas, provided that perfect channel estimation is available at the receiver. However, in real world perfect channel estimation is never known a priori [6].

An interesting perspective of multiple antennas is that a channel affected by fading can be turned into an additive white Gaussian noise (AWGN) channel by increasing the number of antenna diversity branches and using maximum ratio combining (MRC). This observation that was indeed investigated and verified by analysis and simulations by the authors in [7]-[9], required the knowledge of channel estimates at the receiver. An effective and practical way to approach the promised capacity of MIMO channels is to employ space-time coding (STC). STC is a coding technique performed in both spatial and temporal domains and designed for use with multiple transmit antennas. There are various approaches in STC structure; space-time block codes (STBC) and space-time trellis codes (STTC) are the two prevailing techniques.

The first bandwidth efficient coding and modulation technique designed for multiple antenna schemes was proposed in [10], and it included the transmit diversity scheme of [11] as a special case. STTC for two transmit antennas was introduced in [12], and later generalized to more transmit antennas in [13], [14]. These STTC schemes provide full diversity with coding gain however; the coding gain is obtained at the cost of increased decoding complexity [15]. *Alamouti* proposed STBC in [16] as a full rate code for two transmit antennas, which appeared as a simple way to achieve diversity gain as MRC with low decoding complexity. Motivated by the simplicity of the Alamouti scheme, and despite the performance penalty compared with STTC, STBC now extends to an arbitrary number of transmit antennas and varying code rates [17], [18].

A central issue in all these schemes is the exploitation of multipath effects in order to achieve high spectral efficiencies and performance gains [19]-[21]. Space-time codes were originally designed and analyzed assuming flat fading channels and the availability of accurate channel estimates at the receiver. However, the assumption of known flat fading channels is not always justified, especially for MIMO wireless systems. MIMO channels often undergo frequency selective fading that causes the transmitted symbols to overlap, resulting in intersymbol interference (ISI) in the received signals. ISI significantly limits the performance of MIMO wireless systems because an irreducible bit error rate (BER) results if no mitigation mechanism is employed.

Equalization is an effective measure against ISI in wireless communications systems. The term equalization is used to describe any signal processing for finding the coefficients of a digital filter or equalizer in order to minimize distortion. The Zero forcing (ZF) and minimum mean square error (MMSE) criteria are the most used techniques to optimize the equalizer coefficients. For any sequence detection criterion, the knowledge of the channel coefficients is required, which can be provided by a separate channel estimator connected in parallel with the detector algorithm. The channel estimator is a replica of the equivalent discrete time channel filter that models the ISI. The channel estimation is based on the known training sequence of bits that is repeated in every transmission burst [22].

A broad class of equalizers, known as channel estimation based equalizers [23]-[25], require an estimate of the discrete time equivalent channel in order to minimize the BER. The performance of these equalizers have been analyzed in [26]-[30] and shown to be sensitive to the accuracy of the channel estimates. A simple channel estimation algorithm for STC was firstly proposed in [31], using observations associated with orthogonal pilot symbols. Proposed technique did not lead to efficient estimation, especially when used over a frequency selective channel. The performance analysis of the STTC in [12] was done in [32] assuming a flat fading channel. The analysis showed that the design criteria of STTC is still optimum when used over a frequency selective fading channel.

While designing a whitened filter is well known for classical equalization problem, it is not yet known when STC with multiple antennas is used. ZF and MMSE equalization schemes exploiting the structure of STBC with two transmit antennas and using highly idealistic assumptions were presented in [33]. A least-squares (LS) algorithm designed for a training sequence was employed to compute the channel coefficients. Although the ZF equalizer eliminated all interference, it produced higher error rates than the MMSE equalizer. However, it was possible to express the BER exactly with a Q-function for the ZF case, but an upper bound was given for the MMSE case.

A different approach that is receiving increasing interest recently is the investigation of joint channel estimation and data detection methods. Here the data decision obtained from decoding is used as additional training to refine the channel estimate. Based on this approach, a low complexity space-time receiver for linear MIMO channels was proposed in [34]. Results suggest that the receiver can approach within 1 dB the performance of optimal decoder with perfect channel knowledge. Combining joint data detection and channel estimation with an iterative algorithm is a method that is gaining popularity. Iterative channel estimation proposed for MIMO systems use strategies that are based on the expectation-maximization (EM) algorithm [35]-[39], overlay pilots [40], and pilot embedding [41]-[43].

These suggested approaches require a large number of pilot symbols and can be complex to implement. Due to the high computational complexity of matrix inversion and the fact that the matrix inversion has to be calculated in each iteration, optimal maximum likelihood (ML) channel estimate for MIMO channels is difficult to implement in practice [36], [38], [39]. To avoid matrix inversion, an iterative decision directed channel estimation method for MIMO systems was proposed in [44]. Simulation results suggest that at high signal-to-noise ratio (SNR), performance is within 0.5 dB of optimal decision directed channel estimate performance with known channel estimation.

In blind channel estimation techniques no explicit training signals are used. Instead the receiver estimates the channels from the signals received during normal data transmission. Substantial work has been done on blind estimation for MIMO channels and a number of leverages including cyclo-stationarity, finite alphabet, and constant modulus were used to estimate the channel. The use of blind methods has generally not been popular in practical systems. Although, a considerable amount of work has been done in this area, yet very few of the existing methods can be directly applied because of the high implementation complexity and their low suitability for practical MIMO communication systems.

Performance analysis of the Alamouti scheme [16] has been discussed in many papers in recent years, see for instance [45]-[47]. Even if a lot of effort is spent on the analysis in these and other papers, all of them assume perfect channel estimates. Since good channel estimation turns out to be very important for accurate investigation of MIMO systems performance. It is of interest to know how the performance depends on uncertainties in the channel estimates over channels with ISI. Channel estimation for MIMO systems is a major challenge and requires additional effort. When the number of antennas increases, accurate channel estimation becomes more difficult because of the increase in the number of parameters to be estimated [48]-[50]. In addition, equalization techniques for general space-time codes are still an open problem.

1.2 Problem Statements and Objectives

As mentioned in the previous section, MIMO diversity systems employ STC to achieve the promised increase in capacity of MIMO channels and to provide reliable and flexible transmissions of various data rates. However, in order to achieve high spectral efficiencies and performance gains, flat fading channels and the availability of perfect channel estimation at the receiver were assumed.

The performance analysis of MIMO schemes that utilize STC and in particular the Alamouti scheme [16] was mostly presented assuming known flat fading channels. When the inaccurate channel estimation was investigated, the effects of the ISI created by multipath within the frequency selective fading channels were ignored. Channel estimation for MIMO channels is a major challenge due to the increased number of parameters to be estimated and reduced transmit power for each transmit antenna. Substantial work has been done in this area using a number of various methods and different approaches. However, these suggested approaches can be complex to implement and yet very few of the existing methods are suitable for practical MIMO channels that suffer from channel induced ISI.

Observing the above problems, our goal is to develop a new channel estimation technique suitable for MIMO diversity systems that utilize STC. A tradeoff between implementation complexity and practical high performance channel estimator will be considered with the emphasis on frequency selective fading channels.

Four specific objectives are listed below:

- An analysis and comparison into the BER degradations of simple transmit diversity (STD) and MRC schemes due to multipath channel estimation errors are presented. Performance curves to determine the impact of channel estimation errors at various numbers of multipath components are obtained.
- The effects of imperfect channel estimation on the BER of MIMO communication systems utilizing STC are investigated. A multipath channel between each transmit and receive antenna pair that results in ISI is considered. A closed form expression for the BER performance that is a function of the channel estimation error and the multipath components is derived.
- A novel channel estimation technique for STD scheme employing STC now applied in MIMO communication systems is developed. The technique will focus on estimating the channel parameters for a two branch transmit diversity scheme operating in multipath environment that induces ISI. The proposed estimation technique will be based on an iterative procedure derived through the maximum a posteriori (MAP) probability approach. Unlike classic estimation techniques, iteration is done on the probabilities of the different coefficients rather than on the values of the coefficients.
- Low complexity algorithms that are based on the developed channel estimation technique are designed. Two practical approaches will be proposed that are also based on the MAP principle and are simple to implement in practical MIMO systems. The two suggested algorithms would be assessed by a combined analysis/ simulation and results will be compared against those of conventional techniques.

1.3 Contributions

In this dissertation, the effects of imperfect multipath channel estimation on the performance of a MIMO diversity system employing STC is investigated. Then a new technique to estimate the channel parameters for a two branch transmit diversity scheme is derived. In addition, two practical low complexity algorithms are proposed for MIMO channels that are characterized as frequency selective fading channels.

An STD scheme utilizing two transmit and one receive antennas was shown to have the same error performance in non-time-selective fading channels as MRC when perfect channel knowledge is available at the receiver. In practice, it is impossible to achieve the perfect channel estimation at the receiver. Hence, an analysis and comparison are presented into the BER degradations of the STD and MRC schemes due to imperfect multipath channel estimation at the receiver. A theoretical approach to investigate the BER performance of the STD and MRC schemes with perfect and imperfect channel estimation is proposed. BER curves will show that the performance deterioration of the STD scheme increases quite rapidly compared to the degradation in the MRC performance when the channel estimation and multipath components are incremented. Thus, the practical implementation of the STD scheme should be carefully considered under such conditions.

The effect of imperfect channel estimation on the BER of MIMO communication systems utilizing STC is investigated. A multipath channel is considered and a given level for the channel estimation error is assumed. The receiver employs an equalization process to reduce the ISI in the received signal. A closed form expression for the BER performance of the scheme that employs the results of imperfect channel estimation is derived. Results are applicable to any channel estimation technique. Performance curves show that the deterioration of performance in the multiple transmit antenna scheme outweighs the benefits achieved over the single antenna case when the SNR and channel estimation error are large. Results show the deleterious effects of inaccurate channel estimation on the performance of MIMO systems. Hence, novel channel estimation approaches are desired for MIMO communication systems using STC.

The effectiveness of STC schemes and in particular the Alamouti scheme relies on accurate multi-channel estimation at the receiver in order to achieve diversity advantage and coding gain. In reality, reducing the channel estimation error entails using a multitude of techniques. Many of those would require complex processing and a large number of iterations to reduce the channel estimation error. This may become prohibitive for real time applications such as voice and video in fast fading situations. Channel estimation for MIMO systems is a major challenge and requires additional effort; especially for MIMO channels that often undergo frequency selective fading. Thus, STC schemes require the development of practical and effective channel estimation algorithms that can accurately estimate a large number of channel parameters.

A new MAP based channel estimation technique that is amenable to the STC scheme employing two transmit antennas and operating in multipath fading channels that result in ISI is proposed. The channel parameters are the attenuation and delay incurred by the signal transversal along the propagation paths. The complex channel parameters are treated as two real-valued tap coefficients; each taking one of M possible amplitude levels with equal probability. First, we derive the various expressions required to compute the a posteriori probabilities for each coefficient. Based on the MAP criterion, then select as a coefficient value the one that gives the largest probability. The obtained algorithm involves a large amount of computation per received signal, especially if the number M of amplitude levels $\{A_m\}$ is large. To alleviate the computational operations, two practical algorithms are introduced instead.

In the first algorithm, we reduce the number of multiple summations used to perform averaging over all coefficients except the one being optimized in the current iteration. The optimization of the first coefficient would still require averaging over the still unknown probabilities of all possible levels of all other coefficients. Then assign as a first coefficient value the level that corresponds to the largest probability. In estimating the following channel coefficients, we use the selected value of the previously estimated coefficients and set the probabilities of previously estimated levels of the channel coefficients to one. This reduces the number of summations as we proceed, until reaching the last coefficient, which will be optimized using a simple expression. In the second algorithm, a further simplification is achieved by randomly selecting one of the assumed levels for each channel coefficient except the first one. To improve the reliability of the initial estimates of the coefficients, different iterative procedures are used for the two algorithms. The performance of the two algorithms has been assessed by a hybrid analysis and simulation based technique. Simulation results show that significant improvement over conventional channel estimation techniques can be obtained using these iterative procedures.

1.4 Organization

This dissertation is organized as follows. In Chapter 2, we discuss preliminaries that will be useful for the subsequent development of this dissertation. In Chapter 3, an analysis and comparison into the BER degradations of the STD and MRC schemes due to channel estimation errors are presented. Numerical results and discussion of the impact of various channel estimation errors and multipath components on the BER curves of the two schemes are provided. In Chapter 4, the effects of imperfect multipath channel estimation on the performance of MIMO system using STC are examined. The error performance and numerical results of the MIMO system are presented. In Chapter 5, a new channel estimation technique for MIMO systems utilizing STC is introduced. Discrete time model of the MIMO channel that is subject to ISI is described. Expressions required to compute the a posteriori probabilities for each channel coefficient are derived. In Chapter 6, two practical algorithms to reduce computational operations are proposed. Combined analysis/simulation results are presented and compared against the performance of conventional channel estimation techniques. Conclusions and future works are introduced in Chapter 7.

Chapter 2

Background Review

2.1 Wireless Communications

Wireless communications is a broad and dynamic field that has spurred tremendous excitement and technological advances over the last few decades. It is the fastest growing segment of the communication industry. The wireless channel is an unpredictable and difficult communication medium. The characteristics of the channel appear to change randomly with time, which make it difficult to design reliable systems with guaranteed performance.

2.1.1 Fading Channels

Wireless channels pose a severe challenge as a medium for reliable high-speed communication. The fundamental phenomenon that makes reliable wireless communications difficult is the time-varying multipath fading [51]. A signal propagating through the wireless channel usually arrives at the destination along a number of different paths, referred to as multipath. These paths arise from scattering, reflection, refraction, or diffraction of the radiated energy off objects in the environment. The received signal is much weaker than the transmitted signal due to mean propagation loss, long-term fading and short-term fading. The mean propagation loss arises from square law spreading, absorption by foliage, and the effect of ground-generated vertical multipath.

Long-term fading, also known as shadowing, results from signal blocking by buildings and natural features. On the other hand, short-term fading results from multipath in the vicinity of the mobile. Multipath propagation results in the spreading of the signal in different domains, including delay (or temporal) spread, Doppler (or frequency) spread, and angle spread. These spreads have significant effects on the signal. The mean path loss, long-term fading, short-term fading, delay spread, Doppler spread, and angle spread are the main channel effects.

In addition to mean path loss, the received signal exhibits fluctuations in signal level called fading. Fading can be modeled statistically with probability distributions as noted in [52]-[55]. In addition to a statistical description of the fading channel, we can describe the severity of fading in the time, frequency, and spatial domains. These lead to the following different channel characterizations:

- *Flat Fading*: A channel is said to exhibit flat fading if the channel has a constant gain and linear phase response over a bandwidth, which is greater than the bandwidth of the transmitted signal. In this case, all of the received multipath components of a symbol arrive within the symbol time duration; hence the spectral characteristics of the transmitted signal are preserved at the receiver. Flat fading channels are sometimes referred to as *narrowband* channels.

- *Frequency Selective Fading*: A channel is said to exhibit frequency selective fading if the signal bandwidth is greater than the channel coherence frequency, which is the range of frequencies over which the channel passes all spectral components with approximately equal gain. Frequency selective fading channels also known as *wideband* channels. In this case, the received signal includes versions of the transmitted waveform, which are attenuated and delayed in time. Frequency selective fading gives rise to intersymbol interference (ISI); this ISI results in an irreducible error floor that is independent of signal power. The first extensive analysis of the degradation in symbol error probability due to ISI was done by *Bello* and *Nelin* [56].
- *Slow Fading*: A channel is referred to as slow fading if the time duration of the transmitted symbol is smaller than the channel coherence time, which is the expected time duration over which the channel response is essentially invariant. In this case, the channel may be assumed to be static over one or several symbols duration. Slow fading channels are also referred to as *quasi-static* fading channels.
- *Fast Fading*: A channel undergoes fast fading if the time duration of the transmitted symbol is greater than the channel coherence time. In fast fading, the fading character of the channel will change several times during the time span of a symbol, leading to channel induced ISI. Hence, the received signals will be

distorted often resulting in an irreducible error rate.

These channel characterizations are not mutually exclusive, thus a channel could exhibit both fast fading and flat fading or slow and flat fading. Similarly, a channel could be specified as a fast fading and frequency selective fading channel or slow fading and frequency selective channel [57], [58]. Often the time varying impulse response channel model is too complex for simple analysis. In this case a discrete-time approximation for the multipath model can be used. In fact most continuous-time systems can be converted to discrete-time systems via sampling. However, care must be taken in choosing the appropriate sampling rate for this conversion.

A discrete-time model that approximates a wide range of multipath environments has been developed by *Turin* [59], which has been successfully used in mobile radio communications. The discrete-time model is given by:

$$h(t) = \sum_{i=0}^{N-1} \alpha_i \, \delta(t - iT) \tag{2.1}$$

where N is the number of multipath components (bins), $\delta(.)$ is the Dirac delta function, and α_i is the complex gain associated with the ith multipath component. The statistics of α_i have been characterized empirically by *Turin* [59] for wireless channels.

2.1.2 Diversity Techniques

Severe attenuation in a multipath wireless environment makes it extremely difficult for the receiver to determine the transmitted signal unless the receiver is provided with some form of diversity. Diversity techniques are widely used in wireless communications to reduce the effects of multipath fading and improve the reliability of transmission without increasing the transmitted power or sacrificing the bandwidth. Diversity techniques can be applied in the transmitter or the receiver. Diversity implies the use of more than one copy of the signal, such that a more robust signal results in reduction of the effects of short-term fading. *A. de Haas* first discovered this effect while experimenting with two spaced antennas [60].

It was observed that the probability that the signals on both antennas will fade simultaneously is much smaller relative to the probability of fading on either of channels alone. Examples of diversity techniques are (but are not restricted to) introduced below.

- *Temporal Diversity*: Channel coding in conjunction with time interleaving is used. Thus replicas of the transmitted signal are provided to the receiver in the form of redundancy in temporal domain.
- *Frequency Diversity*: The fact that waves transmitted on different frequencies induce different multipath structure in the propagation media is exploited. Thus replicas of the transmitted signal are provided to the receiver in the form of redundancy in the frequency domain.
- *Antenna Diversity*: Spatially separated or differently polarized antennas are used. The replicas of transmitted signal are provided to the receiver in the form of redundancy in spatial domain. This can be provided with no penalty in bandwidth efficiency.

In most scattering environments, antenna diversity is a practical, effective and widely applied technique for reducing the effect of multipath fading [51]. The classical approach is to use multiple antennas at the receiver and perform signal recombining. If several replicas of the transmitted signal over uncorrelated fading channels are available, then the receiver can exploit them to increase signal-to-noise ratio (SNR) and thereby reduce the bit error rate (BER). While both diversity and coding improve system performance, the nature of these gains is very different. Diversity gain improves the slope of the BER curve, while coding gain shifts the BER curve to the left. In receiver diversity, the independent fading paths associated with multiple antennas are combined to obtain a signal that is then passed through a standard demodulator. The method of combining two or more independent fading signals is termed as diversity combining. There are several diversity combining techniques:

1. Selection Combining (SC) means that the best of the two or more received signals is selected according to signal level, power or SNR.
2. Switched Combining means that one of the diversity signals is selected, based on a given threshold level in one receiver. This signal is received until it falls below threshold level and the process is again initiated.
3. Maximal Ratio Combining (MRC) first proposed by *Kahn* [61], the signals from all diversity branches are weighted and then combined. The SNR is maximized at the combiner output and it produces the best statistical reduction of fading of any known linear diversity combiner.
4. Equal Gain Combining (EGC) means that baseband signals are summed. In EGC there is no signal scaling like in the MRC, the branch weights are all set to unity

but the signals are co-phased to provide EGC diversity.

Receiver diversity is a well-known technique to improve the performance of wireless communications in fading channels. A large body of work has addressed approximations and numerical techniques for computing the integrals associated with the average probability of symbol error for different modulations, fading distributions, and combining techniques of receiver diversity (see [62] and the references therein). The main advantage of receiver diversity is that it mitigates the fluctuations due to fading so that the channel appears more like an additive white Gaussian noise (AWGN) channel. However, the major problem with using the receive diversity approach is the cost, size, and power of the remote units. The use of multiple antennas and radio frequency (RF) chains makes the remote units larger and more expensive.

As a result, diversity techniques have almost exclusively been applied to the base stations to improve their reception quality. Transmit diversity is desirable in systems where more space, power, and processing capability is available on the transmit side than on the receive side, as best exemplified by cellular systems. For this reason, transmit diversity schemes are very attractive. In transmit diversity there are multiple transmit antennas and the transmit power is divided among these antennas. Transmit diversity has been studied extensively as a method of combating impairments in wireless fading channels [4], [5], [6], [11], [12], [32], [63]-[67]. It is particularly appealing because of its implementation simplicity and the feasibility of multiple antennas at the base station.

2.2 MIMO Systems

Transmit diversity can be combined with receive diversity to further improve the system performance. Band-limited wireless channels are narrow pipes that do not accommodate rapid flow of data. Deploying multiple transmit and receive antennas broadens these data pipes. Communication systems with multiple antennas at the transmitter and receiver are commonly referred to as Multiple-Input Multiple-Output (MIMO) systems.

2.2.1 MIMO Channels

In MIMO systems the multiple antennas can be used to increase data rates through multiplexing and to improve performance through diversity gains. The growing demand for MIMO communication systems makes it important to determine the capacity limits of the underlying channels for these systems. The maximum error-free data rate that a channel can support is called the channel capacity. The channel capacity for AWGN channels was first derived by *Shannon* in 1948 [68]. In contrast to scalar AWGN channels, MIMO channels exhibit fading and encompass a spatial dimension. The capacity results for MIMO channels have been developed only in the past few years.

The research activity in this area stems from the breakthrough results in [4], [5], [64], [65] indicating that the capacity of a MIMO fading channel is N times larger than the channel capacity without multiple antennas, where $N = min\ (N_t, N_r)$ for N_t and N_r the number of transmit and receive antennas, respectively. These predictable spectral efficiency gains require accurate knowledge of the MIMO channel at the receiver and sometimes at the transmitter as well. Where there is no accurate estimation of the channel at either the transmitter or the receiver, the linear growth in capacity as a function of transmit and receive antennas disappears, and in some cases adding additional antennas provides negligible capacity gain.

The capacity of general fading channel models without transmitter or receiver channel state information (CSI) was investigated in [69]-[71]. These works indicate that at moderate to high SNR, capacity is limited by channel estimation error. At very high SNR, there is no capacity gain for slowly varying channels without receiver CSI. The multiple antennas can be used to obtain array and diversity gain instead of capacity gain, which is also referred to as multiplexing gain. The diversity-multiplexing trade-off or the trade-off between data rate, probability of error, and complexity of MIMO systems has been extensively studied in terms of space-time code designs [12], [72]-[74]. This work has primarily focused on block fading channels with receiver CSI only.

Consider a MIMO system with N_t transmit and N_r receive antennas as shown in Fig. 2.1. The MIMO channel must be modeled properly in order to examine the performance of the system. The channel models that are considered in this thesis are presented in this section. The primary MIMO channel model under consideration is the quasi-static, Flat or frequency non-selective, and frequency selective fading channel model.

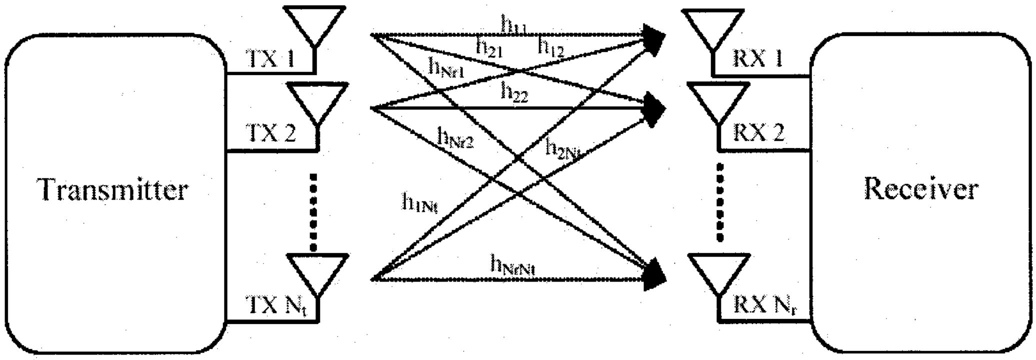

Figure 2.1: MIMO system model.

The MIMO channel is described by an $N_r \times N_t$ complex matrix, denoted by H that is given by

$$H = \begin{bmatrix} h_{11}(t) & \cdots & h_{1N_t}(t) \\ \vdots & \ddots & \vdots \\ h_{N_r1}(t) & \cdots & h_{N_rN_t}(t) \end{bmatrix} \tag{2.2}$$

where $h_{ij}(t)$ is the channel impulse response between the jth transmit and the ith receive antennas. For a flat fading channel, a single path exists between each transmit and receive antenna pair. If a quasi-static channel is further assumed then the channel remains constant over the length of a frame, we can thus write

$$h_{ij}(t) = h_{ij}(t + T_f) = \alpha_{ij} \quad, i = 1, \ldots, N_r \,, j = 1, \ldots, N_t \tag{2.3}$$

where T_f is the frame time and α_{ij} is the complex fading coefficient representing the channel path gain from transmit antenna j to receive antenna i.

MIMO channels are often frequency selective that results in significant number of resolvable paths (L). Such channel can be modeled as a finite impulse response (FIR) filter with memory L. If the frequency selective MIMO channel is assumed to be quasi-static, then the impulse response of the channel between transmit antenna j and receive antenna i can be written as

$$h_{ij}(t) = \sum_{l=0}^{L-1} \alpha_{ij}(l)\delta(t-\tau_l) \quad ,i=1,\ldots,N_r\,, j=1,\ldots,N_t \tag{2.4}$$

where $\delta(.)$ is the Dirac delta function, $\alpha_{ij}(l)$ and τ_l are the complex path gain and time delay of the lth path, respectively. If the antennas on both ends of the MIMO system are separated by more than half of a wavelength, it is usually safe to assume that their path gains are independent of each other [75].

Equalization is a signal processing technique used at the receiver to alleviate the ISI problem caused by multipath delay spread. However, equalization is much more complex in MIMO channel because it must be done over both time and space. The structure of the code can be used to convert the MIMO equalization problem to a single input single output problem for which well-established equalizer designs can be used [75]. Equalization function requires an estimate of the channel impulse response to cancel ISI. A known training sequence is sent in order for the receiver to reliably estimate the channel. Zero forcing (ZF) and minimum mean square error (MMSE) criteria are the most used criteria for linear equalization. The ZF criterion enables the equalizer to completely suppress ISI but can lead to noise enhancement. The MMSE criterion provides a better balance between ISI mitigation and noise enhancement but does not lead to a closed form expression for the BER. While under the ZF criterion the BER can be expressed with Q-function [33].

2.2.2 Space-time Coding

Space-time coding (STC) is a coding technique designed for wireless systems that employ multiple transmit antennas and single or multiple receive antennas. STC elegantly combines temporal and spatial correlation into the transmitted symbols in an effective way to approach the promised capacity of MIMO channels. STC also realizes diversity and coding gain without increasing the total transmitted power or transmission bandwidth. Space-time trellis codes (STTC) and space-time block codes (STBC) are the two prevailing STC techniques.

STTC is an extension of conventional trellis codes that have been primary applied to MIMO systems. STTC combine modulation and trellis coding to achieve full diversity and offer substantial coding gain on flat fading MIMO channels [12]-[14]. However, these gains are obtained at the cost of increased decoding complexity [15]. Fig. 2.2 shows a block diagram of STTC scheme with N_t transmit and N_r receive antennas. The space-time trellis encoder maps the source binary data into N_t streams of modulation symbols that are simultaneously transmitted using N_t transmit antennas.

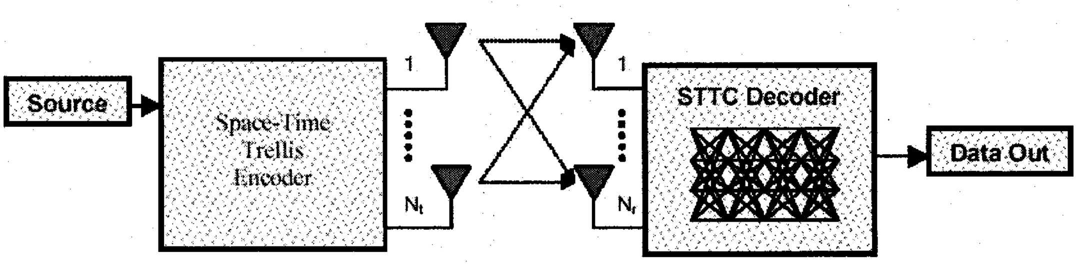

Figure 2.2: Block diagram of space-time trellis codes scheme.

The received signal at each receive antenna is a linear superposition of the N_t transmitted symbols perturbed by noise.

STBC are an alternative technique that can also extract diversity gain with low decoding complexity. Originally proposed by the Alamouti scheme, which yields the same diversity advantage as MRC scheme with linear receiver complexity. Fig. 2.3 shows a block diagram of STBC scheme with two transmit and one receive antenna.

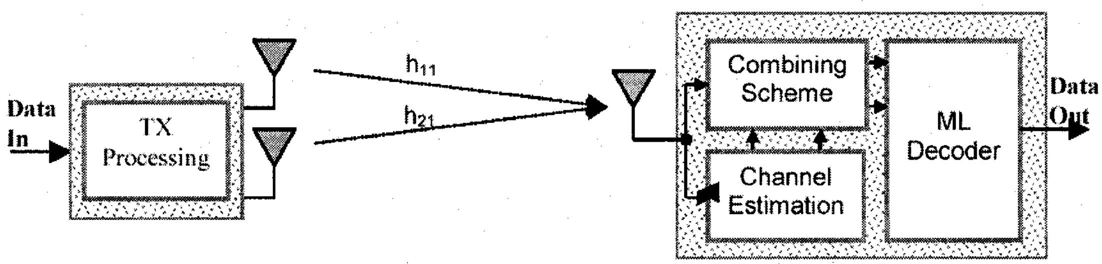

Figure 2.3: Block diagram of space-time block codes scheme.

This scheme was generalized to STBC that achieve full diversity order with an arbitrary number of transmit antennas in [17], [18]. However, while these codes achieve full diversity order, they do not provide coding gain and thus have inferior performance to STTC, which achieve both full diversity and coding gains.

Most space-time codes are designed for known flat fading channels however; successful implementation of STC over frequency selective channels requires the development of practical and high performance signal processing algorithms for channel estimation. This task is challenging due to the long delay spread of frequency selective channels, which increases the number of channel parameters to be estimated.

2.3 Estimation Theory

Estimation is the process of extracting information from data, which can be used to infer the desired information and may contain errors. The development of data processing methods for dealing with random variables can be traced back to *Gauss* who invented the technique of deterministic least-squares (LS). LS error techniques were also devised independently by *Legendre* as a method for estimating parameters from noisy measurements. The next significant contribution to the broad subject of estimation theory occurred when *Fisher* introduced the approach of maximum likelihood estimation. *Wiener* set forth a procedure for the frequency domain design of statistical optimal filters. *Kalman* and others advanced optimal recursive filter techniques based on time domain formulations. This approach, now known as *Kalman* filter, is in essence a recursive solution to *Gauss* original LS problem [76].

Estimation problems can be roughly stated as the approximation of an unknown quantity from a combination of known quantities. An optimal estimator is a computational algorithm that processes data to deduce a minimum error estimate of information produced at another source by utilizing knowledge of the source and assumed statistics of the system noise. A performance criterion that measures the quality of the estimation and leads to practical implementations is the mean square error (MSE). Often this criterion results in optimum or near optimum performance compared to other measures of performance. In the case of estimating a signal received in white Gaussian noise, minimization of the MSE is equivalent to maximum likelihood estimation.

2.3.1 Likelihood Decision

For a signal transmitted over an AWGN channel, the Bayes' theorem allows us to expresses the a posteriori probability (APP) in terms of a continuous-valued random variable x in the following form:

$$P(d=m \mid x)=\frac{p(x \mid d=m)P(d=m)}{p(x)} \qquad m=1,\ldots,M \tag{2.5}$$

and

$$p(x)=\sum_{m=1}^{M} p(x \mid d=m)P(d=m) \tag{2.6}$$

where $p(x|d = m)$ indicates the probability density function (pdf) of a received continuous-valued data-plus-noise signal, x, conditioned on the signal $d = m$, and $d = m$ represents data d belonging to the mth signal class from a set of M classes. The a priori probability $P(d = m)$, is the probability of occurrence of the mth signal. The random variable x represents the observed signal at the output of a demodulator or some other signal processor. The pdf of the received signal x, denoted by $p(x)$, is obtained by averaging over all the classes in the space. The computation of the APP, $P(d = m|x)$, is based on the observation of the received signal, and requires the knowledge of a priori probability $P(d = m)$ and the conditional pdf $p(x|d = m)$, or some statistical knowledge of the signal classes to which the signal may belong. The calculation of the APP can be thought of as a "refinement" of our prior knowledge about the data.

Consider the transmission of binary signal over an AWGN channel. The transmitted data bit is indicated by the variable d, which takes on values +1 or −1 that correspondence to the binary elements 1 or 0, respectively. The conditional pdfs also known as likelihood functions are illustrated in Fig. 2.4. The functions $p(x|d=+1)$ and $p(x|d=-1)$ are used to represent the pdf of the random variable x given that $d=+1$ and $d=-1$ were transmitted, respectively. In Fig. 2.4, one such arbitrary value x_k is shown where the index denotes an observation in the kth time interval. A well-known hard-decision rule, known as *maximum likelihood* (ML), is to choose the data $d_k=+1$ or $d_k=-1$ associated with the larger of the two intercept values ℓ_1 or ℓ_2, respectively. For each data bit at time k, this is tantamount to deciding that $d_k=+1$ if x_k falls on the right side of the decision line labeled γ_0, otherwise deciding that $d_k=-1$ [58].

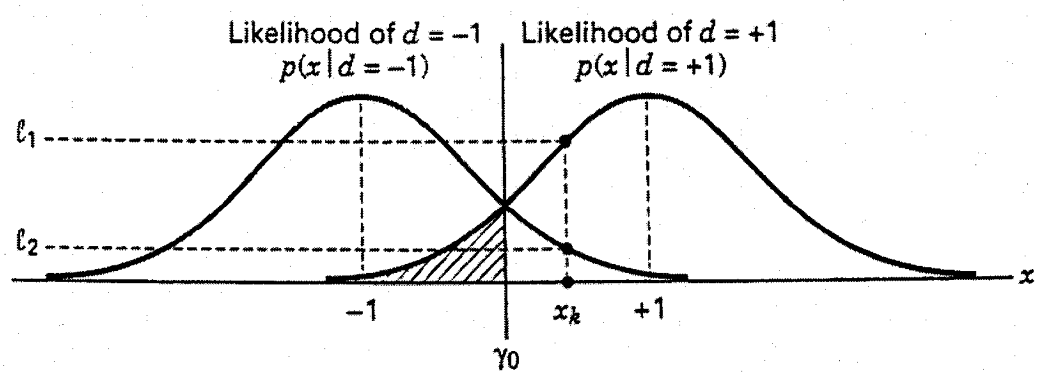

Figure 2.4: Likelihood functions.

A similar decision rule, which takes into account the a priori probability of the data, is called the *maximum a posteriori* (MAP). It can be shown that this rule minimizes the probability of error. Labeling the two possible choices as H_0 and H_1; when a 0 is sent we call it H_0, when a 1 is sent we call it H_1, The general expression for the MAP rule in terms of APPs is

$$p(d=+1\,|\,x) \underset{H_0}{\overset{H_1}{\gtrless}} p(d=-1\,|\,x) \tag{2.7}$$

The decision rule indicates that one should choose the hypothesis H_1, if $P(d=+1\,|\,x)$ exceeds $P(d=-1|x)$. Otherwise, one should choose hypothesis H_0. We can be replace the APPs in (2.7) by their equivalent expressions using the Bayes' theorem of (2.5), yielding

$$p(x\,|\,d=+1)P(d=+1) \underset{H_0}{\overset{H_1}{\gtrless}} p(x\,|\,d=-1)P(d=-1) \tag{2.8}$$

where the pdf $p(x)$ appearing on both sides of the inequality in (2.5) has been canceled.

The principle advantage of LS algorithms compared to MAP and ML estimation techniques, is that they require little information on the statistics of the data, and are usually simple to implement. MAP estimation, on the other hand, requires both the a priori probability pdf and a posteriori pdf of the random variable to be estimated. ML estimation assumes that the a priori pdf is unavailable. However, as a result of the relaxed statistical description required for the LS methods, these methods do not always provide the best performance [77].

2.3.2 Discrete Model of a Continuous-Time System

Digital data transmission involves the transmission of an information sequence consisting of discrete symbols through a bandpass channel. The channel used in this thesis is characterized in general as a quasi-static channel. Thus, the channel can be considered as being essentially time invariant over a large number of signaling intervals. A mathematical model used to represent the digital communication system is depicted in Fig. 2.5. Digital signal **a** is transmitted over a fading multipath channel **h** with L different paths. The channel is characterized as frequency selective case with maximum channel length of L, after which the signal has a memory of L symbols. Flat fading channel corresponds to L equal to zero case.

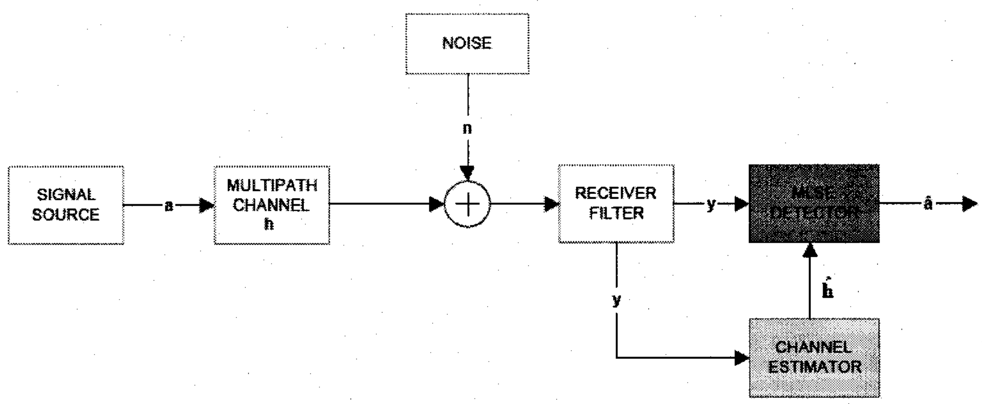

Figure 2.5: Model of digital communication system.

At the output of the channel noise **n** is added, which is assumed to be AWGN with zero mean and variance σ_n^2. At the receiver, the output of a filter matched to the received signal is sampled every signaling interval then passed through a detection algorithm to detect the transmitted bits **a** from the received signal **y**. Besides the received signal the detector needs also the channel estimates $\hat{\mathbf{h}}$, which is provided by a specific channel estimator device. Since the transmitter and receiver operate with discrete time symbols, it is reasonable to develop an equivalent discrete-time model of the communication system. The cascade of the transmitter filter, the channel, the receiver filter, and the sampler can then be represented as a discrete-time transversal filter having L tap coefficients spaced at the sampling interval [77].

The receiver uses a known training sequence to estimate the L unknown channel coefficients. In practice, these channel coefficients are continuous in terms of amplitude characteristics as a function of discrete-time. To obtain a model suitable for the utilization of the MAP based channel estimation technique, it is convenient to use a sampling representation. Also, envisioning a digital implementation of the receiver algorithms, a discrete amplitude model is necessary for the channel coefficients, such that each coefficient takes one of M values, equally spaced around zero.

There are several quantization methods designed to quantize signals depending on the characteristics of the signals source [22]. Amplitude quantizing is the task of mapping samples of a continuous amplitude waveform to a finite set of amplitudes. A typical uniform quantization operation that exhibits equally spaced quantized output levels of a sampled waveform is shown in Fig. 2.6.

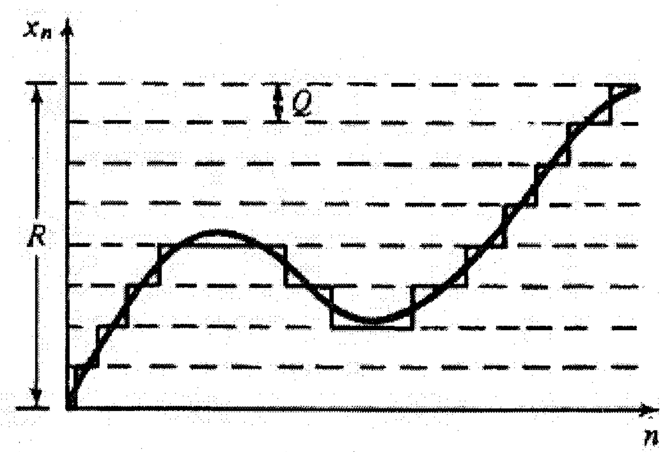

Figure 2.6: Uniform quantizer.

The quantizer allocates M levels to the task of approximating the continuous range of inputs with a finite set of outputs. The spacing between levels and the overall range of amplitude variation of the signal are denoted by Q and R, respectively. Each sample of the signal is assigned to the quantization level closest to the value of the sample. The process replaces the true signal with an approximation; the approximation will result in an error no larger than Q/2 or –Q/2 in either direction. Quantization of the amplitudes of the signal results in some distortion of the waveform. This distortion is referred to as quantization noise, which is inversely proportional to the number of levels used in quantization process.

If we assume that the quantization error is uniformly distributed over a single quantile interval Q-wide. The quantizer variance that represents the quantizer noise or error power for the zero mean error is found to be [58]

$$\sigma^2 = \frac{Q^2}{12} \tag{2.9}$$

It is intuitively satisfying to see that the quantization noise degrades as a function of the quantile interval (Q is less than one) squared. A typical quantile interval of 0.2 and 0.1 will result in quantization noise of 3.3×10^{-3} and 8.3×10^{-4}, respectively. The performance quality, on the basis of discrete channel coefficients estimation, will have no significance degradation when adding the negligible quantization error to the MSE arising from our channel estimation technique. In addition, one should expect that the channel estimation methods dealing with continuous valued coefficients have to be implemented in one quantized form or another, thus suffering from the same quantization error.

2.3.3 Channel Estimation

Channel estimation is an important part of a communications system. Channel estimates are required by detection algorithms such as maximum likelihood sequence estimation (MLSE) or MAP estimation. Channel estimates are also needed by equalizers that minimize the BER to ensure successful removal of ISI and can be used to compute the coefficients of lower complexity equalizers such as ZF or MMSE equalizers [57]. To estimate the unknown channel, a channel estimator is connected in parallel with the detection algorithm, as shown in Fig. 2.5. The channel estimator is a replica of the equivalent discrete-time channel filter discussed in the previous Section. Usually a training sequence of known bits is sent and repeated in every transmission burst for the purpose of channel estimation. The channel estimator is able to reliably estimate the channel for each burst separately by exploiting the known symbols and the corresponding received samples [22].

STC schemes run open loop that makes them very attractive for MIMO wireless transmission. Nevertheless, CSI is still required at the receiver to perform key receiver functions as decoding and equalization. Since our focus in this work is on quasi-static channels, we consider STC processing, where CSI is estimated at the receiver using a training sequence embedded in each transmission block. If the channel has delay, more channel parameters have to be estimated. Additional challenge in channel estimation for multiple transmit antenna systems are the increased number of channel parameters (proportional to N_t) to be estimated at reduced transmit power (by a factor N_t) for each transmit antenna [6].

The problem of channel estimation for flat fading space-time channels was first considered in [31] using orthogonal pilot symbols. A channel estimation algorithm was proposed that used only observations associated with pilot symbols, however such technique did not lead to efficient estimation. Channel estimation for frequency selective space-time channels has been often done using interpolation techniques. Usually the receiver estimates the channel at discrete points adequately spaced in time or frequency, then the full channel is determined through interpolation. There are a few different approaches to channel estimation such as LS or MMSE methods.

Consider a MIMO system utilizing STC transmission over quasi-static channel with two transmit and one receive antennas that is only corrupted by additive noise. Fig. 2.7 illustrates the system model for training based MIMO channel estimation methods. Training symbols are transmitted from each transmit antenna simultaneously in blocks of length M_t symbols. The channel between each transmit and receive antenna pair is modeled as frequency selective with channel length of L.

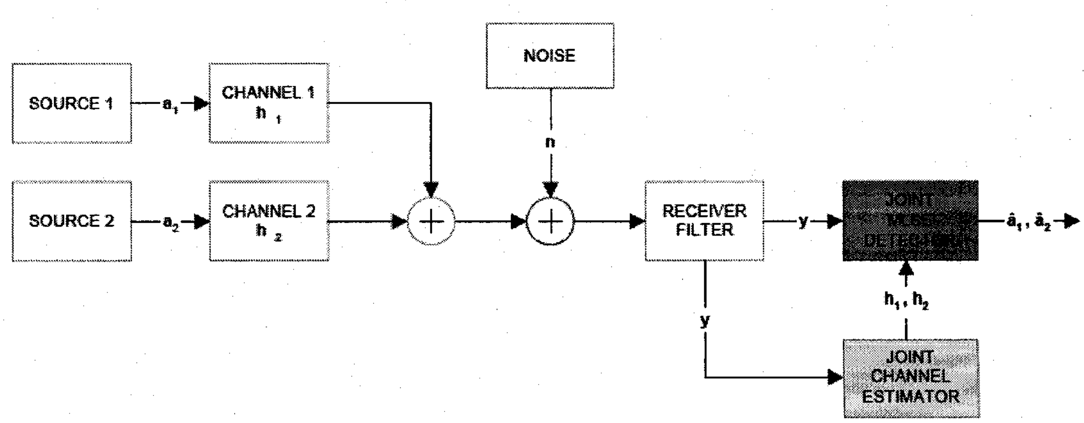

Figure 2.7: System model for a training based MIMO channel estimation.

Two known training sequences $\mathbf{a}_1$ and $\mathbf{a}_2$ are transmitted simultaneously from the first and second antennas over two channels $\mathbf{h}_i = [h_1(0) \quad \cdots \quad h_2(L-1)]^T$ for i = 1, 2, where $(\ )^T$ stands for transpose. An AWGN noise $\mathbf{n}$ with zero mean and variance $\sigma_n^2/2$ per dimension is added to the sum of the two signals to form the received signal $\mathbf{y}$. The receiver uses the training sequences to estimate the $2L$ unknown channel coefficients, which are assumed to remain constant over the transmission block. The vector of observation for the received signal $\mathbf{y}$ can be written in matrix form as

$$\mathbf{y} = \mathbf{Ah} + \mathbf{n} = [\mathbf{A}_1 \quad \mathbf{A}_2]\begin{bmatrix}\mathbf{h}_1 \\ \mathbf{h}_2\end{bmatrix} + \mathbf{n} \tag{2.10}$$

where $\mathbf{y} = [y(L) \quad \cdots \quad y(M_t)]^T$ and $\mathbf{n} = [n(L) \quad \cdots \quad n(M_t)]^T$ each with dimension $(M_t - L + 1) \times 1$. $\mathbf{A}_1$ and $\mathbf{A}_2$ are an $(M_t - L + 1) \times L$ Toeplitz matrices consisting of the training symbols,

$$\mathbf{A}_i = \begin{bmatrix} a_i(L-1) & \cdots & a_i(0) \\ a_i(L) & \cdots & a_i(1) \\ \vdots & \ddots & \vdots \\ a_i(M_t-1) & \cdots & a_i(M_t-L) \end{bmatrix}, \quad i = 1,2 \tag{2.11}$$

The LS estimate of the channel is found by minimizing the following quantity

$$\hat{\mathbf{h}} = \arg\left(\min_{\mathbf{h}} \|\mathbf{y} - \mathbf{Ah}\|^2\right) \tag{2.12}$$

For white Gaussian noise the solution is given by

$$\hat{\mathbf{h}}_{LS} = \left(\mathbf{A}^H\mathbf{A}\right)^{-1}\mathbf{A}^H\mathbf{y} = \mathbf{h} + \left(\mathbf{A}^H\mathbf{A}\right)^{-1}\mathbf{A}^H\mathbf{n} \tag{2.13}$$

where $(\ )^H$ and $(\ )^{-1}$ denote the Hermitian and inverse matrices, respectively. For the channel impulse response to be identifiable, the auto-correlation matrix $(\mathbf{A}^H\mathbf{A})$ in

(2.13) has to be invertible. Thus, the training sequence matrix **A** has to be of full column rank, such condition is satisfied if $(M_t - L + 1) \geq 2L$ [6], [67], [77].

The computational complexity of the LS estimation depends on the number of transmit antennas and the number of channel coefficients being estimated. To achieve acceptable BER performance, more training symbols are required as the number of transmit antennas increases. A longer training sequence matrix results in a more complex matrix inversion. The computational complexity of finding the inverse of an $M_t \times L$ training sequence matrix is $O(M_t^2 L + L^2 M_t + \min(M_t, L))$ [57], [75].

An alternative to LS estimation that yields comparable performance is the MMSE channel estimation. The optimal MMSE estimate of **h** based on the received signal vector **y** and transmitted training sequence matrix **A** is the conditional mean, which is equivalent to the MAP estimate of **h**. The MAP estimate of **h** is found to be

$$\hat{\mathbf{h}}_{MAP} = \left(\mathbf{A}^H \mathbf{R}_{\mathbf{n}}^{-1} \mathbf{A} + \mathbf{R}_{\mathbf{h}}^{-1}\right)^{-1} \mathbf{A}^H \mathbf{R}_{\mathbf{n}}^{-1} \mathbf{y} \tag{2.14}$$

where $\mathbf{R_n}$ and $\mathbf{R_h}$ are the noise covariance matrix and channel covariance matrix, respectively. Applying the matrix inversion lemma, (2.14) can be expressed as

$$\hat{\mathbf{h}}_{MAP} = \mathbf{R}_{\mathbf{h}} \mathbf{A}^H \left(\mathbf{A} \mathbf{R}_{\mathbf{h}} \mathbf{A}^H + \mathbf{R}_{\mathbf{n}}\right)^{-1} \mathbf{y} \tag{2.15}$$

The total computational cost to find the MAP estimate consists of $(2L^2 M_t + L M_t^2 + LM_t + M_t^2)$ multiplications, $(2L^2 M_t + L M_t^2 + LM_t + 2 M_t^2)$ additions, and $O(M_t^3)$ operations for the inversion of an $M_t \times M_t$ training sequence matrix, which is highly undesirable for practical implementation. Furthermore, such large computational complexity is unsuitable and may be prohibitive in most channels of practical interest [44], [67].

Chapter 3

Performance Comparison between Receive and Transmit Diversity Schemes with Imperfect Multipath Channel Estimation

3.1 Introduction

The simple transmit diversity (STD) scheme utilizing space-time coding (STC) proposed in [16] is now applied in Multiple-Input Multiple-Output (MIMO) communication systems. STD scheme was shown to have the same error performance in non-time-selective channels as maximal ratio combining (MRC) when perfect channel knowledge is available at the receiver [78]. In practice, it is impossible to achieve the perfect channel estimation at the receiver. Hence, an analysis and comparison are presented into the bit error rate (BER) degradations of the STD and

MRC schemes due to imperfect multipath channel estimation at the receiver.

A theoretical approach investigating the BER performance of STD and MRC schemes with perfect and imperfect channel estimation was proposed in [79]. Numerical results showed that the STD scheme has the same BER as the MRC scheme with perfect or imperfect channel evaluation in Gaussian channel. Performance analysis in [80] illustrated that the BER degradation for the STD scheme can be significantly worse than that for an MRC scheme in Rayleigh fading channel and that STD suffers a 3 dB degradation relative to MRC when the channel estimation error is dominant. The performance of the STD scheme with channel estimation error out performs no diversity (ND) system with perfect channel estimation regardless of modulation technique or channel model [81].

BER expressions that are directly dependent on the mean square error of the channel estimator were derived for the STD scheme in [82]. Simulation results showed that the gap in performance caused by the error of the imperfect channel evaluation is within 1 dB compared to the case of perfect channel estimation for the STD scheme over nonselective Rayleigh fading channel. The sensitivity of the STD scheme to channel estimation errors was investigated in [83] to determine the effects on performance as the levels of the channel error and multipath components are varied. A performance comparison of the STD and MRC schemes in multipath channel that leads to frequency selective fading causing intersymbol interference (ISI) in the received signal has not been studied.

An analytical approach is proposed to obtain a closed form expression for the BER in each scheme based on the derived SNR value at the output of the equalizer that is employed to compensate for the ISI in the received signal. Performance curves are obtained to analyze the impact of channel estimation errors and multipath components under the assumption that transmit branches are independent with the same number of multipath components. Results show that the degradation in the STD scheme is almost 8 dB compared to the MRC scheme to achieve a BER of 10^{-4} with 5% channel estimation error.

In this analysis, it is assumed that the channels are independent, remain stationary, and the noise is following Gaussian distribution. Thus, the channel multipath coefficients change so slowly for the channel estimation technique to converge with small estimation errors, which are a percentage of the actual values of the channel coefficients. The obtained BER is the probability of error in Gaussian channels, where each channel coefficient denoted by α_i has scalar quantity. For a Rayleigh fading channel, α_i has a Rayleigh distribution, the obtained BER in Gaussian channel can be viewed as a conditional error probability, where the condition is that α_i are fixed. Averaging the Gaussian BER over the Rayleigh distributed of α_i for all i, yields the probability of error for Rayleigh fading channel. However, in later case it will be difficult to obtain a closed form expression to evaluate the exact BER for a Rayleigh fading channel. To compute the Rayleigh BER expression, which will include multiple integrations (one integration for each channel coefficient) numerical integration methods can be employed.

3.2 No Diversity System

Figure 3.1 shows the baseband representation of a typical no diversity system with one transmit and one receive antennas. At a given time, a signal s_0 is sent from the transmitter. For binary phase-shift keying (BPSK) modulation s_i is real and hence $s_i^* = s_i$, where $*$ is the complex conjugate operation. The transmitted signal is assumed either +1 or −1.

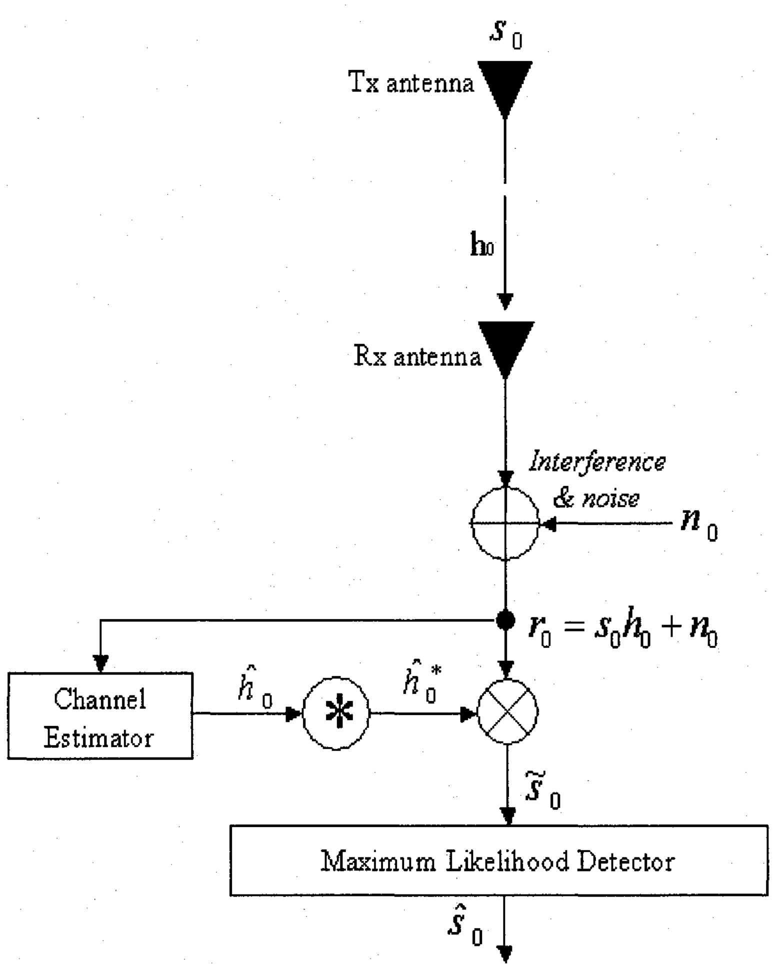

Figure 3.1: System configuration of a classical no diversity scheme.

The multipath channel including the effects of the transmit chain, the propagation channel, and the receive chain is denoted by h_0. The received signal at time t denoted by r_0 is given by

$$r_0 = h_0 s_0 + n_0 \tag{3.1}$$

where n_0 is the additive white Gaussian noise (AWGN) with zero mean and variance σ_n^2. Throughout this analysis, all random variables are complex valued with variance equal to the sum of variances of its real and imaginary components. The channel estimator produces an estimate of the multipath channel coefficients denoted by $\hat{h}_0$. Based on a linear combination of the received signal, the receiver forms the following decision statistic

$$\tilde{s}_0 = \hat{h}_0^* r_0 \tag{3.2}$$

Substituting (3.1) in (3.2), we obtain

$$\tilde{s}_0 = \hat{h}_0^* h_0 s_0 + \hat{h}_0^* n_0 \tag{3.3}$$

In (3.3) the first term is the desired signal, whereas the second term is effectively interference and thermal noise. Based on $\tilde{s}_0$, the maximum likelihood detector selects an estimate of s_0 in order to minimize the BER. If the real part of $\tilde{s}_0$, $\mathrm{Re}(\tilde{s}_0)$ is greater than zero, $s_0 = +1$ is chosen; otherwise $s_0 = -1$ is chosen, for BPSK modulation [78].

3.2.1 System Model

The analysis herein deals with channel estimation errors and is not related to Fig. 3.1, which was shown above for just completion purposes. Due to multipath propagation and time dispersion, the multipath channel is characterized as frequency selective channel that results in ISI. The channel model considered in this analysis has an impulse response that can be expressed as

$$h(t) = \sum_{i=0}^{m} \alpha_i \, \delta(t - iT) \tag{3.4}$$

with a total of (m+1) path components (bins), $\delta(.)$ is the Dirac delta function, and α_i is a complex valued coefficient with zero mean. The sum of all variances of different α_i is normalized to one [57]. Equalization mitigates the ISI created by multipath within the dispersive channel. A linear transversal filter that matches the estimated channel is often used for equalization. Following the estimation of the channel coefficients at the receiver, the equalizer is accordingly adjusted. The baseband impulse response of an equalizer with (n+1) taps is given by

$$\hat{h}(t) = \sum_{j=0}^{n} \beta_j \, \delta(t - jT) \tag{3.5}$$

The required length of the filter (number of tap weights) is a function of how much smearing the channel introduces. For an equalizer with finite length, it is possible to select the tap weights to minimize the ISI. One can use a zero forcing equalizer; the coefficients β_j (resulting from the process of channel estimation) are chosen to force the samples of the combined channel and equalizer impulse response to zero at all but one of the spaced sample points in the tapped delay filter [58].

Without loss of generality, we consider the first term in (3.3), the corresponding equalizer output denoted by $\hat{s}_0(t)$, is found by convolution as follows:

$$\hat{s}_0(t) = s_0 \otimes h(t) \otimes \hat{h}^*(t) = \sum_{i=0}^{m}\sum_{j=0}^{n} s_0\, \alpha_i\, \beta_j^*\, \delta(t - iT - jT) \tag{3.6}$$

where $n \geq m$, and $\otimes$ denotes the convolution operation. The equalizer output at the kth sampling interval can be expressed in the form

$$\begin{aligned} &\hat{s}_0(k) = s_0\, \alpha_0\, \beta_0^*, && k = 0 \\ &\hat{s}_0(k) = \sum_{i=0}^{m}\sum_{j=0}^{n} s_0\, \alpha_i\, \beta_j^*, && k = i + j = 1, 2, \ldots, m + n \end{aligned} \tag{3.7}$$

The desired output of the equalizer corresponds to the k = 0 term while, the other terms corresponds to ISI. An equalizer of sufficient length will use (3.7) to solve simultaneously for the set of (n+1) complex weights β_j. Such that the second term in (3.7) is eliminated for all k [83].

3.2.2 BER Analysis

Due to imperfect channel estimation, the equalizer coefficients can be expressed as

$$\hat{\beta}_j = \beta_j + \varepsilon_j \tag{3.8}$$

where ε_j represents the channel estimation error. The estimation error in each channel coefficient value is modeled as an independent Gaussian random variable with zero mean and variance σ_ε^2, which is equal to the mean square error (MSE) of the channel coefficients estimation process (σ_ε^2 = MSE). Replacing β_j in (3.7) by $\hat{\beta}_j$ from (3.8) and as result of the zero forcing equalizer, we have

$$\begin{aligned} \hat{s}_0 &= s_0\, \alpha_0\, \beta_0^* \\ \gamma_k &= \sum_{i=0}^{m}\sum_{j=0}^{n} s_0\, \alpha_i\, \varepsilon_j^*, \qquad k = i + j = 0,1,2,\ldots,m+n \end{aligned} \tag{3.9}$$

The second term in (3.9) denoted by γ_k is a sum of extra AWGN noise due to channel estimation errors and their reflections on the signal following the equalizer. Since ε_j is independent of α_i and the variance-summed terms of (3.9) are independent. Thus, the sum of the mean square values of γ_k for all values of *k*, is given by

$$\sum_{k=0}^{n+m} E\left[\gamma_k^2\right] = |s_0|^2 (n+1)\left(E\left[\alpha_0^2\right] + E\left[\alpha_1^2\right] + \cdots + E\left[\alpha_m^2\right]\right)\sigma_\varepsilon^2 \tag{3.10}$$

Imposing the normalizing condition $\sum_{i=0}^{m} E\left[\alpha_i^2\right] = 1$ obtains

$$\sum_{k=0}^{n+m} E\left[\gamma_k^2\right] = |s_0|^2 (n+1)\sigma_\varepsilon^2 \tag{3.11}$$

Therefore the intended data symbol term $\hat{s}_0$ has a signal term of peak equal to s_0 and noise term due to the channel estimation error of variance given in (3.11), because $E[\gamma_k] = 0$ for all *k* [83].

The second term $\hat{h}_0^* n_0$ in (3.3) is basically AWGN with zero mean and variance

σ_n^2. Thus, the total noise variance of $\tilde{s}_0$ in (3.3), is given as

$$\sigma_{\tilde{s}_0}^2 = |s_0|^2 (n+1)\sigma_\varepsilon^2 + \sigma_n^2 \tag{3.12}$$

For BPSK signals $|s_0|^2 = |s_1|^2 = E_s$, where E_s is the energy of the signal, (3.12) becomes

$$\sigma_{\tilde{s}_0}^2 = E_s(n+1)\sigma_\varepsilon^2 + \sigma_n^2 \tag{3.13}$$

The mean of $\tilde{s}_0$ in (3.3) is given by

$$E[\tilde{s}_0] = |s_0| \tag{3.14}$$

The required SNR at the output of the equalizer due to $\tilde{s}_0$ can be written as

$$SNR_{\tilde{s}_0} = \frac{\left(\dfrac{E_s}{\sigma_n^2}\right)}{1+(n+1)\left(\dfrac{E_s}{\sigma_n^2}\right)\sigma_\varepsilon^2} \tag{3.15}$$

It is interesting to see that in (3.15) with perfect channel estimation $(\sigma_\varepsilon^2 = 0)$, we obtain $SNR_{\tilde{s}_0} = E_s/\sigma_n^2$ as it should be. The BER of the ND scheme is $Q\left(\sqrt{2SNR_{\tilde{s}_0}}\right)$ for BPSK modulation [22].

3.3 MRC System Performance

Figure 2.2 shows the baseband representation of the classical two-branch MRC scheme with one transmit and two receiving antennas.

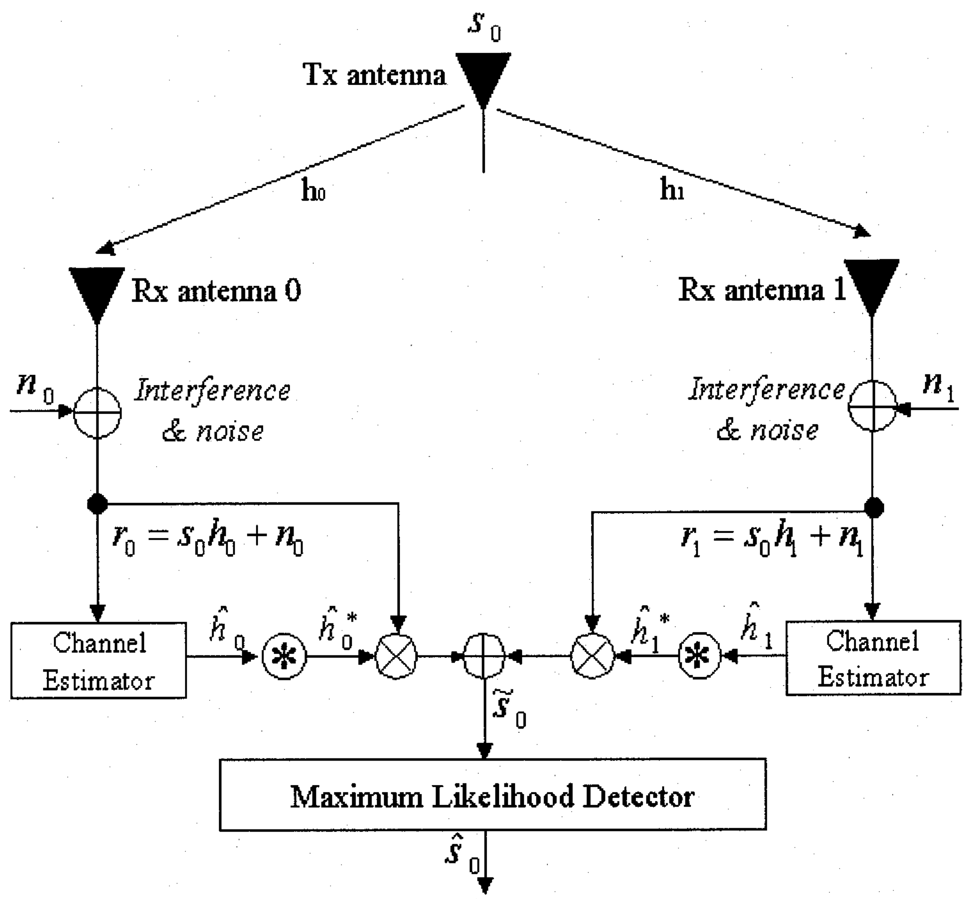

Figure 3.2: System configuration of two branch maximal ratio combining scheme.

Even though two-branch MRC scheme is discussed, the analysis herein can be easily generalized to any arbitrary number of receiving antennas. At a given time, a signal s_0 is sent from the transmitter. The multipath channel coefficients between transmit antenna and receive antennas zero and one that are denoted by h_0 and h_1, respectively. The corresponding received signals at receive antennas zero and one are given as

$$r_0 = h_0 s_0 + n_0$$
$$r_1 = h_1 s_0 + n_1 \tag{3.16}$$

where n_i $(i = 0,1)$ is the AWGN in the ith channel branch with zero mean and variance σ_n^2. The channel estimators produce an estimate of the multipath channels, which are denoted by $\hat{h}_0$ and $\hat{h}_1$. The receiver-combining scheme for two-branch MRC is

$$\tilde{s}_0 = \hat{h}_0^* r_0 + \hat{h}_1^* r_1 \tag{3.17}$$

Substituting the received signals given in (3.16) into (3.17), we obtain

$$\tilde{s}_0 = \left(\hat{h}_0^* h_0 + \hat{h}_1^* h_1\right)(s_0) + \hat{h}_0^* n_0 + \hat{h}_1^* n_1 \tag{3.18}$$

In (3.18) the first two terms describe the combined desired signals from the two receive antennas, whereas the remaining terms are effectively interference and thermal noise. For BPSK modulation, if the real part of $\tilde{s}_i$, Re($\tilde{s}_i$) is greater than zero, $s_i = +1$ is chosen; otherwise $s_i = -1$ is chosen by the maximum likelihood detector in order to minimize BER [79].

3.3.1 Channel Characteristics

The analysis herein deals with channel estimation errors and is not related to Fig. 3.2, which was shown above for just completion purposes. Due to multipath propagation and time dispersion, the channel between transmit and each receive antenna pair is characterized as frequency selective channel that results in ISI. The receiving antennas must be separated far enough to ensure independently fading channels from transmit to each receive antenna. The delay spread per transmit and receive antenna pair is assumed to be the same for all channels [84]. The channel model has an impulse response that can be expressed as

$$h_k(t) = \sum_{i=0}^{m} \alpha_i \, \delta(t - iT), \qquad k = 0,1 \tag{3.19}$$

Each of the kth channel branch has a total of $(m+1)$ path components (bins), $\delta(.)$ is the Dirac delta function, and α_i is a complex valued coefficient with zero mean. The sum of all variances of different α_i is normalized to one [57]. The index k is dropped from the channel model given above, since all channels are assumed to have the same statistical characteristics, i.e. $h(t) = h_k(t)$ for $k = 0,1$.

Following the estimation of the channel coefficients at the receivers, the equalizers are accordingly adjusted. Recall we have few channel estimators not just one as in Fig. 3.2, but occasionally we mention only one for clarity purposes. The equalization process in this case will be similar to that derived for the ND case. This is justified by the fact that each of the two channels has the same mathematical formulation as the channel model for the ND system. The equalizer output samples will consist of a signal term and noise term due to the channel estimation error [58].

3.3.2 SNR Evaluation of MRC Scheme

A close look at (3.18) shows that each of the terms $\hat{h}_0^* h_0$ and $\hat{h}_1^* h_1$ is similar to the first term given in (3.3). Therefore, the combined noise variance for these two terms in

(3.18) can be easily obtained as

$$\sigma_y^2 = 2|s_0|^2 (n+1)\sigma_\varepsilon^2 \tag{3.20}$$

The last two terms $\hat{h}_0^* n_0$ and $\hat{h}_1^* n_1$ in (3.18) are just AWGN each with zero mean and variance σ_n^2. Consequently, the total noise variance of $\tilde{s}_0$ in (3.18), is given as

$$\sigma_{\tilde{s}_0}^2 = 2|s_0|^2 (n+1)\sigma_\varepsilon^2 + 2\sigma_n^2 \tag{3.21}$$

The mean of $\tilde{s}_0$ in (3.18) is given by

$$E[\tilde{s}_0] = 2|s_0| \tag{3.22}$$

For BPSK signals $|s_0|^2 = |s_1|^2 = E_s$, where E_s is the energy of the signal. The required SNR at the output of the equalizer due to $\tilde{s}_0$ in (3.18) can be written as

$$SNR_{\tilde{s}_0} = \frac{2\left(\frac{E_s}{\sigma_n^2}\right)}{1+(n+1)\left(\frac{E_s}{\sigma_n^2}\right)\sigma_\varepsilon^2} \tag{3.23}$$

We notice that in (3.23) with perfect channel estimation $(\sigma_\varepsilon^2 = 0)$, we obtain $SNR_{\tilde{s}_0} = 2E_s/\sigma_n^2$, which is twice that of ND case. The BER for the MRC scheme is $Q\left(\sqrt{2SNR_{\tilde{s}_0}}\right)$, where $SNR_{\tilde{s}_0}$ is given in (3.23) for BPSK modulation [22].

3.4 STD System Analysis

Figure 3.3 shows the baseband representation of the STD scheme with two transmit and one receive antennas. A typical STD system is discussed even though the analysis herein is applicable to an arbitrary number of receive antennas. As in [16], two signals are simultaneously transmitted from two transmit antennas during two consecutive symbol intervals. In the first symbol period at time *t*, signals s_0 and s_1 are transmitted from antennas zero and one, respectively. In the next symbol period at time $t+T$, signals $\left(-s_1^*\right)$ and s_0^* are transmitted from antennas zero and one respectively, where *T* is the duration of a symbol.

The multipath channel coefficients at time *t* between transmit antennas zero and one to the receive antenna are denoted by $h_0(t)$ and $h_1(t)$, respectively. Assuming that the channels remain stationary across at least two consecutive symbols, thus multipath channel coefficients change slowly and look like constants to be estimated (details for such estimators will be presented in Chapters 5 and 6). Then, we can write

$$h_i(t) = h_i(t+T) = h_i, \qquad i = 0,1 \tag{3.24}$$

The received signals at time *t* and time *t* + *T* can be expressed as

$$\begin{aligned} r_0 &= h_0 s_0 + h_1 s_1 + n_0 \\ r_1 &= -h_0 s_1^* + h_1 s_0^* + n_1 \end{aligned} \tag{3.25}$$

where n_i (i = 0,1) is the AWGN in the *i*th channel branch with zero mean and variance σ_n^2.

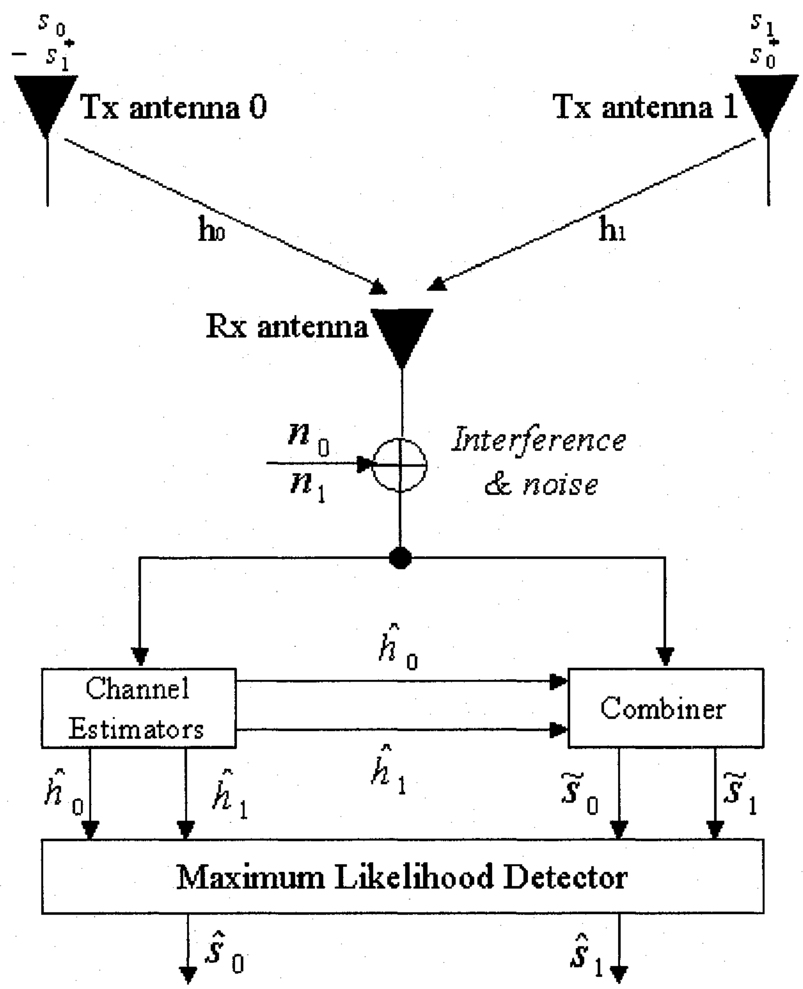

Figure 3.3: System configuration of a simple transmit diversity scheme with one receiving antenna.

As shown in Fig. 3.3, the combiner receives the estimates of the multipath channel coefficients from the channel estimator denoted by $\hat{h}_i (i = 0,1)$. The combiner builds the following two decision statistics

$$\begin{aligned} \tilde{s}_0 &= \hat{h}_0^* r_0 + \hat{h}_1 r_1^* \\ \tilde{s}_1 &= \hat{h}_1^* r_0 - \hat{h}_0 r_1^* \end{aligned} \tag{3.26}$$

Substituting (3.25) in (3.26), we get

$$\widetilde{s}_0 = \left(\hat{h}_0^* h_0 + \hat{h}_1 h_1^*\right)(s_0) + \left(\hat{h}_0^* h_1 - \hat{h}_1 h_0^*\right)(s_1) + \hat{h}_0^* n_0 + \hat{h}_1 n_1^* \tag{3.27}$$

$$\widetilde{s}_1 = \left(\hat{h}_0 h_0^* + \hat{h}_1^* h_1\right)(s_1) + \left(\hat{h}_1^* h_0 - \hat{h}_0 h_1^*\right)(s_0) + \hat{h}_1^* n_0 - \hat{h}_0 n_1^* \tag{3.28}$$

In (3.27) and (3.28) the first two terms describe the combined desired signals, whereas the remaining terms are effectively interference and thermal noise. For BPSK modulation, the maximum likelihood detector computes Re($\widetilde{s}_i$), if it is greater than zero, $s_i = +1$ is chosen; otherwise $s_i = -1$ is chosen to minimize BER [80].

3.4.1 Impact of Channel Estimation Error

The analysis herein deals with channel estimation errors and is not related to Fig. 3.3, which was shown above for just completion purposes. Due to multipath propagation and time dispersion, the channel between each transmit and receive antenna pair is characterized as frequency selective channel that results in ISI. The transmit antennas must be separated far enough to ensure independently fading channels from each transmit antenna to receive antenna. The delay spread per transmit and receive antenna pair is assumed to be the same for all channels. If the transmit antennas are physically co-located at the same station, then this assumption is justified by the fact that the number of multipath components with different delays is dictated by large structures and reflections [85]. The channel model has an impulse response that was given in (3.19).

Since all channels are assumed to have the same statistical characteristics, the impact of the imperfect channel estimation on the equalization process will be similar to that outlined in the MRC case. Thus, the equalizer output samples will again consist of signal terms and noise terms. Without loss of generality, we consider the first two terms in (3.27), which are similar to the first two terms in (3.18). Therefore, their combined noise variance at the equalizer output is given in (3.20). Now for terms like $\hat{h}_0^* h_1 s_1$ in (3.27), notice that $\hat{h}_0$ tries to approximate as much as possible h_0. However, $\hat{h}_0$ has no relation whatsoever to h_1. Developing two such terms in (3.27), e.g.,

$$
\begin{aligned}
z &= \hat{h}_0^* \otimes h_1 \otimes s_1 - \hat{h}_1 \otimes h_0^* \otimes s_1 \\
z &= \sum_{i=0}^{m}\sum_{j=0}^{n} s_1 b_i (a_j^* + \Delta a_j^*)\,\delta(t - iT - jT) - \sum_{i=0}^{m}\sum_{j=0}^{n} s_1 a_i^* (b_j + \Delta b_j)\,\delta(t - iT - jT) \\
z &\geq \sum_{i=0}^{m}\sum_{j=0}^{n} (s_1 b_i \Delta a_j^*)\,\delta(t - iT - jT) - \sum_{i=0}^{m}\sum_{j=0}^{n} (s_1 a_i^* \Delta b_j)\,\delta(t - iT - jT)
\end{aligned}
\tag{3.29}
$$

where a_i, b_j resemble the role of α_i, β_j, while α_i, β_j were related as stated before, a_i, b_j on the other hand have no relation. In (3.29) equality holds if $n = m$, Δa_j, Δb_i are all independent noise terms due to imperfect channel estimation. Clearly $E[z] = 0$, and each term of (3.29) is similar to γ_k in (3.9), so the variance due to imperfect channel estimation is given as

$$\sigma_z^2 \geq 2|s_1|^2 (n+1)\,\sigma_\varepsilon^2 \tag{3.30}$$

The remaining two terms in (3.27), $\hat{h}_0^* n_0$ and $\hat{h}_1 n_1^*$ are basically AWGN each with zero mean and variance σ_n^2. For PSK signals (equal energy constellations),

$|s_0|^2 = |s_1|^2$, the total noise variance of $\tilde{s}_0$ in (3.27) is obtained as

$$\sigma_{\tilde{s}_0}^2 \geq 4|s_0|^2 (n+1)\sigma_\varepsilon^2 + 2\sigma_n^2 \tag{3.31}$$

The equality holds if $n = m$. The mean of $\tilde{s}_0$ in (3.27) is given by

$$E[\tilde{s}_0] = 2|s_0| \tag{3.32}$$

3.4.2 BER Derivation for STD System

The required SNR at the output of the equalizer due to $\tilde{s}_0$ can be written as

$$SNR_{\tilde{s}_0} \leq \frac{2|s_0|^2}{2|s_0|^2 (n+1)\sigma_\varepsilon^2 + \sigma_n^2} \tag{3.33}$$

By symmetry the SNR at the output of the equalizer for the symbol $\tilde{s}_1$ in (3.28) is

$$SNR_{\tilde{s}_1} \leq \frac{2|s_1|^2}{2|s_1|^2 (n+1)\sigma_\varepsilon^2 + \sigma_n^2} \tag{3.34}$$

In (3.33) and (3.34) equality holds if $n = m$. In the STD scheme being considered, two symbols are simultaneously transmitted. By symmetry, the performances for both transmit symbols are the same, so we need only consider the BER for the symbol s_0. Assuming the average symbol energy as E_S and with $n = m$, then the output SNR is

$$SNR_0 = \frac{2\left(\frac{E_S}{\sigma_n^2}\right)}{1 + 2(n+1)\left(\frac{E_S}{\sigma_n^2}\right)\sigma_\varepsilon^2} \tag{3.35}$$

It is interesting to see that in (3.35) with perfect channel estimation $(\sigma_\varepsilon^2 = 0)$, we obtain $SNR_0 = 2E_S/\sigma_n^2$ which is the equivalent MRC conclusion in (3.23). The BER for the STD scheme is $Q\left(\sqrt{2SNR_0}\right)$ for BPSK modulation [22].

In this work, the BER performance for each diversity scheme is evaluated, if the channels are independent, remain stationary, and the noise is following Gaussian distribution. To compute the probability of error for a Rayleigh fading channel, the obtained BER for either MRC or STD schemes is viewed as a conditional error probability, where the condition is that α_i are fixed. Averaging over the Rayleigh distributed α_i, yields the probability of error for Rayleigh fading channel [57].

3.5 Numerical Results and Discussions

For presentation of the numerical results, two parameters are considered in our investigation. The number of channel multipath components (n) and the channel estimation error (σ_ε^2 = MSE). The BER performance curves are computed for the ND, the STD, and the MRC schemes at a certain value of n with various values of MSE. Fig. 3.4 shows the BER performance comparison with both n and MSE set equal to zero, which represents the case of a single path to be estimated between each pair of transmit and receive antennas and perfect channel estimation. In this case, both the STD and MRC systems outperform the ND system.

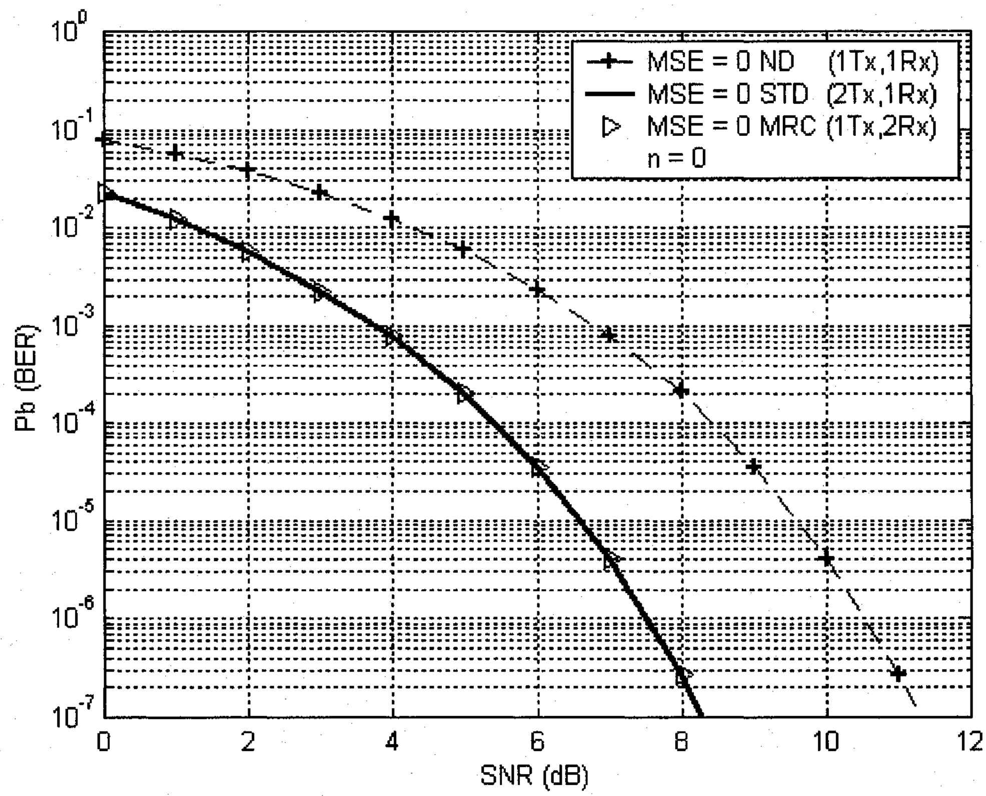

Figure 3.4: The BER performance comparison of the ND scheme with MRC and STD schemes at zero estimation error and $n = 0$.

For $n = 0$, Fig. 3.5 and Fig. 3.6 compare the BER curves for the three schemes with MSE of 5% and 10%, respectively. As the estimation error is increased, the MRC scheme outperforms the STD scheme for SNR values above 2 dB. However, both MRC and STD systems perform better than the ND case. For comparison purpose and at a target BER of 10^{-4}, there is a one dB SNR loss in STD scheme relative to MRC scheme with 5% MSE as shown in Fig. 3.5. However, at the same target BER, the SNR loss increases to about 4 dB in STD compared to MRC with 10% MSE as shown in Fig. 3.6.

The BER performance comparison for the different schemes versus the SNR are illustrated in Fig. 3.7 and Fig. 3.8 at estimation error of 5% and 10% respectively, with n set equal to 2. Note that the BER curves for the case of perfect channel estimation (MSE equal to zero) are those shown in Fig. 3.4 regardless of the value of n. By increasing the value of n, the performance of the STD scheme becomes worse than that of the MRC scheme for all values of SNR. The degradation in STD scheme is almost 8 dB relative to MRC scheme at an estimation error of 5% to achieve a BER of 10^{-4} as seen in Fig. 3.7. In Fig. 3.8, it is found that the deterioration in the performance of the STD scheme increases rapidly as SNR increases, such that the BER curve of the STD scheme approaches that of the ND system.

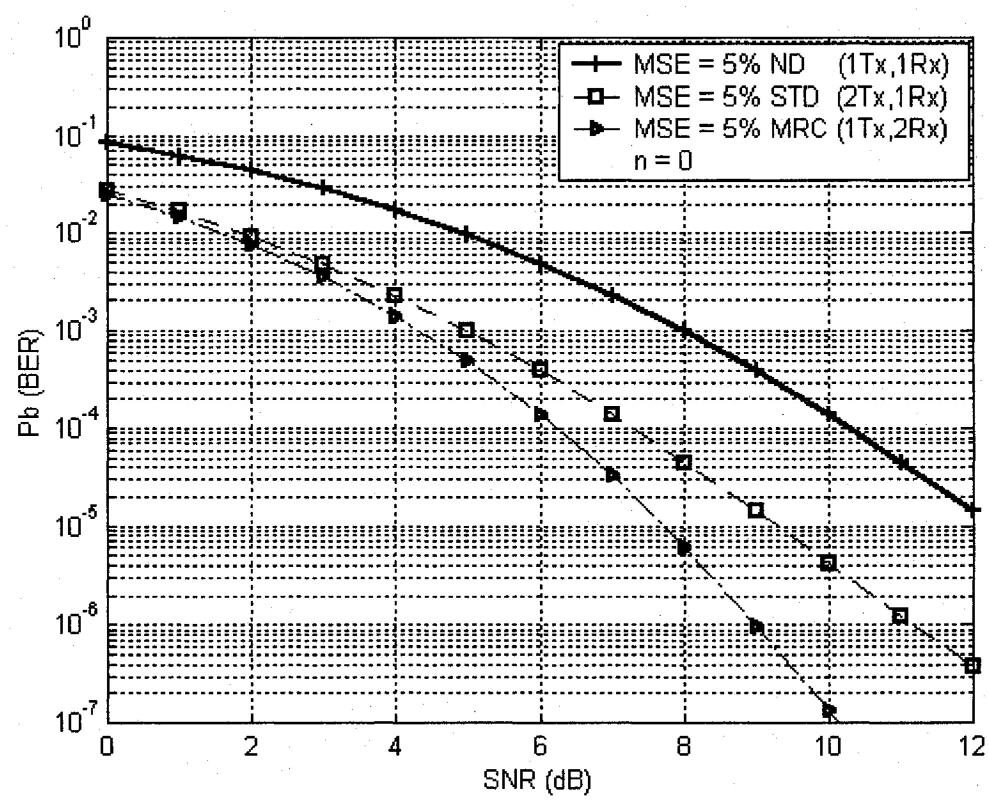

Figure 3.5: The BER performance comparison of the ND scheme with MRC and STD schemes at 5% estimation error and $n = 0$.

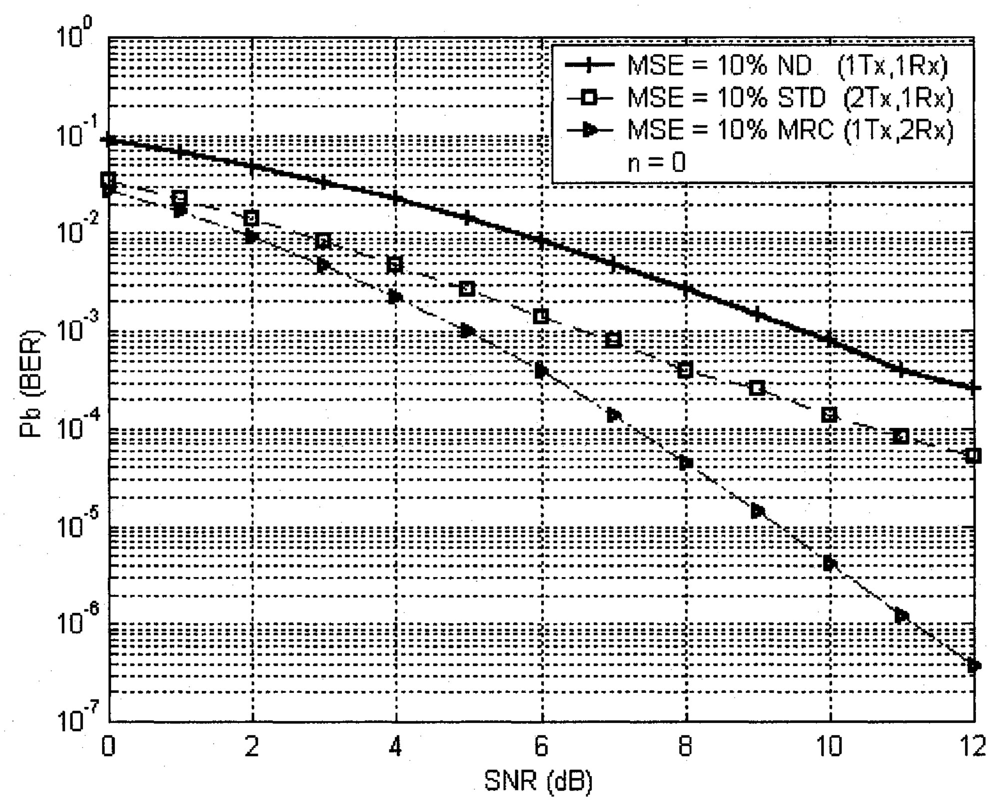

Figure 3.6: The BER performance comparison of the ND scheme with MRC and STD schemes at 10% estimation error and $n = 0$.

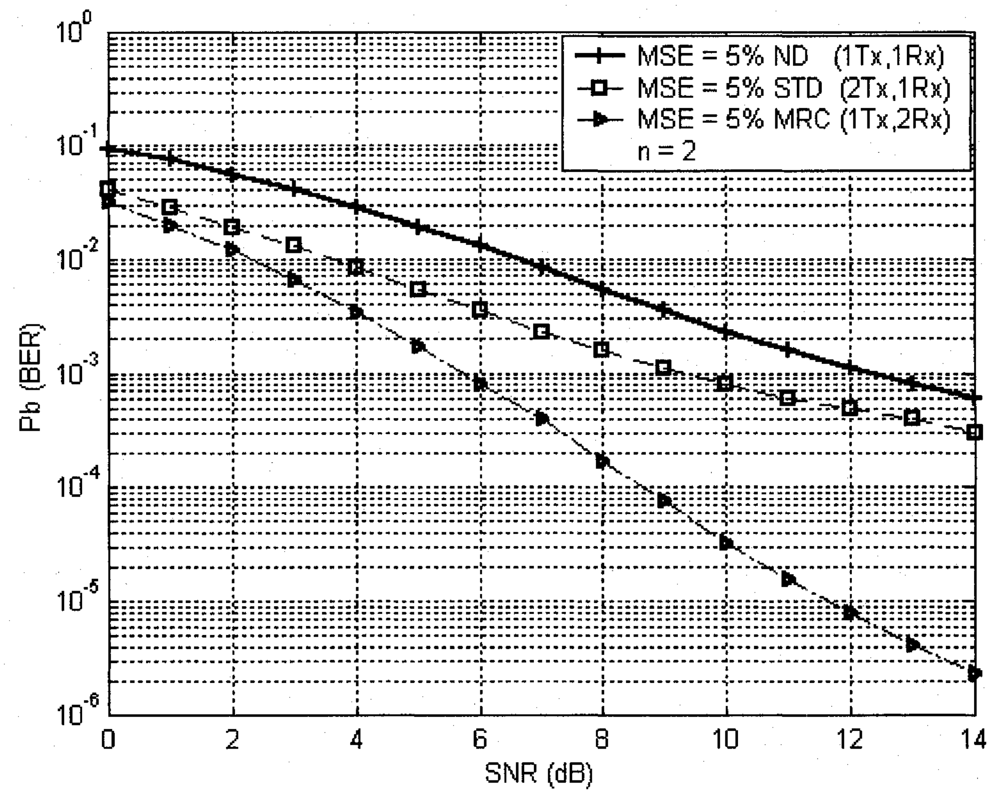

Figure 3.7: The BER performance comparison of the ND scheme with MRC and STD schemes at 5% estimation error and $n = 2$.

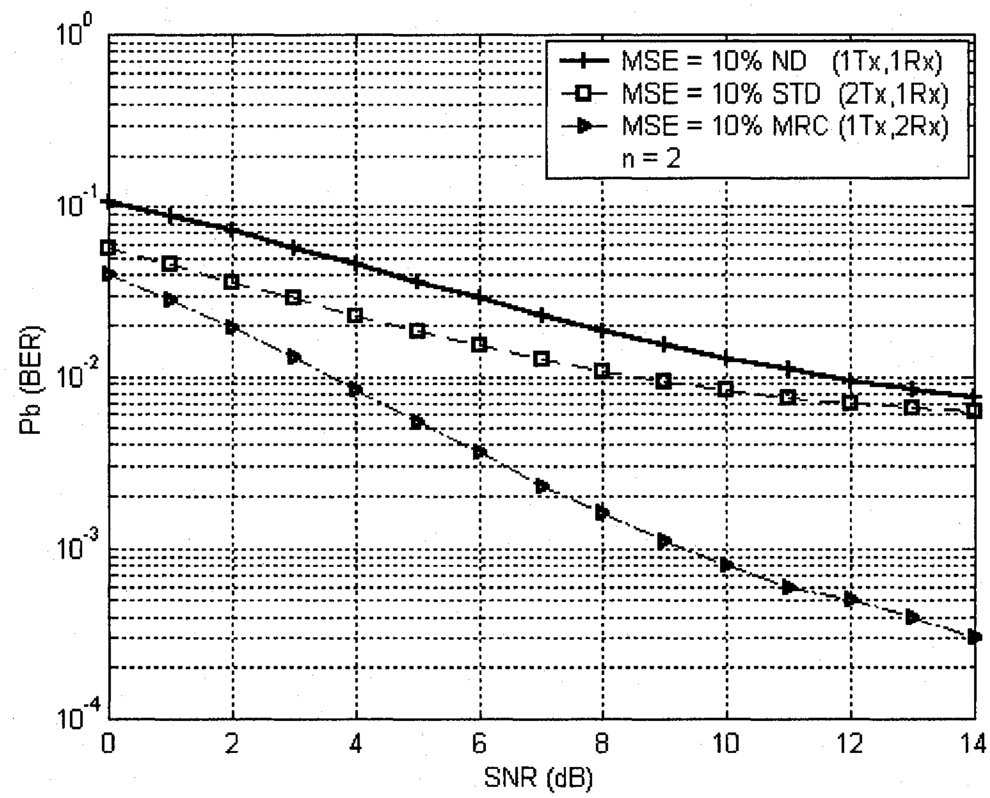

Figure 3.8: The BER performance comparison of the ND scheme with MRC and STD schemes at 10% estimation error and $n = 2$.

For $n = 4$, The BER performance curves of the three schemes are shown in Fig. 3.9 and Fig. 3.10 with channel estimation error of 5% and 10%, respectively. The STD scheme exhibits the same degradation behavior in performance compared to the performance of the MRC scheme. At 5% estimation error the loss in SNR of the STD performance exceeds 8 dB relative to that of the MRC scheme to achieve a BER of only 3×10^{-3}. Amazingly, the deterioration in the STD performance outweighs the benefits achieved over the ND case when the SNR exceeds 10 dB with 10% estimation error. In addition, there is an irreducible error floor in the STD performance when SNR is above 10 dB.

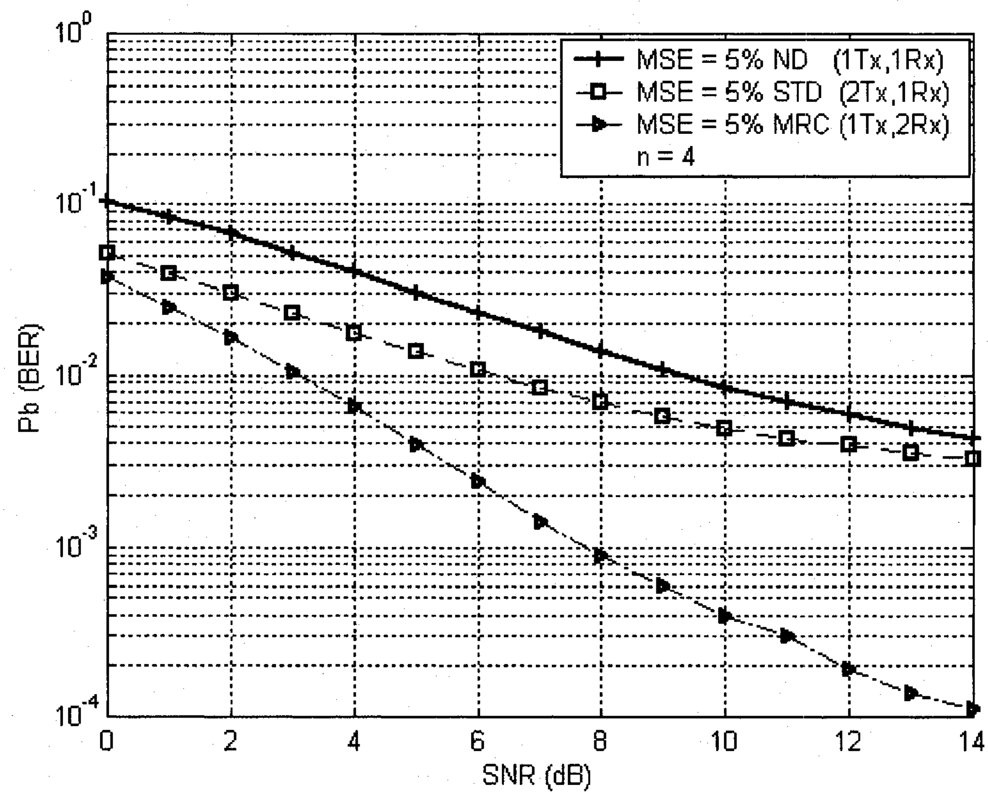

Figure 3.9: The BER performance comparison of the ND scheme with MRC and STD schemes at 5% estimation error and $n = 4$.

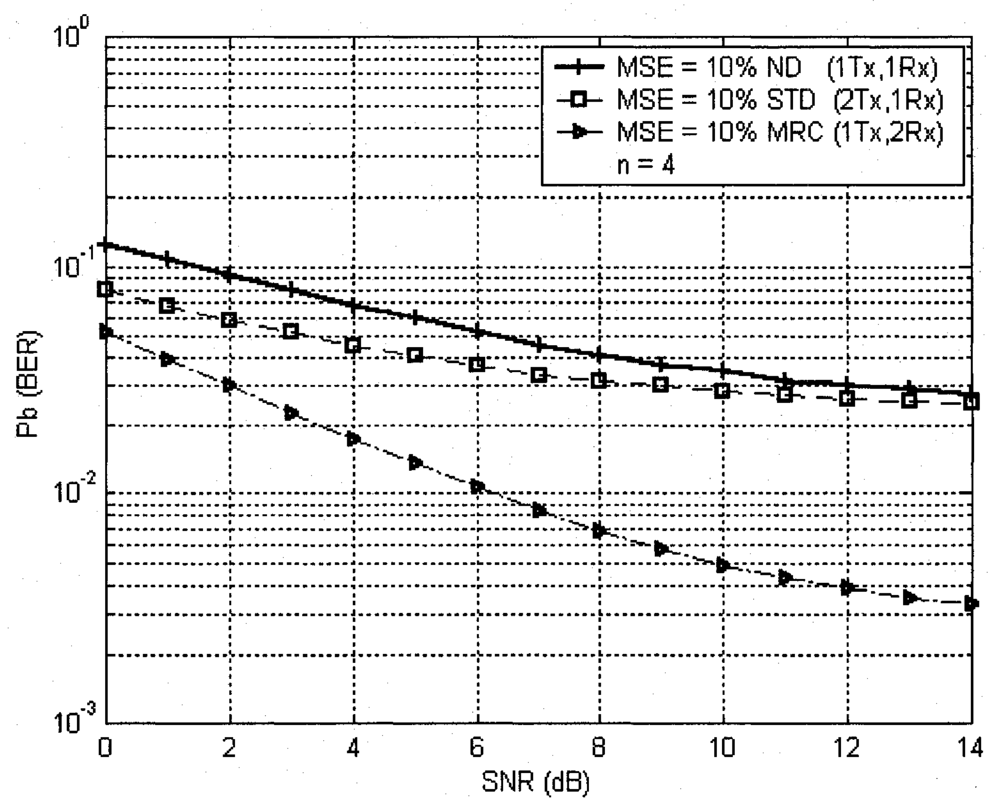

Figure 3.10: The BER performance comparison of the ND scheme with MRC and STD schemes at 10% estimation error and $n = 4$.

3.6 Conclusions

An analysis and comparison of the error performance of the STD and MRC diversity schemes as well as the ND scheme under estimation inaccuracies of multipath channels were presented. A theoretical approach was proposed to obtain the SNR at the output of the equalizer employed to compensate for the ISI in the received signals for the three schemes. The derived expressions showed that the SNR is inversely proportional to two parameters, the number of channel multipath components (n) and the multipath channel estimation error (σ_{ε}^2 = MSE).

To analyze the impact of various estimation errors, the BER performance curves for the three schemes were plotted, if channels are independent and exhibit the same statistical characteristics. Performance curves showed that the STD scheme is significantly more susceptible to errors than the MRC scheme. To achieve a BER of 10^{-4}, the degradation in the STD is almost 8 dB compared to MRC with estimation error of 5% at n = 2. The deterioration in the STD performance increases rapidly relative to the MRC performance with n = 4 and 5% estimation error.

At an estimation error of 10%, there is an irreducible error floor exhibited in the BER performance curve of the STD scheme, such that the deterioration in its performance outweighs the benefits achieved over the ND scheme. Thus, the practical implementation of the STD system should be carefully considered in channels with large number of channel multipath components and channel estimation error exceeding 5% at the receiver.

Chapter 4

Effects of Multipath Channel Estimation Error on Space-Time Coding Performance

4.1 Introduction

In this chapter, the effect of imperfect channel estimation on the bit error rate (BER) of Multiple-Input Multiple-Output (MIMO) communication system utilizing Space-time coding (STC) is investigated. A simple transmit diversity (STD) scheme using two transmit and multiple receive antennas was first proposed in [16]. It was shown to yield the same diversity order as a maximal ratio combining (MRC) scheme under the assumption that perfect channel estimation is available at the receiver.

In reality, reducing the channel estimation error entails a multitude of techniques. Many of those would require complex processing and a large number of iterations to reduce the channel estimation error. This may become prohibitive for real time applications such as voice and video in fast fading situations. Effects of channel estimation errors on the BER were first discussed in [32]. In the simulation, quadrature phase-shift modulation (QPSK) modulation on slow Rayleigh fading channels with two transmit and two receive antennas were employed.

It is assumed that the channel can be described by a slow Rayleigh fading model with constant coefficients over a frame of 130 symbols. The simulation results showed that the STC performance degraded by at most 1 dB compared to the case of ideal channel state information. However, as the numbers of transmit antennas increases, the sensitivity of the system to channel estimation error increases [32]. An analysis and comparison of the performance degradations of the STD scheme and MRC receiver due to errors in estimating the channel parameters were presented in [78]. Performance curves showed that the STD scheme is significantly more susceptible to errors in the channel estimates than that of an MRC receiver.

It was found that an STD system with one receive antenna has a 3 dB performance degradation compared with MRC when SNR and channel estimation error-to-signal ratio (ESR) are large [80]. An ESR of –15 dB results in an SNR loss of 7 dB compared with the perfect channel estimation case at a target BER of 10^{-3} [80]. Theoretical analysis of the error performance of STD scheme for Rayleigh and Rician fading using QPSK was carried out in [81]. The sensitivity of the system to channel estimation error was investigated by analyzing the effects of reducing the error caused by noise when estimating the channel. It was found that the deterioration of performance in STD scheme is about 5 dB due to channel estimation error [81].

In this analysis, a multipath channel that lead to frequency selective fading is considered. Frequency selective fading caused by multipath time delay spread causes intersymbol interference (ISI), which leads to an increased probability of bit error at the receiver. The receiver employs an equalizer to compensate or reduce the ISI in the received signal [57]. No particular channel estimation technique is discussed herein and the results apply to any channel estimation technique. Curves for the BER are obtained based on the derived SNR at the output of the equalizer. Results show that the degradation of performance in a two transmit and two receive antennas scheme exceeds 8 dB when the channel estimation error is 5% to achieve a BER of 10^{-4}.

4.2 STD Scheme with Two Receivers

A typical STD system with two receive antennas is discussed even though the analysis herein is applicable to an arbitrary number of receive antennas. As in [16], two signals are simultaneously transmitted from two transmit antennas during two consecutive symbol intervals. Fig. 4.1 shows the block diagram of the two-branch Alamouti scheme [16] with two receivers. The signals transmitted from antennas zero and one are denoted by s_0 and s_1, respectively. During the next symbol period signals $\left(-s_1^*\right)$ and s_0^* are transmitted from antennas zero and one respectively, where $*$ is the complex conjugate operation.

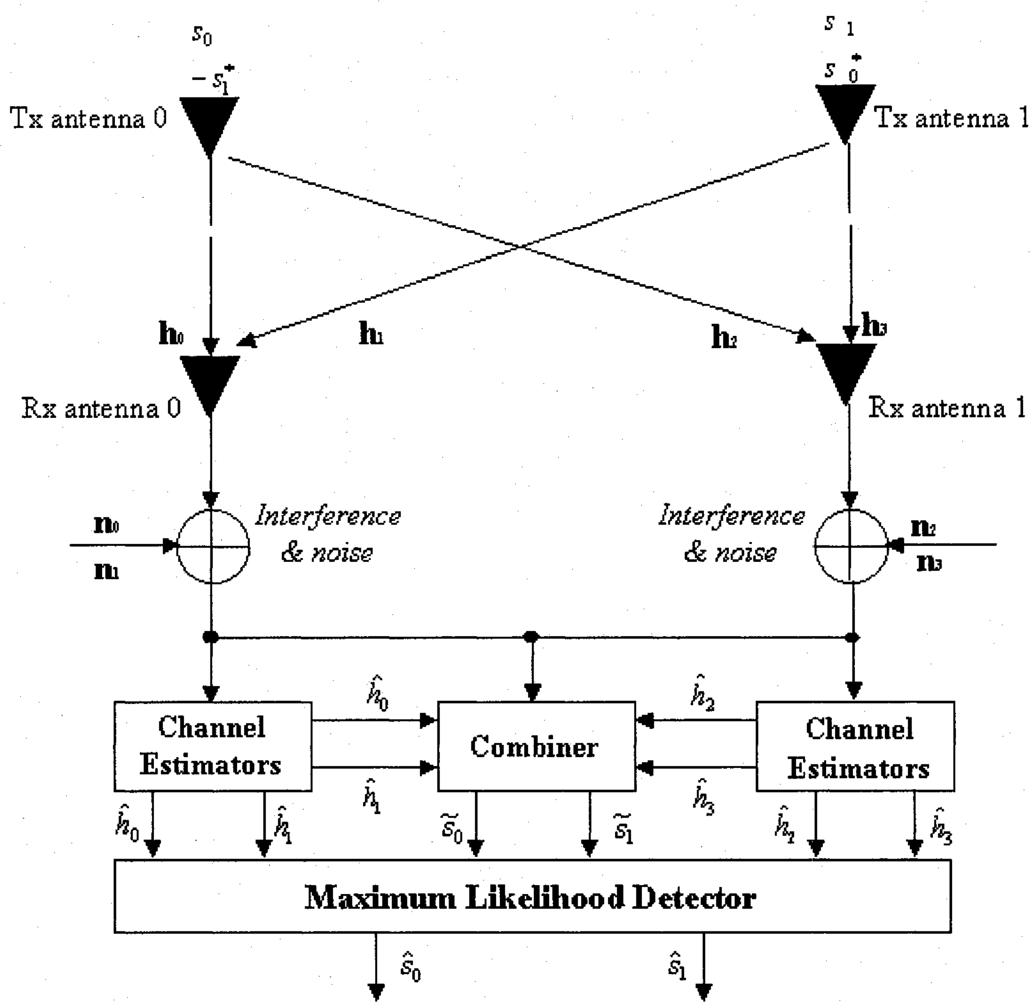

Figure 4.1: Alamouti two-branch transmit diversity scheme with two receivers.

The multipath channel coefficients between transmit and receive antennas are denoted by h_i (i = 0,1,2,3), where h_0 and h_1 represent the channel at time t between transmit antennas zero and one to receive antenna zero, respectively. Similarly, h_2 and h_3 represent the channels to receive antenna one as illustrated in Fig. 4.1. Assuming that the channels remain stationary across at least two consecutive symbols, thus

$h_i(t) = h_i(t+T) = h_i$, where T is the duration of a symbol. Many wireless channels typically have a coherence time larger than one packet time, which is the time to transmit a large number of consecutive symbols [57], [58]. Thus multipath channel coefficients change slowly and look like constants to be estimated. The received signals at time t and time $t + T$ at receive antenna zero are denoted by r_0 and r_1, respectively. The received signals at time t and time $t + T$ at receive antenna one are denoted by r_2 and r_3, respectively. Thus, the received signals can be expressed as:

$$\begin{aligned} r_0 &= h_0 s_0 + h_1 s_1 + n_0 \\ r_1 &= -h_0 s_1^* + h_1 s_0^* + n_1 \\ r_2 &= h_2 s_0 + h_3 s_1 + n_2 \\ r_3 &= -h_2 s_1^* + h_3 s_0^* + n_3 \end{aligned} \tag{4.1}$$

where n_i (i = 0,1,2,3) is the AWGN in the ith channel branch with zero mean and variance σ_n^2. Throughout this analysis, all random variables are complex valued with variance equal to the sum of variances of its real and imaginary components. Note that for BPSK modulation s_i is real and hence $s_i^* = s_i$. The transmitted signal is assumed either +1 or −1. The combiner [16] receives the estimates of the multipath channel coefficients from the channel estimator denoted by $\hat{h}_i (i = 0,1,2,3)$. Two decision statistics based on the linear combination of the received signals are formed. The decision statistics, denoted by $\tilde{s}_0$ and $\tilde{s}_1$ are given by

$$\begin{aligned} \tilde{s}_0 &= \hat{h}_0^* r_0 + \hat{h}_1 r_1^* + \hat{h}_2^* r_2 + \hat{h}_3 r_3^* \\ \tilde{s}_1 &= \hat{h}_1^* r_0 - \hat{h}_0 r_1^* + \hat{h}_3^* r_2 - \hat{h}_2 r_3^* \end{aligned} \tag{4.2}$$

Substituting (4.1) in (4.2), we get

$$\begin{aligned}\tilde{s}_0 &= \left(\hat{h}_0^* h_0 + \hat{h}_1 h_1^* + \hat{h}_2^* h_2 + \hat{h}_3 h_3^*\right)(s_0) + \left(\hat{h}_0^* h_1 - \hat{h}_1 h_0^* + \hat{h}_2^* h_3 - \hat{h}_3 h_2^*\right)(s_1) \\ &\quad + \hat{h}_0^* n_0 + \hat{h}_1 n_1^* + \hat{h}_2^* n_2 + \hat{h}_3 n_3^*\end{aligned} \tag{4.3}$$

$$\begin{aligned}\tilde{s}_1 &= \left(\hat{h}_0 h_0^* + \hat{h}_1^* h_1 + \hat{h}_2 h_2^* + \hat{h}_3^* h_3\right)(s_1) + \left(\hat{h}_1^* h_0 - \hat{h}_0 h_1^* + \hat{h}_3^* h_2 - \hat{h}_2 h_3^*\right)(s_0) \\ &\quad + \hat{h}_1^* n_0 - \hat{h}_0 n_1^* + \hat{h}_3^* n_2 - \hat{h}_2 n_3^*\end{aligned} \tag{4.4}$$

In (4.3) and (4.4) the first four terms describe the combined desired signals from the two receive antennas, whereas the remaining terms are effectively interference and thermal noise. If the real part of $\tilde{s}_i$, Re($\tilde{s}_i$) is greater than zero, $s_i = +1$ is chosen; otherwise $s_i = -1$ is chosen by the maximum likelihood detector to minimize BER, for BPSK modulation [79].

4.2.1 Multipath Channel Model and Equalization Filter

The analysis herein deals with channel estimation errors and is not related to Fig. 4.1, which was shown above for just completion purposes. Due to multipath propagation and time dispersion, the channel between each transmit and receive antenna pair is characterized as frequency selective channel that results in ISI. The antennas on both ends are separated far enough to ensure independently fading channels from each transmit to each receive antenna. The delay spread per transmit and receive antenna pair is assumed to be the same for all channels. If the transmit antennas are physically co-located at the same station, then this assumption is justified by the fact that the number of multipath components with different delays is dictated by large structures and reflections [85].

The channel model has an impulse response that can be expressed as

$$h_k(t) = \sum_{i=0}^{m} \alpha_i \, \delta(t - iT), \qquad k = 0,1,2,3 \tag{4.5}$$

Each of the k channels has a total of $(m+1)$ path components (bins), $\delta(.)$ is the Dirac delta function, and α_i is the complex valued coefficient with zero mean. The sum of all variances of different α_i is normalized to one [57]. Since all channels are assumed to have the same statistical characteristics, we consider a channel $h(t) = h_k(t)$ for k = 0,1 to illustrate the equalization process.

Following the estimation of the channel coefficients at the receiver (a process that should take a small period of time), the equalizers are accordingly adjusted. Equalization cancels the ISI created by multipath within the dispersive channels. A linear transversal filter that matches the estimated channel is often used for equalization. The baseband impulse response of an equalizer with $(n+1)$ taps is given by

$$\hat{h}(t) = \sum_{j=0}^{n} \beta_j \, \delta(t - jT) \tag{4.6}$$

The required length of the filter (number of tap weights) is a function of how much smearing the channel introduces. For an equalizer with finite length it is possible to select the tap weights to minimize the ISI. One can use a zero forcing equalizer; the coefficients β_j (resulting from the process of channel estimation) are chosen to force the samples of the combined channel and equalizer impulse response to zero at all but one of the spaced sample points in the tapped delay filter [22].

Without loss of generality, we consider one of the first four terms in (4.3), the corresponding equalizer output denoted by $\hat{s}_0(t)$, is found by convolution as follow:

$$\hat{s}_0(t) = s_0 \otimes h(t) \otimes \hat{h}^*(t) = \sum_{i=0}^{m}\sum_{j=0}^{n} s_0\, \alpha_i\, \beta_j^*\, \delta(t - iT - jT) \tag{4.7}$$

where $n \geq m$, and $\otimes$ denotes the convolution operation. Each equalizer output at the kth sampling interval can be expressed in the form

$$\begin{aligned} \hat{s}_0(k) &= s_0\, \alpha_0\, \beta_0^*, & k &= 0 \\ \hat{s}_0(k) &= \sum_{i=0}^{m}\sum_{j=0}^{n} s_0\, \alpha_i\, \beta_j^*, & k &= i + j = 1, 2, \ldots, m+n \end{aligned} \tag{4.8}$$

The desired output of the equalizer corresponds to the k = 0 term while, the other terms corresponds to ISI. The ISI terms may be eliminated by a zero forcing equalizer of sufficient length [22].

Equation (4.8) is used to solve simultaneously the set of (n +1) complex weights β_j such that for k = 0, $\mathrm{Re}(s_0\alpha_0\beta_0^*) = s_0$ and $\mathrm{Im}(s_0\alpha_0\beta_0^*) = 0$, which leads to the following:

$$\begin{aligned} \mathrm{Re}(\alpha_0)\,\mathrm{Re}(\beta_0) + \mathrm{Im}(\alpha_0)\,\mathrm{Im}(\beta_0) &= 1 \\ \mathrm{Im}(\alpha_0)\,\mathrm{Re}(\beta_0) - \mathrm{Re}(\alpha_0)\,\mathrm{Im}(\beta_0) &= 0 \end{aligned} \tag{4.9}$$

Knowing $\mathrm{Re}(\alpha_0)$ and $\mathrm{Im}(\alpha_0)$ enables one to compute $\mathrm{Re}(\beta_0)$ and $\mathrm{Im}(\beta_0)$ from (4.9). Similarly, for $k \geq 1$, we have

$$\mathrm{Re}\left(\sum_{i=0}^{m}\sum_{j=0}^{n} s_0 \alpha_i \beta_j^*\right) = 0$$
$$\mathrm{Im}\left(\sum_{i=0}^{m}\sum_{j=0}^{n} s_0 \alpha_i \beta_j^*\right) = 0 \qquad k = i + j = 1, 2, \ldots, m+n \qquad (4.10)$$

Each equalizer will solve the above real and imaginary equations to find the components of the complex β_j, such that the second term in (4.8) is eliminated for all k.

4.2.2 Impact of Imperfect Channel Estimation

Channel estimation is needed to adjust the tap weights of the equalizer to eliminate ISI. Due to imperfect channel estimation, the equalizer coefficients can be expressed as

$$\hat{\beta}_j = \beta_j + \varepsilon_j \qquad (4.11)$$

where ε_j represents the channel estimation error. The estimation error in each channel coefficient value is modeled as an independent Gaussian random variable with zero mean and variance σ_ε^2, which is equal to the mean square error (MSE) of the channel coefficients estimation process ($\sigma_\varepsilon^2 = \mathrm{MSE}$).

Replacing β_j in (4.8) by $\hat{\beta}_j$ from (4.11) and as result of the zero forcing equalizer, which imposes the condition in (4.10), we have

$$\hat{s}_0 = s_0\, \alpha_0\, \beta_0^*$$
$$\gamma_k = \sum_{i=0}^{m}\sum_{j=0}^{n} s_0\, \alpha_i\, \varepsilon_j^*, \qquad k = i + j = 0,1,2,...,m+n \tag{4.12}$$

The first term in (4.12) is the signal term, which is typically s_0 due to the action of the equalizer that follows the channel estimation. The second term denoted by γ_k is a sum of extra AWGN noise due to channel estimation errors and their reflections on the signal following the equalizer. Note that ε_j is independent of α_i and the variance-summed terms of (4.12) are independent. Thus, the mean square value of the *k*th term of γ_k, is given by

$$E[\gamma_0^2] = |s_0|^2 \sigma_\varepsilon^2 E[\alpha_0^2]$$
$$E[\gamma_1^2] = |s_0|^2 \sigma_\varepsilon^2 (E[\alpha_0^2] + E[\alpha_1^2])$$
$$\vdots$$
$$E[\gamma_{n+m}^2] = |s_0|^2 \sigma_\varepsilon^2 E[\alpha_m^2]$$

Noting the independence of all ε_j, the total reflection of the ISI terms of one data symbol on neighboring data symbols or vice versa, the ISI encountered by the symbol s_0 is the sum of the $E[\gamma_k^2]$ for all values of *k*, i.e.

$$\sum_{k=0}^{n+m} E[\gamma_k^2] = |s_0|^2 (n+1)\left(E[\alpha_0^2] + E[\alpha_1^2] + \cdots + E[\alpha_m^2]\right)\sigma_\varepsilon^2 \tag{4.13}$$

Imposing the normalizing condition $\sum_{i=0}^{m} E[\alpha_i^2] = 1$ obtains

$$\sum_{k=0}^{n+m} E\left[\gamma_k^2\right] = \left|s_0\right|^2 (n+1)\sigma_\varepsilon^2 \tag{4.14}$$

Therefore the intended data symbol term $\hat{s}_0$ has a signal term of peak equal to s_0 and noise term due to the channel estimation error of variance given in (4.14), because $E[\gamma_k] = 0$ for all *k*. The analysis of the remaining terms of (4.3) is similar and yields the same signal peak and noise variance as that of $\hat{s}_0$. Thus their combined variance is given as

$$\sigma_y^2 = 4\left|s_0\right|^2 (n+1)\sigma_\varepsilon^2 \tag{4.15}$$

4.3 BER Performance and Numerical Analysis

Now for terms like $\hat{h}_0^* h_1 s_1$ in (4.3), notice that $\hat{h}_0$ tries to approximate as much as possible h_0 (there is only a small noisy difference between them). However, $\hat{h}_0$ has no relation whatsoever to h_1. Developing two such terms in (4.3), e.g.,

$$\begin{aligned}
z &= \hat{h}_0^* \otimes h_1 \otimes s_1 - \hat{h}_1 \otimes h_0^* \otimes s_1 \\
z &= \sum_{i=0}^{m}\sum_{j=0}^{n} s_1 b_i (a_j^* + \Delta a_j^*)\delta(t - iT - jT) - \sum_{i=0}^{m}\sum_{j=0}^{n} s_1 a_i^* (b_j + \Delta b_j)\delta(t - iT - jT) \\
z &\geq \sum_{i=0}^{m}\sum_{j=0}^{n} (s_1 b_i \Delta a_j^*)\delta(t - iT - jT) - \sum_{i=0}^{m}\sum_{j=0}^{n} (s_1 a_i^* \Delta b_j)\delta(t - iT - jT)
\end{aligned} \tag{4.16}$$

where a_i, b_j resemble the role of α_i, β_j, while α_i, β_j were related as stated before, a_i, b_j on the other hand have no relation. In (4.16) equality holds if $n = m$, Δa_j, Δb_i

are all independent noise terms due to imperfect channel estimation. Clearly $E[z]=0$, and each term of (4.16) is similar to γ_k in (4.12), so the variance due to imperfect channel estimation is given as

$$\sigma_z^2 \geq 2|s_1|^2(n+1)\sigma_\varepsilon^2 \tag{4.17}$$

The same procedure is followed for the other two similar terms given in (4.3), i.e.

$$g = \hat{h}_2^* \otimes h_3 \otimes s_1 - \hat{h}_3 \otimes h_2^* \otimes s_1 \tag{4.18}$$

The noise variance due to estimation error is expressed as

$$\sigma_g^2 \geq 2|s_1|^2(n+1)\sigma_\varepsilon^2 \tag{4.19}$$

The remaining four terms in (4.3), $\hat{h}_0^* n_0$, $\hat{h}_1 n_1^*$, $\hat{h}_2^* n_2$, and $\hat{h}_3 n_3^*$ are basically AWGN each with zero mean and variance σ_n^2.

Therefore the total noise variance of $\tilde{s}_0$ given in (4.3) is obtained as

$$\sigma_{\tilde{s}_0}^2 \geq 4|s_0|^2(n+1)\sigma_\varepsilon^2 + 4|s_1|^2(n+1)\sigma_\varepsilon^2 + 4\sigma_n^2 \tag{4.20}$$

For PSK signals (equal energy constellations) $|s_0|^2 = |s_1|^2$, then (4.20) becomes

$$\sigma_{\tilde{s}_0}^2 \geq 4|s_0|^2(n+1)\sigma_\varepsilon^2 + 4\sigma_n^2 \tag{4.21}$$

The equality holds if $n = m$.

4.3.1 BER Evaluation with Imperfect Channel Estimation

The mean of $\tilde{s}_0$ in (4.3) is given by

$$E[\tilde{s}_0] = 4|s_0| \tag{4.22}$$

The required SNR at the output of the equalizer due to $\tilde{s}_0$ can be written as

$$SNR_{\tilde{s}_0} \le \frac{4|s_0|^2}{2|s_0|^2(n+1)\sigma_\varepsilon^2 + \sigma_n^2} \tag{4.23}$$

By symmetry the SNR at the output of the equalizer for the symbol $\tilde{s}_1$ in (4.4) is

$$SNR_{\tilde{s}_1} \le \frac{4|s_1|^2}{2|s_1|^2(n+1)\sigma_\varepsilon^2 + \sigma_n^2} \tag{4.24}$$

In (4.23) and (4.24), equality holds if $n = m$.

In the STD scheme being considered, two symbols are simultaneously transmitted. By symmetry, the performances for the both transmit symbols are the same, so we need only to consider the BER for the symbol s_0. Assuming the average symbol energy as E_S and with $n = m$, then the corresponding output SNR is

$$SNR_0 = \frac{4\left(\frac{E_S}{\sigma_n^2}\right)}{1 + 2(n+1)\left(\frac{E_S}{\sigma_n^2}\right)\sigma_\varepsilon^2} \tag{4.25}$$

It is interesting to see that in (4.25) with perfect channel estimation $(\sigma_\varepsilon^2 = 0)$, we obtain $SNR_0 = 4E_S/\sigma_n^2$ which is the equivalent MRC conclusion in [16]. The BER is $Q\left(\sqrt{2SNR_0}\right)$ for BPSK modulation [22].

In this analysis, it was assumed that multipath components change so slowly for the channel estimation technique to converge with small error σ_ε^2. Even with this assumption, slow Rayleigh fading channel coefficients α_i may occur. To compute SNR_0 in the last-mentioned case, one could assume an additional Rayleigh distribution for each α_i that reflects the effect of channel time fading. The sum of power of α_i in (4.13) will now be conditional and the values SNR_0 and BER will be conditional on α_i values. Averaging over the Rayleigh distributed α_i then yields the average BER [79].

For fast fading channels, even the last analysis may not apply; almost all channel estimation techniques will not converge properly before the fast channel multipath components change again. The result in (4.25) could still apply in such cases, however a much larger σ_ε^2 has to be assumed.

4.3.2 Numerical Results and Discussions

It is assumed that each transmit antenna radiates half the energy in order to ensure the same total radiated power as that of the classical no diversity (ND) case with a single transmit and single receive antennas. With this power normalization, if we define the SNR for the ND case as $\Gamma = 1/\sigma_n^2$, then the SNR_0 for the STD scheme with perfect channel estimation is 2Γ. The benefit of using transmit diversity is outweighed by the errors resulting from estimating the channels, when Γ exceeds a certain threshold SNR, denoted by Γ_t.

Figure 4.2 shows the BER curves for various values of σ_ε^2 = MSE with $n = 0$, which represents the case of a single path to be estimated between each pair of transmit and receive antennas. At a target BER of 10^{-5}, the loss in SNR compared to perfect channel estimation is about 0.7 dB, 1.6 dB, and 2.8 dB for σ_ε^2 of 2%, 5% and 10%, respectively. Note that Γ_t corresponds to higher values of 17 dB, 13 dB, and 10 dB for the three different values σ_ε^2.

For $n = 1$, the BER curves of the STD scheme against SNR (Γ) for various σ_ε^2 are illustrated in Fig. 4.3. For comparison purposes, the performance curves for ND case (1 Tx, 1 Rx) and an STD scheme with perfect channel estimation are included. The Γ_t is about 14 dB, 10 dB, and 7 dB for σ_ε^2 = 2%, 5%, 10%, respectively. At BER of 10^{-4} there is about 1 dB, 2.3 dB, and 3.8 dB degradation with the three different values of σ_ε^2 relative to perfect channel estimation case.

At low SNR from 0 dB to 7 dB, the STD scheme with imperfect channel estimation outperforms the ND system. Amazingly, the performance of the STD scheme is worse than that of the ND system with perfect channel estimation at SNR values above Γ_t.

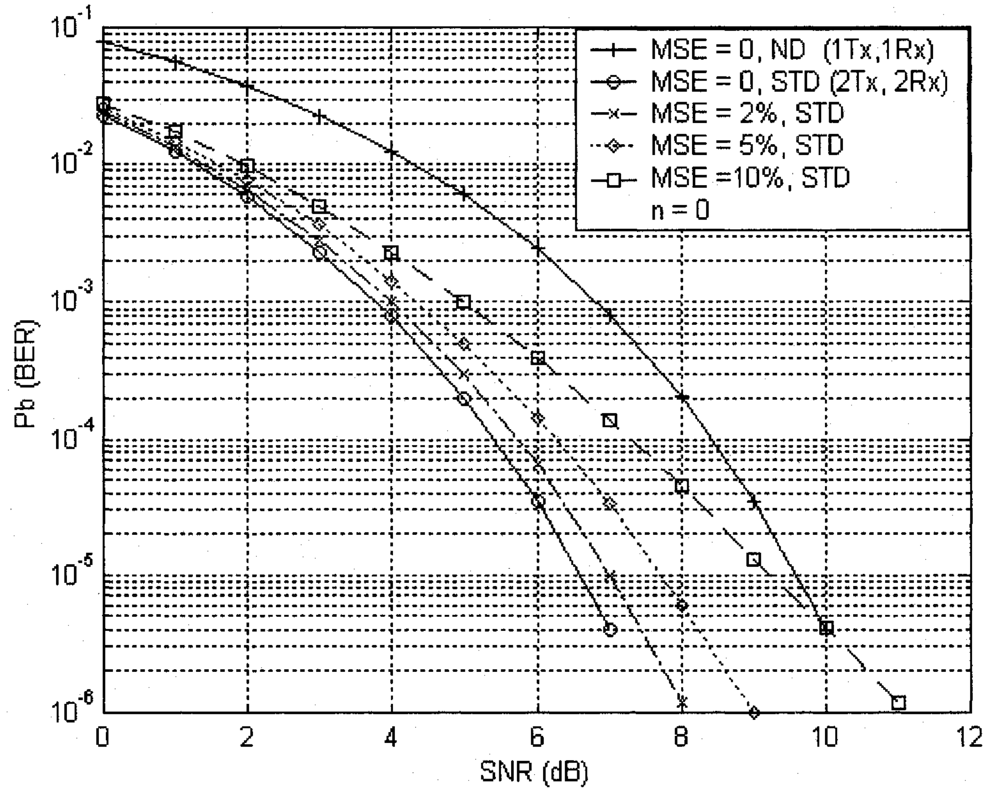

Figure 4.2: BER performance of the STD scheme with two receivers against SNR for various values of channel estimation error at $n = 0$.

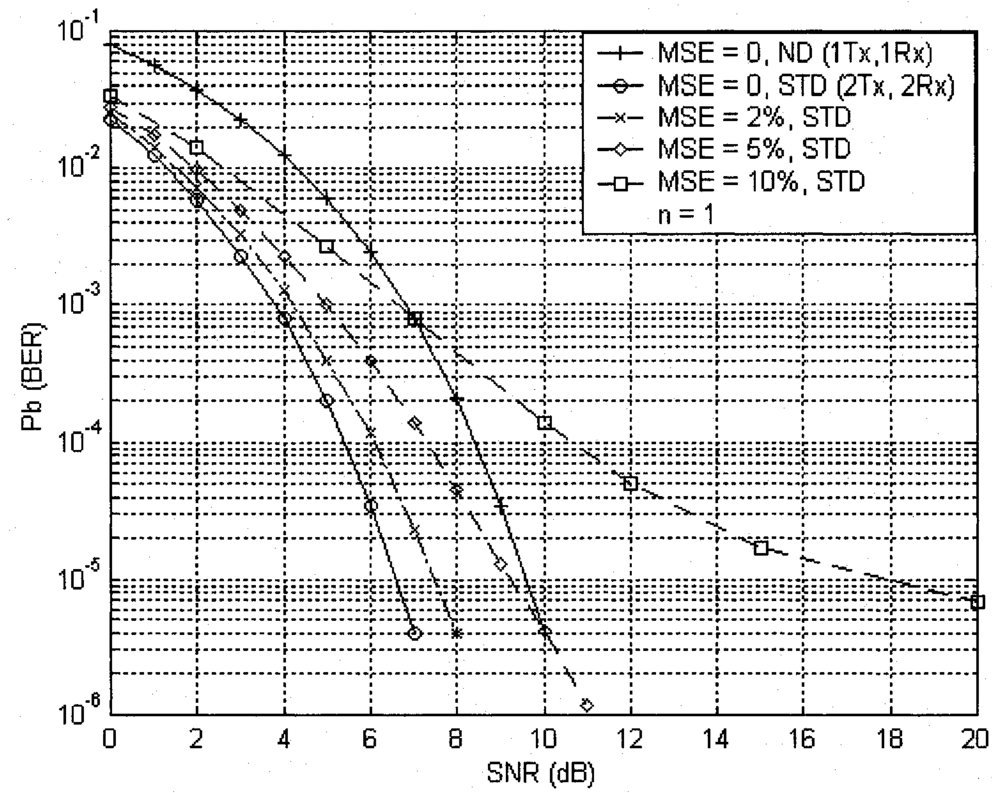

Figure 4.3: BER performance of the STD scheme with two receivers as a function of SNR for various values of channel estimation error at $n = 1$.

In Fig. 4.4, the BER curves are plotted versus the SNR for various values of channel estimation error with n = 2. To achieve a BER of 10^{-3}, it is found that the deterioration in performance is in the range of 1 dB to 5.6 dB as the variance of error is varied. In addition, the BER degrades quite rapidly with increase σ_{ε}^{2} and increase of SNR beyond the threshold SNR Γ_t.

The BER curves of the STD scheme with n = 4 are exhibited in Fig. 4.5. From these curves we can see that an increment in the thermal SNR of 1 dB, 2.2 dB, and 7.7 dB are required to obtain a BER of 5×10^{-3} with the three different values of σ_{ε}^{2} compared to perfect channel estimation case. At high SNR, the deterioration of performance in the STD scheme exceeds the benefits achieved over ND system as the channel estimation error variance increases. The BER difference between perfect and imperfect channel estimation increases quite rapidly as SNR increases. Moreover, there is an irreducible error floor with σ_{ε}^{2} of 10% for SNR above 15 dB.

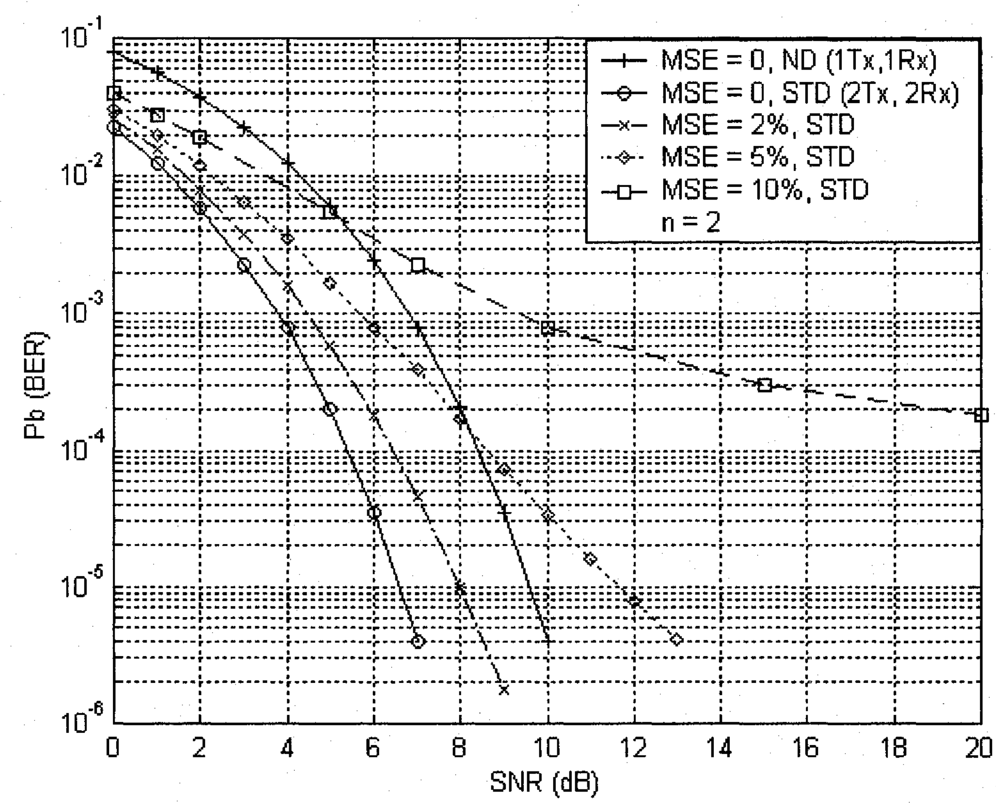

Figure 4.4: The BER curves of the STD scheme with two receivers versus SNR for several values of channel estimation error at $n = 2$.

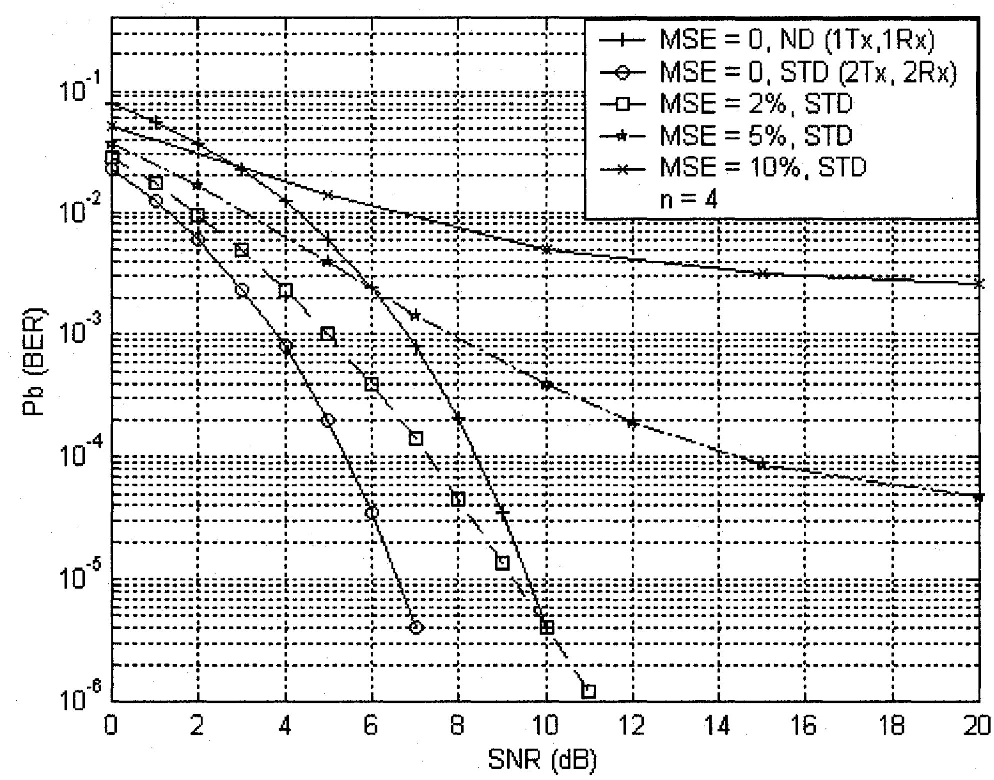

Figure 4.5: The BER curves of the STD scheme with two receivers as a function of SNR for several values of channel estimation error at $n = 4$.

4.4 Conclusions

The error performance of a simple transmit diversity (STD) scheme due to imperfect channel estimation over multipath channels that lead to frequency selective fading was evaluated. An expression was derived for the SNR at the output of the equalizer, which is employed to compensate for the ISI in the received signal. To analyze the effects of channel estimation error, the BER of the STD scheme was obtained, based on the derived SNR. These results are applicable to any channel estimation technique.

Performance curves showed that there is a significant deterioration in BER when the SNR, number of multipath components, and the channel estimation error are large. To achieve a BER of 10^{-4}, the use of STD scheme with $\sigma_{\varepsilon}^{2} = 5\%$ and $n = 4$ result in a SNR loss of about 8 dB compared to the perfect channel estimation case. For $\sigma_{\varepsilon}^{2} = 10\%$ and $n = 4$, there is an irreducible error floor at SNR values higher than 10 dB, which is required to achieve a BER of 5×10^{-3}.

It was also shown that the degradation of performance in the STD scheme exceeds the benefits achieved over no diversity (ND) case when the SNR and the channel estimation error are large. Possible trade-off can be considered between the transmitted power and the channel estimation errors. Increasing the SNR level can compensate the degradation in the BER performance due to large estimation error.

Chapter 5

A New Channel Estimation Technique for Multiple Input Multiple Output Systems

5.1 Introduction

Transmit diversity has emerged in the last decade as an effective mean for achieving spatial diversity in fading channels with an antenna array at the transmitter [86]. Band-limited channels are narrow pipes that do not accommodate rapid flow of data. Deploying multiple transmit antennas broadens this data pipe by exploiting the spatial dimension [87]. Transmit diversity is an attractive area as it enables the designer to move diversity burden from the subscriber unit to the base station where the use of multiple antennas is more feasible [18].

However, the main problem with deploying transmit diversity resides in the lack of instantaneous channel state information (CSI) at the transmitter. Thus, one must employ a channel code that will guarantee good performance over a broad range of channel realizations [88]. An important such class of Multiple-Input Multiple-Output (MIMO) communication system, i.e. Space-time Coding (STC) schemes combine the channel code design and the use of multiple transmit antennas. The encoded data is split into n_T streams that are simultaneously transmitted using n_T transmit antennas.

At a certain reception antenna, the received signal is a linear combination of these simultaneously transmitted symbols (with fading coefficients as weights) corrupted by noise and channel-induced intersymbol interference (ISI). Space-time decoding algorithms follow channel estimation techniques incorporated at the receiver in order to achieve both near perfect diversity advantage and coding gain [12]. Alamouti [16] presented a two-branch transmit diversity scheme, which utilizes two transmit and two receive antennas.

The scheme in [16] achieves the same diversity benefits at the subscriber unit as maximal ratio receiver combining (MRRC) with one transmit antenna and four receive antennas. Although the diversity benefits can be replicated, the gain from coherent combining is not noticeable unless the channel is known at the transmitter. Thus, the effectiveness of [16] and most STC schemes relies on accurate multi-channel estimation techniques at the receiver in order to achieve diversity advantage and coding gain [89]. Moreover, the scheme assumes that there is no ISI in the channel. This will not be the case if the channel experiences multipath and a non-negligible delay spread [90].

5.2 Preliminaries

One method of channel estimation is to turn off all transmit antennas apart from antenna i at some time instant and to send a pilot signal using antenna i. The fade coefficients $\overline{\alpha}_{i,j}$, between transmit antenna i and receive antenna j, are then estimated for $1 \leq j \leq n_R$. This procedure is repeated for $1 \leq i \leq n_T$ until all coefficients $\overline{\alpha}_{i,j}$, $i = 1,2,\ldots,n_T$, $j = 1,2,\ldots,n_R$ are estimated, where the over bar indicates that the $\overline{\alpha}_{i,j}$ consists of n_m channel multipath coefficients. A second method of estimation is to send orthogonal sequences of signals (similar to Walsh functions) for pilot signals, one from each transmit antenna [32].

Minimum mean square error (MMSE) and least-squares (LS) channel estimators have been widely investigated for use in multiple transmit antenna systems. These estimators can be implemented to efficiently estimate the channel given certain knowledge about the channel statistics. The MMSE estimator assumes a priori knowledge of noise variance and channel covariance. It seeks to minimize the mean square error between the estimated and the quantity being estimated. Mean-square estimation is based on statistical averages.

Least-squares estimators require very little information on the statistics of the parameters to be estimated. The LS criterion specifies that one should choose as an estimate the value that minimizes the sum of squared errors. Minimization of the squared error can be achieved either by differentiation or by application of the orthogonality principle. The MMSE estimator has good performance but high complexity. The LS estimator has low complexity but its performance is not as good as that of the MMSE estimator [91], [92].

In [31], orthogonal pilot sequences insertion along with training sequences are used to estimate the channel at the receiver. The estimates of the channel coefficients are chosen such that the mean square error (MSE) is minimized. However, simulation results showed a substantial error flooring due to severe ISI caused by the channel. An MMSE estimator that makes full use of the correlation of the channel frequency response at different times and frequencies was derived in [93]. If the channel parameters are estimated using the method developed in [93], the signals from other transmitter antenna(s) will become interference, and the MSE of the estimation process will be very large.

Consequently, a channel parameter estimation approach was proposed for transmitter diversity using STC in [94]. However, the approach in [94] required the inversion of a large matrix, which involves intensive computation. To simplify computation, Ye Li proposed two reduced-complexity channel estimation techniques in [95], which are based on minimum mean-square error channel estimation for orthogonal frequency division multiplexing (OFDM) systems. Results in [95] show the deleterious effects of inaccurate channel estimation on the performance of ST codes. Hence, novel parameter estimation approaches are desired to improve performance MIMO systems using STC.

5.3 Channel Estimation Based on MAP Approach

The challenge in channel estimation for multiple-transmit-antenna systems over the single-transmit-antenna case is the increased number of channel parameters to be estimated and reduced transmit power (by a factor of two) for each transmit antenna. Intersymbol interference created by the multipath in band-limited time dispersive channels distorts the transmitted signal, causing bit errors at the receiver [57]. The channel is viewed as path gains having complex-valued parameters with unknown deterministic quantities. The channel parameters are the attenuation and delay incurred by the signal transversal along the propagation paths.

The channel is assumed to be of finite length, thus the maximum number of paths is considered known a priori [96], [97]. The complex channel parameters are further treated as two real-valued tap coefficients; each taking one of M possible amplitude levels $\{A_m\}$ with equal probability. In order to evaluate the channel parameters, two known fixed-length training sequences are simultaneously sent from the two transmission antennas. The training sequence is typically a pseudorandom binary signal or a fixed, prescribed bit pattern. Assuming the channel varies very slowly for the duration of the training sequence transmission, the parameters remain constant during that time.

The effectiveness of STC schemes requires the development of practical and high-performance algorithms for channel estimation. This work presents a new channel estimation technique amenable to STC. The objective is to derive the estimates of the MIMO channel parameters for a simple two-branch transmit diversity scheme from the received waveforms at one of the reception antennas. The proposed estimation technique is based on an iterative procedure built around the maximum a posteriori (MAP) probability that is related to the basics of turbo coding. First, we derive the various expressions required to compute the posteriori probabilities for each coefficient. Based on the MAP criterion, we select as a coefficient value the amplitude level that gives the maximum probability [98], [99].

Unlike classic estimation techniques, we iterate on the different probabilities of different coefficients rather than on the coefficient values themselves. The concept is to pass the reliability of the decisions made in one iteration to serve as a refinement of the prior knowledge for the next iteration and repeat this process several times to produce better decisions. The estimation process is done in discrete values framework. In practice, the channel coefficients are continuous but discrete values are used to enable us to implement the proposed MAP approach. Such discretization or quantization is not new to communication fields, e.g. matched filters and other components are based on discrete realization, even if the original model is continuous.

This new MAP approach inherently uses the principle of turbo decoding, which utilizes both the a priori probability and the a posteriori probabilities (APPs). All estimation techniques used to estimate channel coefficients that are continuous apply the Bayes theorem. However, their treatment does not include the a priori probability of the channel coefficient, which we propose herein to improve results iteration after iteration. Both the APPs and the a priori probability are updated which leads to improving the convergence rate and the mean square error of channel estimation.

5.3.1 Two Branch Transmit Diversity Scheme

Diversity is one of the most important factors in providing reliable communications over wireless channels. To improve the quality and data rates, one can use multiple transmit and receive antennas to obtain diversity. Fig. 5.1 shows the baseband representation of the two branch transmit diversity scheme with two receive antennas that focuses on the channel estimation.

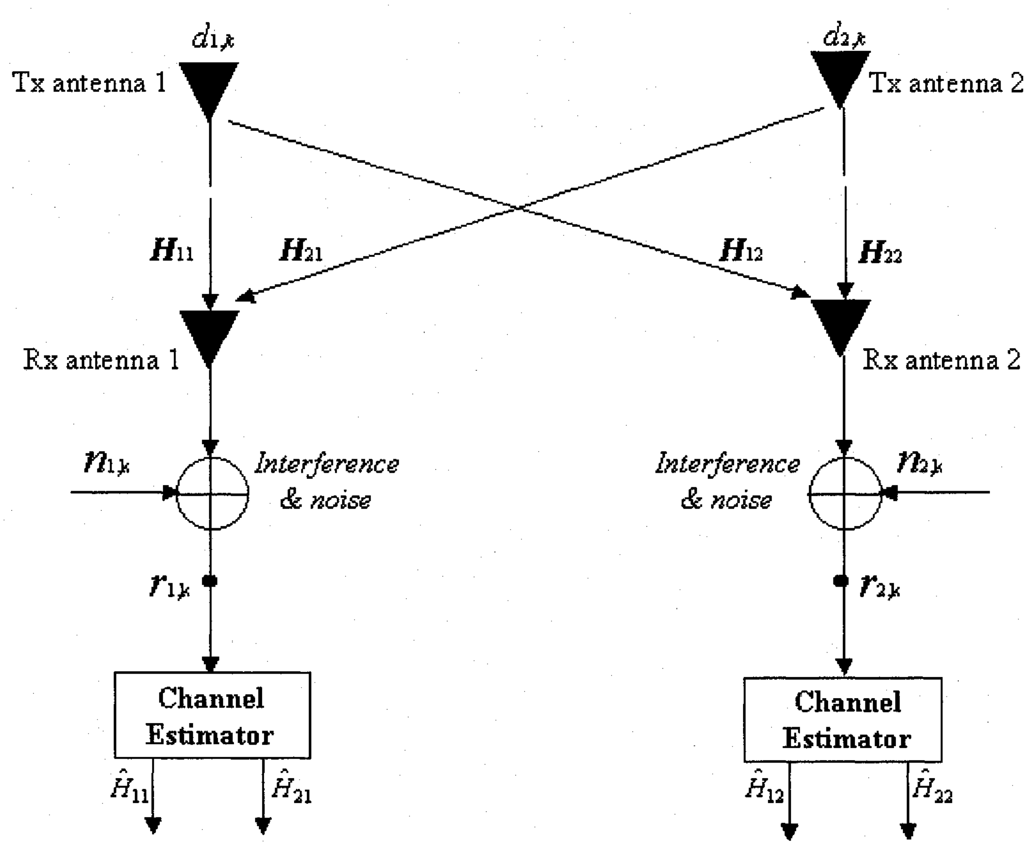

Figure 5.1: The two branch transmit diversity scheme with two receivers.

At a given symbol period, two signals are simultaneously transmitted from the two transmit antennas. The signal transmitted from antennas one and two are denoted by $d_{1,k}$ and $d_{2,k}$, respectively. The transmitted signals are given by

$$\begin{aligned} d_{1,k} &= x_{1,k} + jy_{1,k} \\ d_{2,k} &= x_{2,k} + jy_{2,k} \end{aligned} \tag{5.1}$$

where k is the time index, k = 1, 2, …, K (maximum number of received preamble signals). The channel between the transmit antenna i and the receive antenna j is defined as H_{ij}, i = 1, 2 and j = 1, 2. The received signals are a linear superposition of the simultaneously transmitted symbols with the channel fading parameters as weights. The signals received at receive antennas one and two are denoted by $r_{1,k}$ and $r_{2,k}$, respectively. The received signals can be expressed as

$$\begin{aligned} r_{1,k} &= H_{11}d_{1,k} + H_{21}d_{2,k} + n_{1,k} \\ r_{2,k} &= H_{12}d_{1,k} + H_{22}d_{2,k} + n_{1,k} \end{aligned} \tag{5.2}$$

where $n_{1,k}$ and $n_{2,k}$ are complex random variables representing receiver thermal noise and interference.

The basic functional blocks used for estimating the channels shown in Fig. 5.1 are independent of each other, but performs a similar function. Thus, we focus on one channel estimator to explain the principle used to provide the space-time decoder with a perfect knowledge of the channels. Fig. 5.2 shows a two branch transmit diversity scheme with one receiver. The encoding and transmission sequence of the training symbols for this configuration is identical to the case of two receivers.

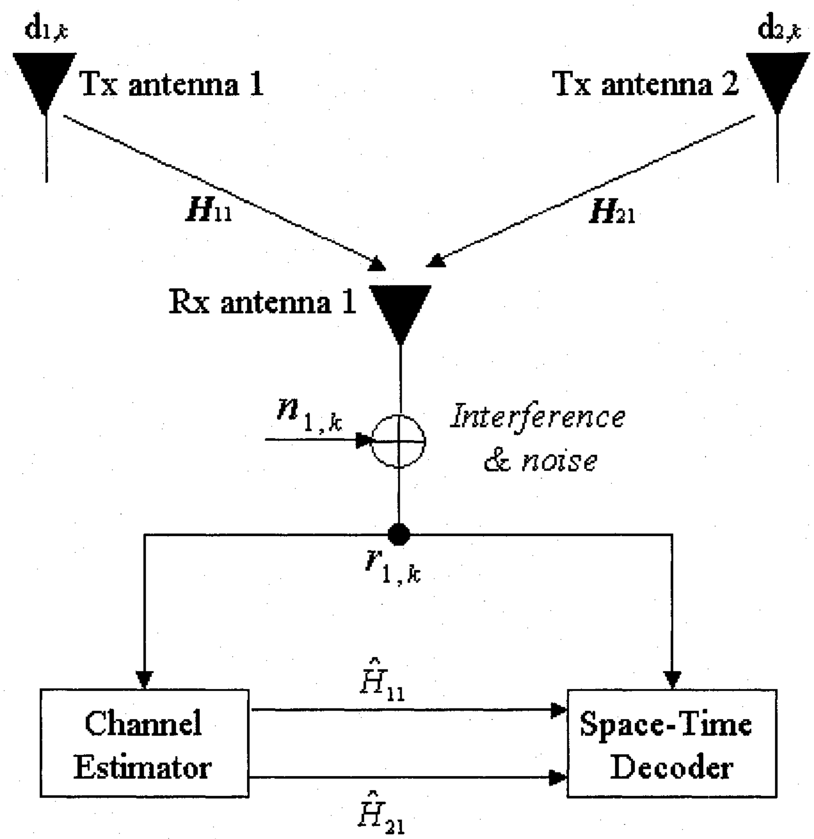

Figure 5.2: The two branch transmit diversity scheme with one receiver.

5.3.2 Model of MIMO Channel with ISI

Due to multipath propagation and time dispersion, the channel between each transmit and receive antenna pair is characterized as a frequency selective channel that results in ISI between each transmit and receive antenna pair. The antennas on both ends must be separated far enough to ensure independently fading channels. The delay spread per transmit and receive antenna pair is assumed to be the same for all channels. If the transmit antennas are physically co-located at the same station, then this assumption is justified by the fact that the number of multipath components with different delays is dictated by large structures and reflecting objects [85].

In dealing with band-limited channels that result in ISI, it is convenient to develop an equivalent discrete-time model for the system. To model digital signal transmission with diversity, the discrete-time channel model is extended to provide parallel multi-channel operation. Fig. 5.3 illustrates the channel model of the equivalent discrete-time system between the two transmit antennas and receive antenna one with four delay paths for each channel.

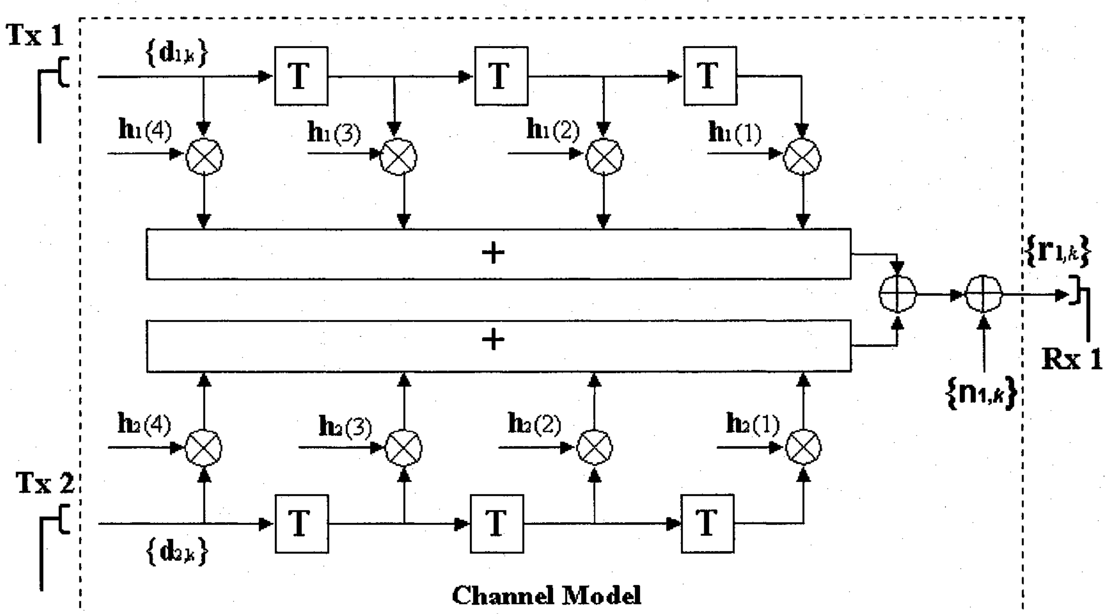

Figure 5.3: Equivalent discrete-time model of MIMO channel with ISI.

The lth tap of the frequency selective channel from transmit antenna i to receive antenna one is denoted $h_i(l)$ for $l = 1,\ldots, L$ and $i = 1, 2$. The tap gain parameters are complex-valued and given by

$$h_i(l) \ = \alpha_{i,l} + j\beta_{i,l} \tag{5.3}$$

The output sequence $\{r_{1,k}\}$ at receive antenna one can be expressed as [100]

$$r_{1,k} = \left[\sum_{i=1}^{2}\sum_{l=1}^{L} h_i(l) d_{i,k+l-L}\right] + n_{1,k} \tag{5.4}$$

Substituting (5.1) and (5.3) in (5.4), the output sequence becomes:

$$\begin{aligned} r_{1,k} &= \left\{\left[\sum_{i=1}^{2}\sum_{l=1}^{L}(\alpha_{i,l} + j\beta_{i,i})(x_{i,k+l-L} + jy_{i,k+l-L})\right] + (\eta_{I,k} + j\eta_{Q,k})\right\} \\ r_{1,k} &= \theta_{1,k} + j\sigma_{1,k} \end{aligned} \tag{5.5}$$

where $\{\eta_{I,k}\}$ and $\{\eta_{Q,k}\}$ are AWGN sequences each with zero mean and variance $N_0/2$ [77], [101].

We have implicitly assumed a static channel, meaning that the channel parameters are either fixed or vary so slowly that they remain constant for the channel estimation period or one data packet. Expanding the output sequence and equating separately the real and imaginary parts, we obtain

$$\begin{aligned} \theta_{1,k} &= \left[\sum_{i=1}^{2}\ \sum_{l=1}^{L}(\alpha_{i,l}x_{i,k+l-L} - \beta_{i,l}y_{i,k+l-L})\right] + \eta_{I,k} \\ \sigma_{1,k} &= \left[\sum_{i=1}^{2}\ \sum_{l=1}^{L}(\beta_{i,l}x_{i,k+l-L} + \alpha_{i,l}y_{i,k+l-L})\right] + \eta_{Q,k} \end{aligned} \tag{5.6}$$

At the receiver, few symbols must be first received in order to start the estimation process; the estimate of each coefficient is based on an observation of a sequence of received signals. Thus, an index k is added to the coefficients' indices to represent the estimation at a certain kth symbol. Thus, $\alpha_{i,l,k}$ and $\beta_{i,l,k}$ represent the kth symbol estimate of the actual channel coefficients $\alpha_{i,l}$ and $\beta_{i,l}$, respectively. So, (5.6) could be rewritten as [101]

$$\begin{aligned}\theta_{1,k+L} &= \left[\sum_{i=1}^{2}\sum_{l=1}^{L}(\alpha_{i,l,k}x_{i,k+l} - \beta_{i,l,k}y_{i,k+l})\right] + \eta_{I,k+L}\\ \sigma_{1,k+L} &= \left[\sum_{i=1}^{2}\sum_{l=1}^{L}(\beta_{i,l,k}x_{i,k+l} + \alpha_{i,l,k}y_{i,k+l})\right] + \eta_{Q,k+L}\end{aligned} \tag{5.7}$$

The above equation represents the received sequence for $k = 0, 1, \ldots, K-4$ [77].

5.4 Principle of the Channel Estimation Algorithm

Successful implementation of STC in wireless systems, which employ multiple transmit antennas and single or multiple receive antennas requires the development of practical and high-performance signal processing algorithms for channel estimation. The estimation technique is based on an iterative procedure derived through the MAP approach. To implement the MAP approach the channel parameters are approximated by discrete values from a finite alphabet. The channel coefficients are continuous in practice, but using discrete realization even though the original model is continuous, enables us to apply the proposed MAP approach.

This new approach inherently uses the principle of turbo decoding, which utilizes both the a priori probability and the a posteriori probabilities (APPs). All maximum likelihood estimation techniques used to estimate channel coefficients, which are continuous, apply the Bayes theorem. However, their treatment does not include the a priori probability of the channel coefficients, which we propose herein to improve the mean square error of the channel estimation. The technique will focus on estimating the channel parameters for a simple two-branch transmit diversity scheme operating in a multipath environment.

Multipath in band-limited time dispersive channels creates ISI, which distorts the transmitted signal. The channel model considered in this work was shown in Fig. 5.3. All taps of the frequency selective channel are viewed as complex-valued parameters having unknown deterministic quantities. The parameters are further treated as two real-valued tap coefficients. Each coefficient takes one of M possible equally spaced amplitude values with equal probabilities. The coefficients $\alpha_{j,l}$ and $\beta_{j,l}$, $j = 1, 2$, $l = 1, 2, \ldots, L$, are the real and imaginary parts of the lth tap gain from transmit antenna j to receive antenna one.

Without the loss of generality, each channel is represented by four taps for convenience purposes to estimate when a known training sequence embedded in the data symbols are transmitted. The training sequence is typically a pseudorandom binary signal or a fixed, prescribed bit pattern. In addition, we assume the channel parameters remain constant for the duration of the transmitted training sequence time. The technique will be illustrated in the context of detecting a pulse amplitude modulated (PAM) signal with M possible levels.

5.4.1 Estimation Process Basis

Suppose it is desired to estimate the coefficient $\alpha_{j,l}$ in the kth symbol interval, and $r_{1,k}$ be the observed received signal. Similar to the approach in [58], we compute the APPs

$$P(\alpha_{j,l,k} = A_m | r_{1,k}) \tag{5.8}$$

for the M possible amplitude levels and choose as a coefficient value the amplitude level with the largest probability. Thus, the MAP criterion for deciding on the estimate of a certain coefficient $\alpha_{j,l,k}$ is

$$\tilde{\alpha}_{j,l,k} = \arg\left\{\max_{A_m} P(\alpha_{j,l,k} = A_m \mid r_{1,k})\right\} \tag{5.9}$$

where *arg* denotes the quantized value of A_m that maximize the right hand side (RHS) of (5.9). The APPs in (5.8) are given by:

$$P(\alpha_{j,l,k} = A_m \mid r_{1,k}) = \\ \underbrace{\sum \cdot\cdot \sum \cdot\cdot}_{other\ \alpha_S} \underbrace{\sum \cdot\cdot \sum}_{\beta_{j,l,k}} \frac{p(r_{1,k} \mid \alpha_{j,l,k} = A_m) P(\alpha_{j,l,k-1} = A_m)}{\sum_{i=1}^{M} p(r_{1,k} \mid \alpha_{j,l,k} = A_i) P(\alpha_{j,l,k-1} = A_i)} Z_{j,l} \tag{5.10}$$

where

j is the transmitting antenna identifier, $j = 1, 2$.

l is the lth path of the channel, $l = 1, 2, \ldots, L = 4$.

k is the time index, $k = 0, 1, \ldots, K- 4$.

m is the number of possible levels for the channel coefficient, $m = 1, 2, \ldots, M$.

A_m is the mth amplitude level.

The multiple summations perform averaging over all possible amplitude levels of the channel coefficients except the one being currently estimated $\alpha_{j,l,k}$, when the training symbols are known. The probability $P(\alpha_{j,l,k-1} = A_m)$ is the a priori probability of the coefficient for the A_m amplitude level. This a priori probability can be obtained from the probability computed in the previous iteration. Thus, it establishes the iterative nature of the estimation algorithm.

For the received signals $r_{1,k}$, the channel taps $h_i(l)$, and the symbols $d_{j,k}$, which are all complex-valued the conditional probability density for additive Gaussian noise is

$$\begin{aligned} &p(r_{1,k} \mid \alpha_{j,l,k} = A_m) = \\ &\frac{1}{\sqrt{2\pi N_0}} \exp\left\{ \frac{-1}{2N_0} \left[\theta_{k+L} - \sum_{j=1}^{2} \sum_{l=1}^{L} \left(\alpha_{j,l,k} x_{j,k+l} - \beta_{j,l,k} y_{j,k+l} \right) \right]^2 \right\} \times \\ &\frac{1}{\sqrt{2\pi N_0}} \exp\left\{ \frac{-1}{2N_0} \left[\sigma_{k+L} - \sum_{j=1}^{2} \sum_{l=1}^{L} \left(\beta_{j,l,k} x_{j,k+l} + \alpha_{j,l,k} y_{j,k+l} \right) \right]^2 \right\} \Bigg| \alpha_{j,l,k} = A_m \end{aligned} \tag{5.11}$$

for $j = 1, 2$ and $l = 1, \ldots, L$. For mathematical convenience, we define

$$\begin{aligned} R_m &= \left[\theta_{k+L} - \sum_{j=1}^{2} \sum_{l=1}^{L} (\alpha_{j,l,k} x_{j,k+l} - \beta_{j,l,k} y_{j,k+l}) \right]^2 \Bigg| \alpha_{j,l,k} = A_m \\ I_m &= \left[\sigma_{k+L} - \sum_{j=1}^{2} \sum_{l=1}^{L} (\beta_{j,l,k} x_{j,k+l} + \alpha_{j,l,k} y_{j,k+l}) \right]^2 \Bigg| \alpha_{j,l,k} = A_m \end{aligned} \tag{5.12}$$

Thus, the conditional probability density in (5.11) can be expressed as

$$p(r_{1,k} \mid \alpha_{j,l,k} = A_m) = \frac{1}{2\pi N_0} \exp\left(-\frac{R_m + I_m}{2N_0} \right) \Bigg| \alpha_{j,l,k} = A_m \tag{5.13}$$

The joint probability $Z_{j,l}$ in (5.10) is computed from the probabilities of the possible amplitude levels being assumed in each of the multiple summations corresponding to various coefficients. The joint probability is given by:

$$Z_{j,l} = \left(\underbrace{\prod_{i=1}^{2} \prod_{n=1}^{4}}_{If \ i=j, then \ n \neq l} P(\alpha_{i,n,k} = A_{1...M}) \right) \cdot \left(\prod_{s=1}^{2} \prod_{t=1}^{4} P(\beta_{s,t,k} = A_{1...M}) \right) \tag{5.14}$$

The notation $A_{1...M}$ is used to represent the amplitude level being assumed by the corresponding coefficient as dictated by the index of the corresponding summation in (5.10). Thus, the general relation to compute the APPs given in (5.10) can be expressed as follows:

$$P(\alpha_{j,l,k} = A_m \mid r_{1,k}) = \\ \underbrace{\sum \cdots \sum \cdots}_{other \ \alpha_S} \underbrace{\sum \cdot \cdot \sum}_{\beta_{j,l,k}} \frac{\exp\left(-\frac{R_m + I_m \left| \alpha_{j,l,k} = A_m \right.}{2N_0} \right) P\left(\alpha_{j,l,k-1} = A_m \right)}{\sum_{i=1}^{M} \exp\left(-\frac{R_i + I_i \left| \alpha_{j,l,k} = A_i \right.}{2N_0} \right) P\left(\alpha_{j,l,k-1} = A_i \right)} \times Z_{j,l} \tag{5.15}$$

In computing the RHS of (5.15), the coefficient being currently estimated will be set equal to amplitude level A_m, while each of the other coefficients will assume an amplitude level set by its corresponding summation value.

5.4.2 First Coefficient Estimation

In order to reduce the computational complexity, we set a constraint on the transmitted preamble symbols without loss of generality. Such that the real and imaginary parts are equal, that to say $x_{j,k} = y_{j,k}$ for j = 1, 2 and k = 1, 2, …, K. This will allow us to eliminate a number of terms and to reduce the number of variables involved in the computation of the RHS of (5.15). Then, for j = 1, 2 and l = 1, …, 4, we let

$$\varphi_{j,l} = \alpha_{j,l,k} x_{j,k+l} = \alpha_{j,l,k} y_{j,k+l} \text{ and } \psi_{j,l} = \beta_{j,l,k} y_{j,k+l} = \beta_{j,l,k} x_{j,k+l}$$

To illustrate the estimation process, we start by investigating a particular coefficient $\alpha_{1,1,k}$. From the general relation given in (5.15), we want to derive the relation required to optimize the quantized value of $\alpha_{1,1,k}$. By letting j = 1 and l = 1, we focus on $\varphi_{1,1}$, which is a function of the currently considered coefficient. Thus, we have

$$\varphi_{1,1} = \alpha_{1,1,k} x_{1,k+1} = \alpha_{1,1,k} y_{1,k+1}$$

The sum of the R_m and I_m given in (5.12), could be rewritten as a summation of terms involving $\varphi_{1,1}$, denoted by $F(\varphi_{1,1})$ and those not involving $\varphi_{1,1}$, denoted by $F(O)$. Terms not involving $\varphi_{1,1}$ can be cancelled out from the numerator and the denominator of (5.15). Thus, the function $F(\varphi_{1,1})$ can be expressed as:

$$\begin{aligned} F(\varphi_{1,1}) = 2\varphi_{1,1}^2 + 4\varphi_{1,1}(\varphi_{1,2} + \varphi_{1,3} + \varphi_{1,4} + \varphi_{2,1} + \\ \varphi_{2,2} + \varphi_{2,3} + \varphi_{2,4}) - 2\varphi_{1,1}(\theta_{k+L} + \sigma_{k+L}) \end{aligned} \tag{5.16}$$

Therefore, the APPs required to optimize coefficient $\alpha_{1,1,k}$ are given by

$P(\alpha_{1,1,k} = A_m \mid r_{1,k}) =$

$$\sum_{\alpha_{1,2,k}} \cdot\cdot \sum_{\alpha_{1,3,k}} \cdot\cdot \sum_{\alpha_{1,4,k}} \cdot\cdot \sum_{\alpha_{2,1,k}} \cdot\cdot \sum_{\alpha_{2,2,k}} \cdot\cdot \sum_{\alpha_{2,3,k}} \cdot\cdot \sum_{\alpha_{2,4,k}} \cdot\cdot \sum_{\beta_{1,1,k}} \cdot\cdot \sum_{\beta_{1,2,k}} \cdot\cdot \sum_{\beta_{1,3,k}} \cdot\cdot \sum_{\beta_{1,4,k}} \cdot\cdot \sum_{\beta_{2,1,k}} \cdot\cdot \sum_{\beta_{2,2,k}} \cdot\cdot$$
$$\sum_{\beta_{2,3,k}} \cdot\cdot \sum_{\beta_{2,4,k}} \frac{\exp\left(-\frac{F(\varphi_{1,1}) \mid \varphi_{1,1} = A_m x_{1,k+1}}{2N_0}\right) P\left(\alpha_{1,1,k-1} = A_m\right)}{\sum_{i=1}^{M} \exp\left(-\frac{F(\varphi_{1,1}) \mid \varphi_{1,1} = A_i x_{1,k+1}}{2N_0}\right) P(\alpha_{1,1,k-1} = A_i)} \times Z_{1,1} \tag{5.17}$$

The joint probability $Z_{1,1}$ is given as follows:

$$Z_{1,1} = \left(\underbrace{\prod_{i=1}^{2} \prod_{n=1}^{4}}_{i=1, n \neq 1} P\left(\alpha_{i,n,k} = A_{1 \ldots M}\right) \right) \cdot \left(\prod_{s=1}^{2} \prod_{t=1}^{4} P\left(\beta_{s,t,k} = A_{1 \ldots M}\right) \right) \tag{5.18}$$

Since no prior knowledge is available, we assume all amplitude levels are equally probable. Thus, the probabilities required to compute the above joint probability are set equal to $1/M$, for all M possible amplitude levels for each of the coefficients. The notation $A_{1 \ldots M}$ is used to represent the amplitude level indicated by the coefficient corresponding to the summation index.

Equation (5.17) will be computed for the M possible amplitude levels being assumed for the channel coefficient. The decision criterion is based on selecting the value corresponding to the maximum of the set of computed APPs. This MAP decision criterion to estimate $\alpha_{1,1,k}$, denoted by $\widetilde{\alpha}_{1,1,k}$ is represented as

$$\widetilde{\alpha}_{1,1,k} = \arg\left\{ \max_{A_m} P(\alpha_{1,1,k} = A_m \mid r_{1,k}) \right\} \tag{5.19}$$

5.4.3 Sequential Coefficients Optimization

Consider next the optimization of the coefficient $\alpha_{1,2,k}$. Thus, we let j = 1, l = 2 and concentrate on the variable $\varphi_{1,2} = \alpha_{1,2,k} x_{1,k+2}$. The sum of R_m and I_m given in (5.12), could be rewritten as a summation of terms involving $\varphi_{1,2}$, denoted by $F(\varphi_{1,2})$ and those not involving $\varphi_{1,2}$, denoted by $F(O')$. Terms not involving $\varphi_{1,2}$ can be cancelled out from the numerator and the denominator of (5.15). Consequently, the function $F(\varphi_{1,2})$ is expressed as follows:

$$\begin{aligned} F(\varphi_{1,2}) = 2\varphi_{1,2}^2 + 4\varphi_{1,2}(\varphi_{1,1} + \varphi_{1,3} + \varphi_{1,4} + \varphi_{2,1} + \\ \varphi_{2,2} + \varphi_{2,3} + \varphi_{2,4}) - 2\varphi_{1,2}(\theta_{k+L} + \sigma_{k+L}) \end{aligned} \tag{5.20}$$

Hence, to estimate the coefficient $\alpha_{1,2,k}$ on the basis of the received sequence, we compute the following APPs

$$\begin{aligned} & P(\alpha_{1,2,k} = A_m \mid r_{1,k}) = \\ & \sum_{\alpha_{1,1,k}} .. \sum_{\alpha_{1,3,k}} .. \sum_{\alpha_{1,4,k}} .. \sum_{\alpha_{2,1,k}} .. \sum_{\alpha_{2,2,k}} .. \sum_{\alpha_{2,3,k}} .. \sum_{\alpha_{2,4,k}} .. \sum_{\beta_{1,1,k}} .. \sum_{\beta_{1,2,k}} .. \sum_{\beta_{1,3,k}} .. \sum_{\beta_{1,4,k}} .. \sum_{\beta_{2,1,k}} .. \sum_{\beta_{2,2,k}} .. \\ & \sum_{\beta_{2,3,k}} .. \sum_{\beta_{2,4,k}} \frac{\exp\left(-\dfrac{F(\varphi_{1,2}) \mid \varphi_{1,2} = A_m x_{1,k+2}}{2N_0} \right) P(\alpha_{1,2,k-1} = A_m)}{\sum_{i=1}^{M} \exp\left(-\dfrac{F(\varphi_{1,2}) \mid \varphi_{1,2} = A_i x_{1,k+2}}{2N_0} \right) P(\alpha_{1,2,k-1} = A_i)} \times Z_{1,2} \end{aligned} \tag{5.21}$$

for the M possible amplitude levels and select as a value for the coefficient being investigated the amplitude level having the largest probability. The joint probability $Z_{1,2}$ in (5.21) is given by:

$$Z_{1,2} = \left(\underbrace{\prod_{i=1}^{2} \prod_{n=1}^{4}}_{i=1, n \neq 2} P(\alpha_{i,n,k} = A_{1...M}) \right) \cdot \left(\prod_{s=1}^{2} \prod_{t=1}^{4} P(\beta_{s,t,k} = A_{1...M}) \right) \tag{5.22}$$

In computing the above joint probability, the involved probabilities for the coefficient $\alpha_{1,1,k}$, $P(\alpha_{1,1,k} = A_{1...M})$ correspondence to the APPs previously computed in (5.17). This is a refinement of our prior knowledge, while the other required probabilities are set equal to $1/M$, for all M possible amplitude levels for each of the other coefficients.

Following the same procedure, it is straightforward to derive the relations required to optimize the remaining α_s coefficients. Moreover, the same process could be followed to develop the required equations in order to evaluate the coefficient $\beta_{j,l,k}$ for j = 1, 2 and l = 1,..., 4. Thus, to optimize the last coefficient $\beta_{2,4,k}$, we let j = 2, l = 4 and concentrate on the variable $\psi_{2,4} = \beta_{2,4,k} y_{2,k+4}$. Terms involving the variable $\psi_{2,4}$ can be written in a function denoted by $F(\psi_{2,4})$, which is expressed as:

$$\begin{aligned} F(\psi_{2,4}) = 2\psi_{2,4}^2 + 2\psi_{2,4}(\psi_{1,1} + \psi_{1,2} + \psi_{1,3} + \psi_{1,4} + \\ \psi_{2,1} + \psi_{2,2} + \psi_{2,3}) + 2\psi_{2,4}(\theta_{k+L} - \sigma_{k+L}) \end{aligned} \tag{5.23}$$

Hence, we compute the following APPs

$$\begin{aligned} &P(\beta_{2,4,k} = A_m \mid r_{1,k}) = \\ &\sum_{\alpha_{1,1,k}} \cdot\cdot \sum_{\alpha_{1,2,k}} \cdot\cdot \sum_{\alpha_{1,3,k}} \cdot\cdot \sum_{\alpha_{1,4,k}} \cdot\cdot \sum_{\alpha_{2,1,k}} \cdot\cdot \sum_{\alpha_{2,2,k}} \cdot\cdot \sum_{\alpha_{2,3,k}} \cdot\cdot \sum_{\alpha_{2,4,k}} \cdot\cdot \sum_{\beta_{1,1,k}} \cdot\cdot \sum_{\beta_{1,2,k}} \cdot\cdot \sum_{\beta_{1,3,k}} \cdot\cdot \sum_{\beta_{1,4,k}} \cdot\cdot \sum_{\beta_{2,1,k}} \cdot\cdot \\ &\sum_{\beta_{2,2,k}} \cdot\cdot \sum_{\beta_{2,3,k}} \frac{\exp\left(-\frac{F(\psi_{2,4}) \mid \psi_{2,4} = A_m y_{2,k+4}}{2N_0} \right) P(\beta_{2,4,k-1} = A_m)}{\sum_{i=1}^{M} \exp\left(-\frac{F(\psi_{2,4}) \mid \psi_{2,4} = A_i y_{2,k+4}}{2N_0} \right) P(\beta_{2,4,k-1} = A_i)} \times V_{2,4} \end{aligned} \tag{5.24}$$

for the M possible amplitude levels and select as a value for the coefficient being investigated the amplitude level having the largest probability. The joint probability $V_{2,4}$ in (5.24) is given by:

$$V_{2,4} = \left(\prod_{s=1}^{2} \prod_{t=1}^{4} P\left(\alpha_{s,t,k} = A_{1...M}\right) \right) \cdot \left(\underbrace{\prod_{i=1}^{2} \prod_{n=1}^{4}}_{i=2, n \neq 4} P\left(\beta_{i,n,k} = A_{1...M}\right) \right) \tag{5.25}$$

All the probabilities required to evaluate the above joint probability are those corresponding to the APPs that were previously computed during the process of optimizing the other coefficients. This is a refinement of our prior assumption that all amplitude levels are equally probable for each of the channel coefficient.

Thus, the estimate of $\beta_{2,4,k}$, denoted by $\widetilde{\beta}_{2,4,k}$, is

$$\widetilde{\beta}_{2,4,k} = \arg\left\{ \max_{A_m} P(\beta_{2,4,k} = A_m \mid r_{1,k}) \right\} \tag{5.26}$$

Therefore, an initial estimate for each coefficient would have been selected at this point. The concept is to pass the reliability of the decisions made in one iteration to serve as a refinement of the a priori values for the next iteration and repeat this process several times to produce better decisions.

5.5 Conclusions

A new MAP based channel estimation technique amenable to STC schemes was unveiled in this chapter. The effectiveness of most STC schemes relies on accurate multi-channel estimation technique at the receiver in order to achieve diversity advantage and coding gain. A simple two-branch diversity scheme operating in band-limited channels with ISI was discussed. An equivalent discrete-time model of MIMO channel was developed.

The channel was parameterized in terms of complex-valued path gains with unknown deterministic quantities. The channel parameters are the attenuation and delay incurred by the signal transversal along the propagation paths. The channel was assumed to be of finite length, thus the maximum number of paths was considered known a priori. Assuming the channel varies very slowly for the duration of the training sequence transmission, the parameters remained constant during that time.

Discrete realization for the channel coefficients were used even though the original channel model is continuous in practice, this enabled us to apply the proposed MAP approach. Unlike classic estimation techniques, the presented technique iterated on the different probabilities of different coefficients rather than on the coefficient values themselves. This new MAP approach inherently used the principle of turbo decoding, which utilizes both the a priori probability and the APPs.

The APPs expressions required to optimize the various coefficients were derived. For each coefficient, the derived APPs were computed for M possible amplitude levels and the level having the largest probability was selected as an estimate of the coefficient being optimized. Initially since no prior knowledge was available, we assumed all amplitude levels were equally probable for each of the coefficients. As we proceeded in the estimation process, the newly computed APPs were used instead. This was seen as a refinement of our prior assumption that all amplitude levels are equally probable.

The concept was to pass the reliability of the decisions made in one coefficient to serve as a refinement of the a priori values for the next coefficient. The proposed technique required successive evaluation of M probabilities for each coefficient, which involved intensive computational operations. In particular, the averaging performed over all coefficients (except the one being currently optimized) if the number M of the amplitude levels was large. We have computed and ran a few simulation results for the new algorithm presented in this chapter, however, the mean square estimation errors were not optimized. Therefore, low complexity algorithms are proposed next to alleviate computational complexity.

Chapter 6

Low Complexity MAP Based Channel Estimation Algorithms for MIMO Systems

6.1 Introduction

In most channels of practical interest, an intensive computation is prohibitively expensive to implement. To alleviate computational operations of the proposed technique, two simplified methods are devised instead. Two practical approaches, denoted by the first algorithm and the second algorithm, are proposed to implement the derived general expressions to estimate the actual channel coefficients in this Chapter.

In the first algorithm, we reduce the number of multiple summations used to perform averaging over all coefficients except the one being currently selected for optimization in the current pass. The optimization of the first coefficient would still require averaging over the still unknown probabilities of all possible amplitude levels of all other coefficients that are related to the coefficient being optimized. Then choose as a coefficient value the amplitude level with the largest probability.

In estimating the following channel coefficients, we use the selected value of the previously estimated coefficients and set their corresponding probabilities to one. Thus, averaging over those previously estimated channel coefficients are no longer needed. This reduces the number of summations as we proceed, until we reach to the last coefficients, which will be optimized using simple expression.

In the second algorithm, a further simplification is achieved by randomly selecting for each channels coefficient except for the first coefficient, one of the assumed amplitude levels as a coefficient value. Consequently, averaging over all possible amplitude levels of the channel coefficients except the one being currently estimated that is performed by multiple summations is no longer needed. Thus, each coefficient will be optimized using a simplified expression to compute the APPs.

To improve the reliability of the initial estimates of the coefficients, an iterative procedure is used for the two algorithms. To ensure the convergence of the iterative procedure an update process is used for the two algorithms. The performances of the two algorithms are assessed by simulations. Combined analysis and simulation results are presented and compared against those of conventional channel estimation techniques. Results show that improvement over conventional techniques can be obtained using the proposed iterative algorithms.

6.2 First Algorithm Structure

In the first algorithm for computing equations like (5.17), (5.21), (5.24), and further optimization of the various channel coefficients, the goal is to reduce the number of multiple summations used to perform averaging over all coefficients except the one being optimized. A close look at the derived general expressions in the previous chapter (e.g. (5.17), (5.12), and (5.24)) shows that to optimize a particular $\alpha_{j,l,k}$ coefficient, we only need the values of the remaining α_s coefficients. The same applies to the β_s coefficients.

Hence, we are able to subdivide the channel coefficients into two separate groups with one group containing the α_s coefficients and the second group containing the β_s coefficients. Suppose it is desired to estimate the coefficient $\alpha_{j,l}$ in the kth symbol interval, and $r_{1,k}$ be the observed received signal. Similar to the approach in previous chapter, we compute the APPs, which can be now expressed as

$$P(\alpha_{j,l,k} = A_m \mid r_{1,k}) = \underbrace{\sum \cdots \sum}_{other\ \alpha_S} \frac{\exp\left(-\dfrac{R_m + I_m \middle| \alpha_{j,l,k} = A_m}{2N_0}\right) P\left(\alpha_{j,l,k-1} = A_m\right)}{\sum_{i=1}^{M} \exp\left(-\dfrac{R_i + I_i \middle| \alpha_{j,l,k} = A_i}{2N_0}\right) P\left(\alpha_{j,l,k-1} = A_i\right)} \times Z'_{j,l} \tag{6.1}$$

for the M possible amplitude levels and choose as a coefficient value the amplitude level with the largest probability. The functions R_m and I_m were given in (5.12) and the joint probability $Z'_{j,l}$ is given by

$$Z'_{j,l} = \underbrace{\prod_{i=1}^{2}\prod_{n=1}^{4}}_{If\ \ i=j, then\ \ n\neq l} P(\alpha_{i,n,k} = A_{1\ldots M}) \tag{6.2}$$

The notation $A_{1\ldots M}$ is used to represent the amplitude level being assumed by the corresponding coefficient as dictated by the index of the corresponding summation.

6.2.1 First Algorithm Process

To illustrate the estimation process of the first algorithm, we start by investigating a particular coefficient $\alpha_{1,1,k}$ based on the received signal $r_{1,k}$. From the general relation given in (6.1), the APPs required to optimize coefficient $\alpha_{1,1,k}$ are given by

$$P(\alpha_{1,1,k} = A_m \mid r_{1,k}) = \sum_{\alpha_{1,2,k}} .. \sum_{\alpha_{1,3,k}} .. \sum_{\alpha_{1,4,k}} .. \sum_{\alpha_{2,1,k}} .. \sum_{\alpha_{2,2,k}} .. \sum_{\alpha_{2,3,k}} ..$$
$$\sum_{\alpha_{2,4,k}} \frac{\exp\left(-\dfrac{F(\varphi_{1,1}) \mid \varphi_{1,1} = A_m x_{1,k+1}}{2N_0}\right) P\left(\alpha_{1,1,k-1} = A_m\right)}{\sum_{i=1}^{M} \exp\left(-\dfrac{F(\varphi_{1,1}) \mid \varphi_{1,1} = A_i x_{1,k+1}}{2N_0}\right) P(\alpha_{1,1,k-1} = A_i)} \times Z'_{1,1} \tag{6.3}$$

where the function $F(\varphi_{1,1})$ was given in (5.16). Equation (6.3) will be computed for the M possible amplitude levels being assumed for the channel coefficient. Based on the MAP criterion, the amplitude value having the largest probability will be selected as an estimate of $\alpha_{1,1,k}$ coefficient.

As a result of examining a particular received signal, we compute the APPs, which can be thought of as a “refinement” of our prior knowledge. Those APPs will serve as a priori values for the next iteration. Computing (6.3) requires the knowledge of the a priori probability $P(\alpha_{1,1,k-1} = A_m)$ for $m = 1, 2,\ldots, M$. For the first iteration, we assume that the M possible amplitude values are equally probable a priori, i.e., $P(\alpha_{1,1,k-1} = A_m) = 1/M$ for all M levels. The joint probability $Z'_{1,1}$ is given as

$$Z'_{1,1} = \underbrace{\prod_{i=1}^{2}\prod_{n=1}^{4}}_{i=1,n\neq 1} P\left(\alpha_{i,n,k} = A_{1\ldots M}\right) \tag{6.4}$$

Since no prior knowledge is available, we assume all amplitude levels are equally probable. Thus, the probabilities required to compute the above joint probability are set equal to $1/M$, for all M possible amplitude levels for each of the coefficients.

Consider next the optimization of the coefficient $\alpha_{1,2,k}$, while fixing a value for the first coefficient that corresponds to the level with MAP, and assigning a probability of one to that level. Thus, averaging over the coefficient $\alpha_{1,1,k}$ is no longer needed and the number of summations is reduced by one. Hence, the APPs to optimize coefficient $\alpha_{1,2,k}$ becomes as:

$$P(\alpha_{1,2,k} = A_m \mid r_{1,k}) = \sum_{\alpha_{1,3,k}}\cdot\cdot\sum_{\alpha_{1,4,k}}\cdot\cdot\sum_{\alpha_{2,1,k}}\cdot\cdot\sum_{\alpha_{2,2,k}}\cdot\cdot\sum_{\alpha_{2,3,k}}\cdot\cdot\sum_{\alpha_{2,4,k}} \frac{\exp\left(-\dfrac{F(\varphi_{1,2}) \mid \varphi_{1,2} = A_m x_{1,k+2}}{2N_0}\right)P\left(\alpha_{1,2,k-1} = A_m\right)}{\sum_{i=1}^{M}\exp\left(-\dfrac{F(\varphi_{1,2}) \mid \varphi_{1,2} = A_i x_{1,k+2}}{2N_0}\right)P(\alpha_{1,2,k-1} = A_i)} \times Z'_{1,2} \tag{6.5}$$

where $F\left(\varphi_{1,2}\right)$ was given in (5.20) with the computed estimate of the coefficient $\alpha_{1,1,k}$ ($\tilde{\alpha}_{1,1,k}$) used in the evaluation of (5.20).

Again, (6.5) needs to be computed for the M possible amplitude levels of the coefficient $\alpha_{1,2,k}$. The required a priori probabilities $P(\alpha_{1,2,k-1} = A_m)$ for computing (6.5) will be set equal to $1/M$ i.e. equally probable for all M levels. A decision based on the MAP criterion will be used to select an estimate of the second coefficient $\alpha_{1,2,k}$. The selected value will correspond to the level of maximum probability. The joint probability $Z_{1,2}^{'}$ in (6.5) is given as:

$$Z_{1,2}^{'} = \underbrace{\prod_{i=1}^{2}\prod_{n=1}^{4}}_{i=1,n\neq 1,2} P(\alpha_{i,n,k} = A_{1...M}) \tag{6.6}$$

In computing the above joint probability, the involved probabilities for the coefficient $\alpha_{1,1,k}$, $P(\alpha_{1,1,k} = A_{1...M})$ corresponding to the APPs previously computed in (6.5). This is a refinement of our a priori values, while the other required probabilities are set equal to $1/M$, for all M possible amplitude levels for each of the other coefficients.

The same procedure will be followed to derive the APPs relations required to optimize the remaining coefficients in this group. Note that with each successive coefficient estimation the number of summations is reduced by one. Such that when we reach to the optimizing of the last coefficient in this group, which is $\alpha_{2,4,k}$, the required APPs relations becomes as

$$P(\alpha_{2,4,k} = A_m \mid r_{1,k}) = \frac{\exp\left(-\frac{F(\varphi_{2,4}) \mid \varphi_{2,4} = A_m x_{2,k+4}}{2N_0}\right) P\left(\alpha_{2,4,k-1} = A_m\right)}{\sum_{i=1}^{M} \exp\left(-\frac{F(\varphi_{2,4}) \mid \varphi_{2,4} = A_i x_{2,k+4}}{2N_0}\right) P\left(\alpha_{2,4,k-1} = A_i\right)} \times 1 \tag{6.7}$$

where $F(\varphi_{2,4})$ is given by

$$F(\varphi_{2,4}) = 2\varphi_{2,4}^2 + 4\varphi_{2,4}(\varphi_{1,1} + \varphi_{1,2} + \varphi_{1,3} + \varphi_{1,4} + \varphi_{2,1} + \varphi_{2,2} + \varphi_{2,3}) - 2\varphi_{2,4}(\theta_{k+L} + \sigma_{k+L}) \tag{6.8}$$

In evaluating the above function, we substitute the estimate of the previously optimized coefficients, i.e. $\alpha_s = \tilde{\alpha}_s$ except for the one being currently estimated that will assume the value set in the left hand side of (6.7). Note that the joint probability is equal to one as seen in (6.7), since no averaging is required at this point.

In a similar fashion, we can derive the relations required to compute the APPs for the second group containing the β_s coefficients. We start with the first coefficient $\beta_{1,1,k}$; the relations to compute it's APPs will be similar to those derived for $\alpha_{1,1,k}$. For the last coefficient in this group $\beta_{2,4,k}$, the relations will be similar to those derived for $\alpha_{2,4,k}$ in (6.7) with function $F(\psi_{2,4})$ that was given in (5.23) instead of $F(\varphi_{2,4})$.

6.2.2 The Iterative Procedure

At the end of the first iteration and based on (6.3), (6.5), and (6.7), we would have computed the APPs for all coefficients and selected based on the MAP criterion an initial estimate for each coefficient of the channels. In order to increase the reliability of the initial estimates, we need to continue the estimation process of the coefficients to ensure convergence of the coefficients to their optimum values. This is accomplished by means of an iterative procedure to adjust the coefficients estimates.

Unlike classic estimation techniques, iteration will take place on the probabilities of the coefficients rather than on the coefficients values. The concept is to pass the reliability of the decisions made in one iteration to serve as a refinement of the a priori values for the next iteration and repeat the process several times to produce sufficiently reliable decisions. The computed APPs in the current iteration will be adjusted and fed as a priori probabilities for the next iteration.

A scale factor Δ is used to control the rate of adjustment, Δ is a positive number chosen small enough to ensure rapid convergence of the iterative procedure. Two update processes will be utilized to achieve the adjustment and to ensure the convergence of the coefficients to their optimum values. In the first update process, the scale factor Δ is set equal to a fixed positive number to adjust all the APPs for all coefficients of the channels.

Assuming that in the current iteration the MAP was found at level i (i =1, …, M) for coefficient $\alpha_{j,l,k}$ (j = 1, 2, l = 1, 2, 3, 4). The a priori probabilities for the coefficient $\alpha_{j,l,k}$ in the next iteration are obtained as follows:

$$\begin{aligned} P(\alpha_{j,l,k-1} = A_i)\,\Big|_{\substack{Next\\ iteration}} &= MAP(\alpha_{j,l,k} = A_i)\,\Big|_{\substack{Current\\ iteration}} + \Delta \\ P(\alpha_{j,l,k-1} = A_m)\,\Big|_{\substack{Next\\ iteration}} &= P(\alpha_{j,l,k} = A_m)\,\Big|_{\substack{Current\\ iteration}} - \Delta/(M-1), \qquad m = 1,2,\ldots,M, m \neq i \end{aligned} \tag{6.9}$$

If a priori probability becomes negative it is set to zero, and if a priori probability becomes greater that one it is set to one. Then normalization is preformed, such that the sum of the a priori probabilities for the M possible amplitude levels of each coefficient is equal to one. The same update is preformed for the β_s coefficients.

Alternatively, the second update process utilizes a dynamic Δ to update the APPs, such that the APPs for each coefficient is adjusted using a particular positive number that might be different from those used in updating the APPs for the other coefficients. Assuming that in the current iteration the MAP of coefficients $\alpha_{j,l,k}$ and $\beta_{j,l,k}$ were found at level i (i =1, ..., M) for j = 1, 2 and l = 1, 2, 3, 4. Then, we compute the sum of the entire MAP, denoted by $SMAP$ as follows:

$$SMAP = \sum_{j=1}^{2} \sum_{l=1}^{4} \left[MAP(\alpha_{j,l,k} = A_i) + MAP(\beta_{j,l,k} = A_i) \right] \tag{6.10}$$

The dynamic Δ denoted by $D\Delta$ is obtained according to the relation

$$\begin{aligned} D\Delta(\alpha_{j,l,k}) &= \{[1 - MAP(\alpha_{j,l,k} = A_i)]\}/SMAP \\ D\Delta(\beta_{j,l,k}) &= \{[1 - MAP(\beta_{j,l,k} = A_i)]\}/SMAP \end{aligned} \tag{6.11}$$

for j = 1, 2 and l = 1, 2, 3, 4. Then, the APPs for $\alpha_{j,l,k}$ (j = 1, 2, l = 1, 2, 3, 4) are updated as follows:

$$\begin{aligned} P(\alpha_{j,l,k-1} = A_i)\,|_{\substack{Next\\iteration}} &= MAP(\alpha_{j,l,k} = A_i)\,|_{\substack{Current\\iteration}} + D\Delta(\alpha_{j,l,k}) \\ P(\alpha_{j,l,k-1} = A_m)\,|_{\substack{Next\\iteration}} &= P(\alpha_{j,l,k} = A_m)\,|_{\substack{Current\\iteration}} - D\Delta(\alpha_{j,l,k})/(M-1), \end{aligned} \tag{6.12}$$

for $m = 1,2,\ldots,M, m \neq i$. If a priori probability becomes negative, it is set to zero, and if a priori probability becomes greater, that one it is set to one. Then normalization is performed, such that the sum of the a priori probabilities for the M possible amplitude levels of each coefficient is equal to one. The updated APPs will serve as a priori values for the next iteration. The same update is performed for the β_s coefficients.

In the second and the following iterations, we again start the process by optimizing the coefficient $\alpha_{1,1,k}$. The required APPs are given by

$$P(\alpha_{1,1,k}=A_m \mid r_{1,k})=\frac{\exp\left(-\frac{F(\varphi_{1,1})\mid \varphi_{1,1}=A_m x_{1,k+1}}{2N_0}\right)P\left(\alpha_{1,1,k-1}=A_m\right)}{\sum_{i=1}^{M}\exp\left(-\frac{F(\varphi_{1,1})\mid \varphi_{1,1}=A_i x_{1,k+1}}{2N_0}\right)P(\alpha_{1,1,k-1}=A_i)} \tag{6.13}$$

where the function $F(\varphi_{1,1})$ was given in (5.16), which will be evaluated using the estimate of the coefficients found in the previous iteration, expect for the coefficient being currently optimized. Equation (6.13) will be computed for the M possible amplitude levels and select an estimate of $\alpha_{1,1,k}$ coefficient the amplitude value having the largest probability.

Similarly, the APPs to optimize the coefficient $\alpha_{1,2,k}$ can be expressed as

$$P(\alpha_{1,2,k}=A_m \mid r_{1,k})=\frac{\exp\left(-\frac{F(\varphi_{1,2})\mid \varphi_{1,2}=A_m x_{1,k+2}}{2N_0}\right)P\left(\alpha_{1,2,k-1}=A_m\right)}{\sum_{i=1}^{M}\exp\left(-\frac{F(\varphi_{1,2})\mid \varphi_{1,2}=A_i x_{1,k+2}}{2N_0}\right)P(\alpha_{1,2,k-1}=A_i)} \tag{6.14}$$

where $F(\varphi_{1,2})$ was given in (5.20), which will be evaluated using the estimate of the coefficient $\alpha_{1,1,k}$ ($\tilde{\alpha}_{1,1,k}$) obtained in current iteration. For the other coefficients, we substitute the optimum values from the previous iteration. Similar relations could be derived for the remaining coefficients in this group.

It is straightforward to repeat the same procedure for the second group that contains the β_s coefficients. Therefore, the structure of the first algorithm encompasses two separate estimation process, one for the α_s coefficients and another for the β_s coefficients. The iterative procedure of the first algorithm proceeds as follows:

1. In the first iteration, set the a priori probabilities = $1/M$, for all M possible amplitude levels and for all coefficients.
2. Estimate the first coefficient using (6.3) and select an optimum value based on MAP criterion.
3. Optimize the second coefficient using (6.5) and choose an optimum value.
4. Repeat step 3 until the last coefficient in this group is estimated utilizing (6.7).
5. Proceed to optimize the second group that contains the β_s coefficients and repeat steps 2 to 4.
6. Update the computed APPs using either (6.9) or (6.12); set them as a priori probability for the coefficients in the next iteration.
7. Start the next iteration by optimizing the first and the second coefficients using (6.13) and (6.14) respectively.
8. Estimate the remaining coefficients until last one in this group is optimized.
9. Repeat steps 7 and 8 for the second group.
10. Repeat steps 6 to 9 enough iterations to yield a reliable decision (certain mean square estimation error of channel coefficients, as will be defined shortly in Section 6.3.1).

6.2.3 Second Algorithm

In order to lessen the computational complexity of the first algorithm and to produce an algorithm that is more practical to implement, we further propose a second algorithm for estimating the actual coefficients of the channels. In this algorithm, again the coefficients are subdivided into two groups, each group contains 8 coefficients. We start the procedure by randomly selecting for each coefficient a value chosen from the range of amplitude levels being assumed, except for the first coefficients in each of the two groups for which we always begin the estimation process.

Thus, we utilize (6.13) to compute the APPs required to optimize the coefficient $\alpha_{1,1,k}$ in the first iteration. The a priori probabilities would still be equally probable, i.e. $P(\alpha_{1,1,k-1} = A_k) = 1/M$ for all M. To estimate the second coefficient, we utilize (6.14) and in computing $F(\varphi_{1,2})$, we set $\alpha_{1,1,k}$ equal to its estimate obtained in the current iteration. For other coefficients, we substitute the values that were randomly selected at the beginning of the algorithm and fixing other coefficients at their most recently estimated values in the current iteration.

This process should be continued until we reach to the last coefficient in this group, which can be optimized using equation (6.7). In evaluating $F(\varphi_{2,4})$ given in (6.7), we substitute the optimized values of each coefficient, which were previously obtained in the current iteration. It is straightforward to repeat the same procedure for the second group that contains the β_s coefficients. The iterative procedure of the second algorithm proceeds as follows:

1. Set the a priori probabilities = $1/M$, for all M possible levels and for all coefficients.

2. Randomly select for each coefficient a value from the set of possible amplitude levels being assumed.

3. Estimate the first and second coefficients using (6.13) and (6.14) respectively.

4. Proceed to optimize the next coefficient, until we optimize the last coefficient in this group.

5. Repeat steps 3 and 4 for the second group of coefficients.

6. Update the obtained APPs using either (6.9) or (6.12) and feed them as a priori probability for the coefficients in the next iteration.

7. Repeat steps 3 to 6 enough number of iterations to yield reliable decision (certain mean square estimation error of channel coefficients, as will be defined shortly in Section 6.3.1).

6.3 Analysis and Simulation Results

Simulations were conducted in order to examine the performance of the two proposed algorithms. Simulations results are presented and compared against the performance of conventional estimation techniques in this section.

6.3.1 Implementation Aspects

It is assumed that each transmit antenna radiates half the energy to ensure the same radiated power as that of the classical no diversity (ND) system. A fixed-length training sequence was sent from each of the two transmit antennas simultaneously. The transmitted sequences $\{d_{1,k}\}$ and $\{d_{2,k}\}$ are complex-valued elements, their real and imaginary components $\in \{1,-1\}$. An observation interval of one hundred training symbols is chosen ($K = 100$).

Without loss of generality, we have set each channel to four paths and the channels were assumed static. Meaning that the channel coefficients are either fixed or vary so slowly that they remain constant over the observation interval. The actual values assigned to each of the channel coefficients were randomly selected from a set of M possible amplitudes levels. The amplitude levels $\{A_m\}$ take discrete values, which are equally spaced around zero.

Throughout the simulation, the carrier-to-noise ratio (CNR) is used as variable input parameter to the simulation. The CNR is defined as

$$CNR(dB) = 10\log_{10}\frac{|d_{1,k}|^2 + |d_{2,k}|^2}{2N_0} \tag{6.15}$$

for a required CNR value, the corresponding noise variance was computed. Two independent blocks of Gaussian random noise with zero mean and computed variance are generated and added separately to the received signals components. The received signals are then obtained in terms of their real and imaginary parts using (5.7).

Since the most meaningful measure of performance for a digital communications system is the average probability of errors, it is desirable to choose the coefficients estimate to minimize this performance index. However, the probability of error is a highly nonlinear function of the channels coefficients. Consequently, the probability of error as a performance index for channel estimation is impractical. A performance criterion that measures the quality of the estimation and leads to practical implementations is the mean square error (MSE). Often this criterion results in optimum or near optimum performance compared to other measures of performance. In the case of estimating a signal received in white Gaussian noise, minimization of the MSE is equivalent to maximum likelihood estimation, which is optimum from a probability of error viewpoint. Thus, we consider the MSE as a criterion to evaluate the two proposed algorithms performance. The normalized MSE is defined as

$$MSE = \sum_{i=1}^{16}\frac{|C_i - \widetilde{C}_i|}{C_i^2} \tag{6.16}$$

where C_i is the actual coefficient value, selected by the randomization process mentioned before, and $\widetilde{C}_i$ is the estimate of that coefficient [22].

To avoid fluctuation in the computed MSE from one iteration to the next, we run a number of trials for each specified CNR. Simulations were repeated enough times for both algorithms and results were averaged (at each iteration) to yield a confidence interval of 90% and an error of 5% (not to be confused with our MSE above).

6.3.2 First Algorithm Performance Evaluation

The first algorithm was implemented using four-level ($M = 4$). Thus, each actual channel coefficient was randomly assigned a value from the set {-0.3, -0.1, 0.1, 0.3}. The value of the fixed scale factor Δ was chosen to be 0.1. Several other values were used, but $\Delta = 0.1$ produced the most stable and smooth convergence of the iterative process compatible with other input parameters such as CNR.

The average MSE of the first algorithm with a fixed Δ of 0.1 is displayed in Fig. 6.1 and Fig. 6.2. The average MSE versus the number of iterations with a variable CNR of 0, 2, 4, and 5 dB is shown in Fig. 6.1. In Fig. 6.2, the average MSE is depicted as a function of CNR for various numbers of iterations. From these plots, we can see that the MSE decreases with the increase in the number of iterations, as well as with the increment in CNR. Thus, the estimated coefficients values converge towards the actual values in a smooth fashion and a satisfactory performance can be achieved with only a few iterations. Possible trade-off can be considered between the number of iterations and the transmitted power. The results were obtained by averaging the MSE in 50 simulations, which were found enough to satisfy a confidence interval of 90% with an error of 5% (85% to 95%).

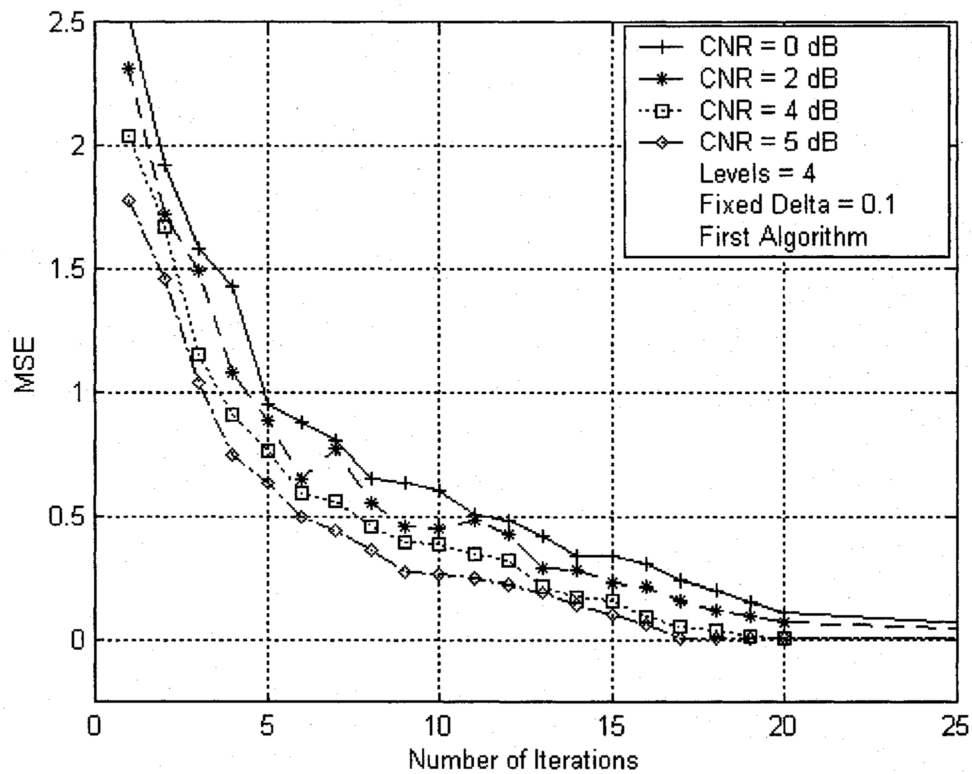

Figure 6.1: The Average MSE of the first algorithm versus the number of iterations with a variable CNR of 0, 2, 4, and 5 dB for fixed Δ of 0.1.

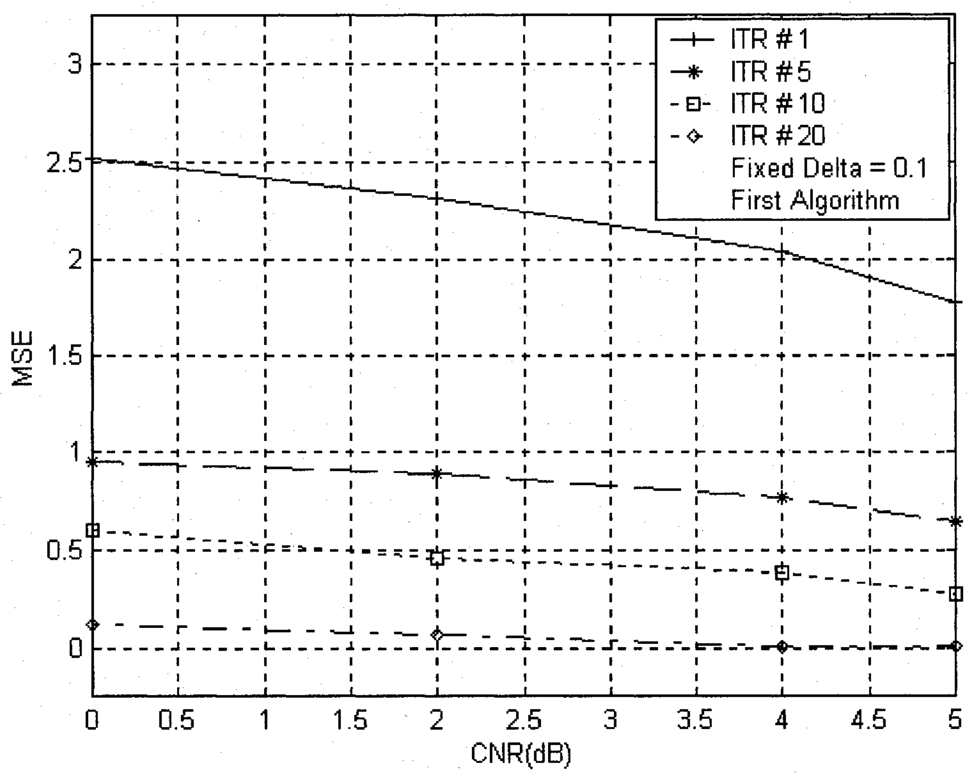

Figure 6.2: Average MSE of the first algorithm as a function of CNR with various numbers of iterations for fixed Δ of 0.1.

Figure 6.3 illustrates the average MSE of the first algorithm with a dynamic Δ against the number of iterations for CNR of 0, 2, 4, and 5 dB. The average MSE seen in Fig. 6.3 exhibits a generally identical convergence behavior to that of the fixed Δ case, as shown in Fig. 6.1. Thus, we can achieve a satisfactory and a desirable convergence rate of the channel coefficients towards their optimum values that is minimum MSE. Possible trade-off can be considered between the number of iterations and the transmitted power. The results were obtained by averaging the MSE in 50 simulations, which were found enough to satisfy a confidence interval of 90% with an error of 5% (85% to 95%).

The average MSE performance comparison of the first algorithm between the cases of fixed and dynamic Δ with a variable CNR of 0 and 5 dB are demonstrated in Fig. 6.4. The performance curves for the two update processes are basically the same as the number of iterations and CNR are increased. It is clear that with a single fixed scale factor or with a slightly more complex adjustable scale factor the algorithm results are almost the same. It requires the same number of iterations in both cases for the algorithm to converge to a low MSE value and achieve a good performance.

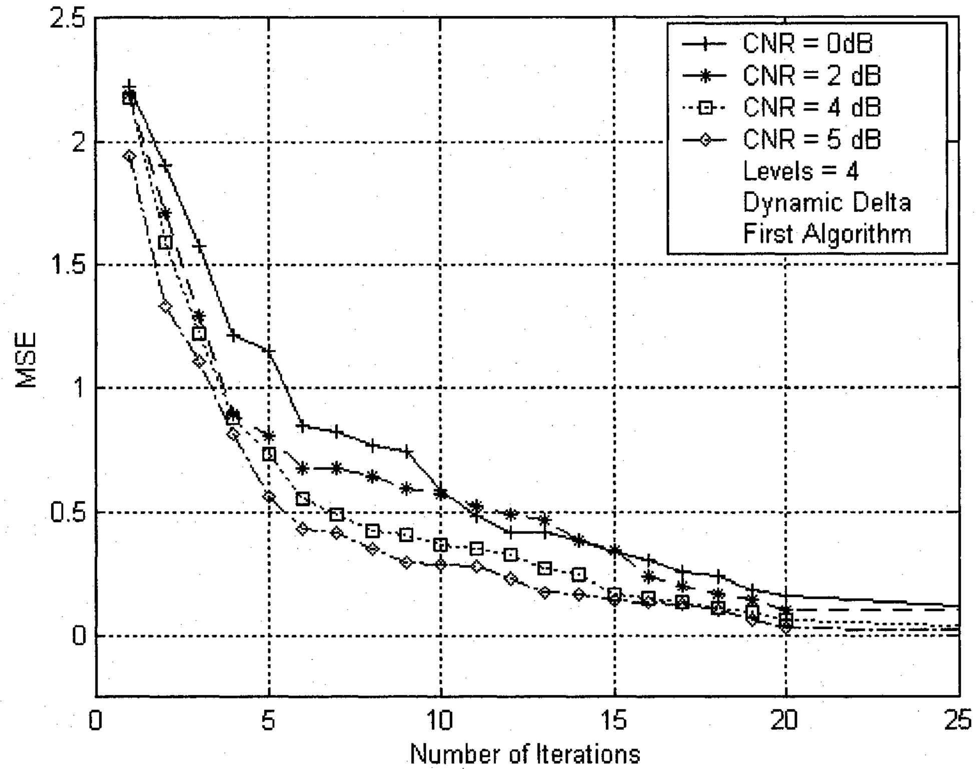

Figure 6.3: Average MSE of the first algorithm against the number of iterations with a CNR of 0, 2, 4, and 5 dB for dynamic Δ.

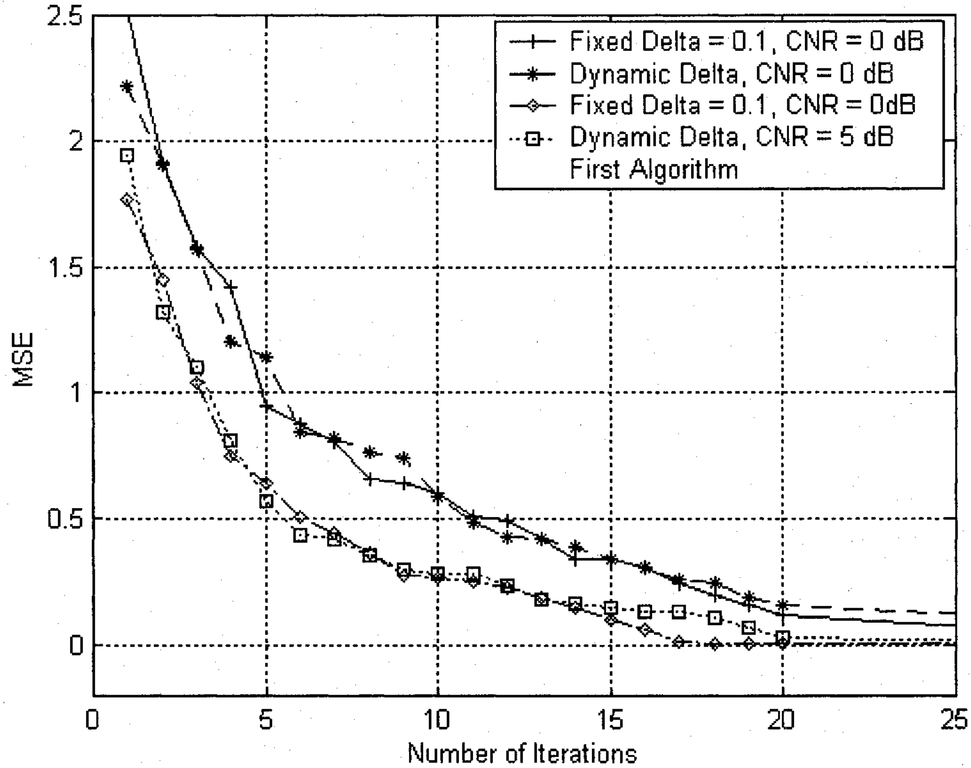

Figure 6.4: Average MSE performance comparison of the first algorithm with fixed and dynamic Δ for CNR values of 0 and 5 dB.

6.3.3 Second Algorithm Performance Results

The second algorithm was implemented using eight-level (M = 8) of estimation. Thus, each actual channel coefficient was randomly assigned a value from the set of amplitude levels {-0.35, -0.25, -0.15, -0.05, 0.05, 0.15, 0.25, 0.35}. The value of the fixed scale factor Δ was chosen to be 0.1. Several other values were used, but $\Delta = 0.1$ produced the most stable and smooth convergence of the iterative process compatible with other input parameters such as CNR.

The average MSE of the second algorithm with a fixed Δ of 0.1 is displayed in Fig. 6.5 and Fig. 6.6. The average MSE versus the number of iterations with a variable CNR of 0, 2, 4, and 5 dB is shown in Fig. 6.5. In Fig. 6.6, the average MSE is depicted as a function of CNR for various numbers of iterations. From these figures, we can see that the MSE decreases with the increase in the number of iterations, as well as with the increment in CNR.

Thus, the estimated coefficients values converge towards the actual values in a relatively fast fashion and a satisfactory performance can be achieved with only few iterations. Possible trade-off can be considered between the number of iterations and the transmitted power. The results were obtained by averaging the MSE in 100 simulations, which were found enough to satisfy a confidence interval of 90% with an error of 5% (85% to 95%).

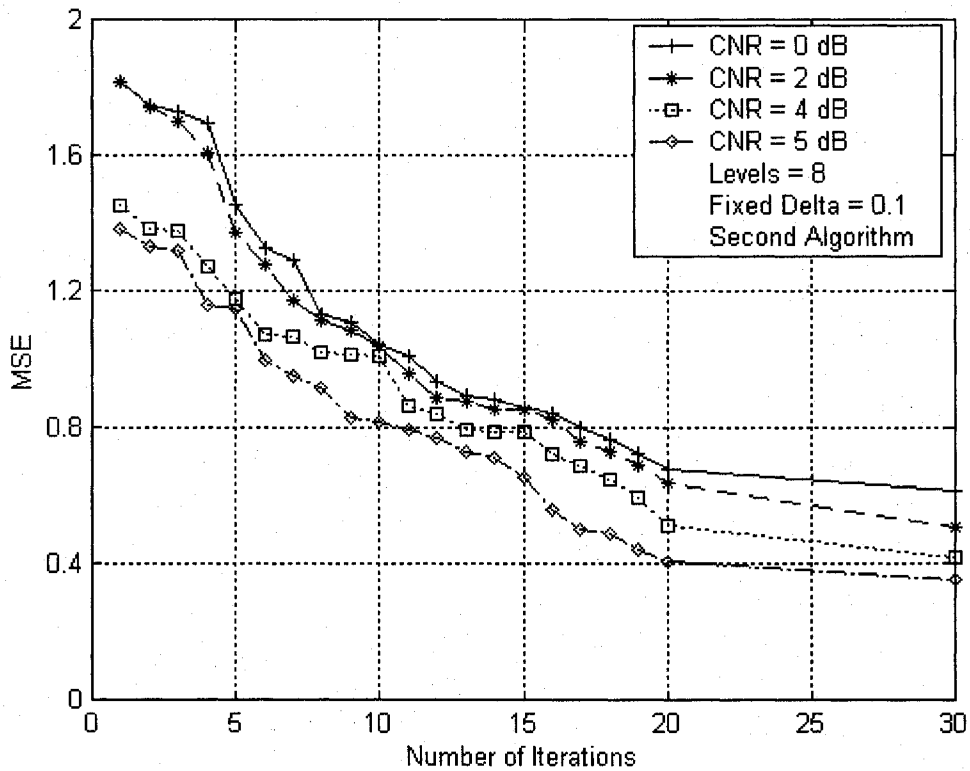

Figure 6.5: The Average MSE of the second algorithm versus the number of iterations with a variable CNR of 0, 2, 4, and 5 dB for fixed Δ of 0.1.

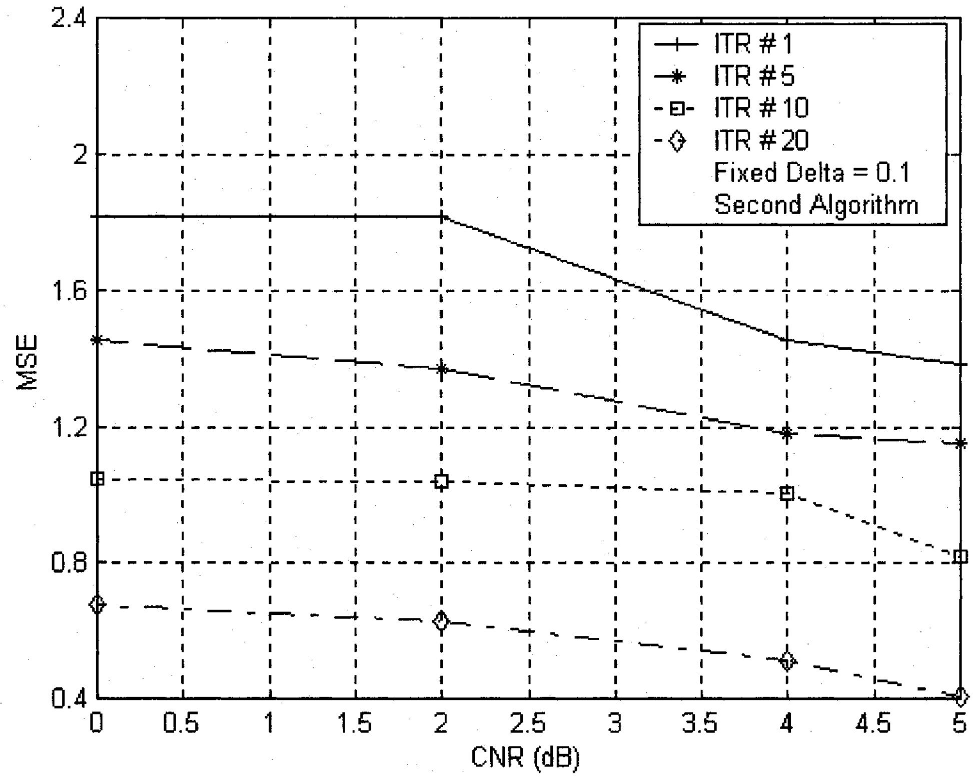

Figure 6.6: Average MSE of the second algorithm as a function of CNR with various numbers of iterations for fixed Δ of 0.1.

Figure 6.7 illustrates the average MSE of the second algorithm with a dynamic Δ against the number of iterations for CNR of 0, 2, 4, and 5 dB. The average MSE seen in Fig. 6.7 exhibits a generally identical convergence behavior to that of the fixed Δ case, as shown in Fig. 6.5. Thus, we can achieve a satisfactory and a desirable convergence rate of the channel coefficients towards their optimum values that is minimum MSE. Possible trade-off can be considered between the number of iterations and the transmitted power. The results were obtained by averaging the MSE in 100 simulations, which were found enough to satisfy a confidence interval of 90% with an error of 5% (85% to 95%).

The average MSE performance comparison of the second algorithm between the cases of fixed and dynamic Δ with a variable CNR of 0 and 5 dB are demonstrated in Fig. 6.8. The performance curves for the two update processes are basically the same as the number of iterations and CNR are increased. It is clear that with a single fixed scale factor or with a slightly more complex adjustable scale factor the algorithm results are almost the same. It requires the same number of iterations in both cases for the algorithm to converge to a low MSE value and achieve a good performance.

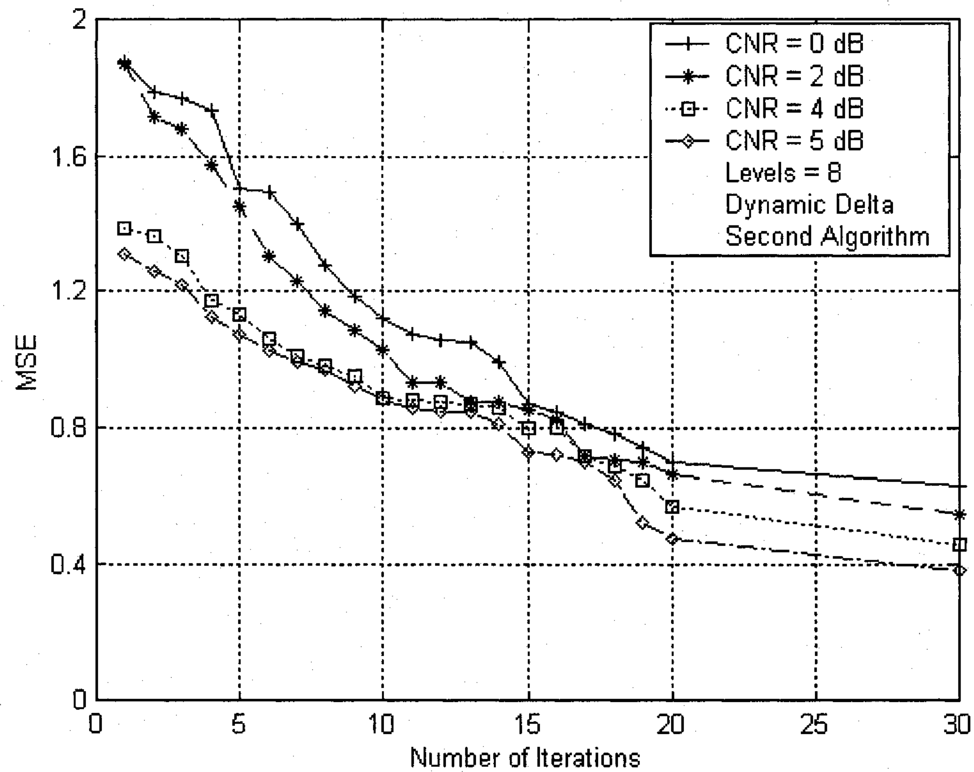

Figure 6.7: Average MSE of the second algorithm against the number of iterations with a variable CNR of 0, 2, 4, and 5 dB for dynamic Δ.

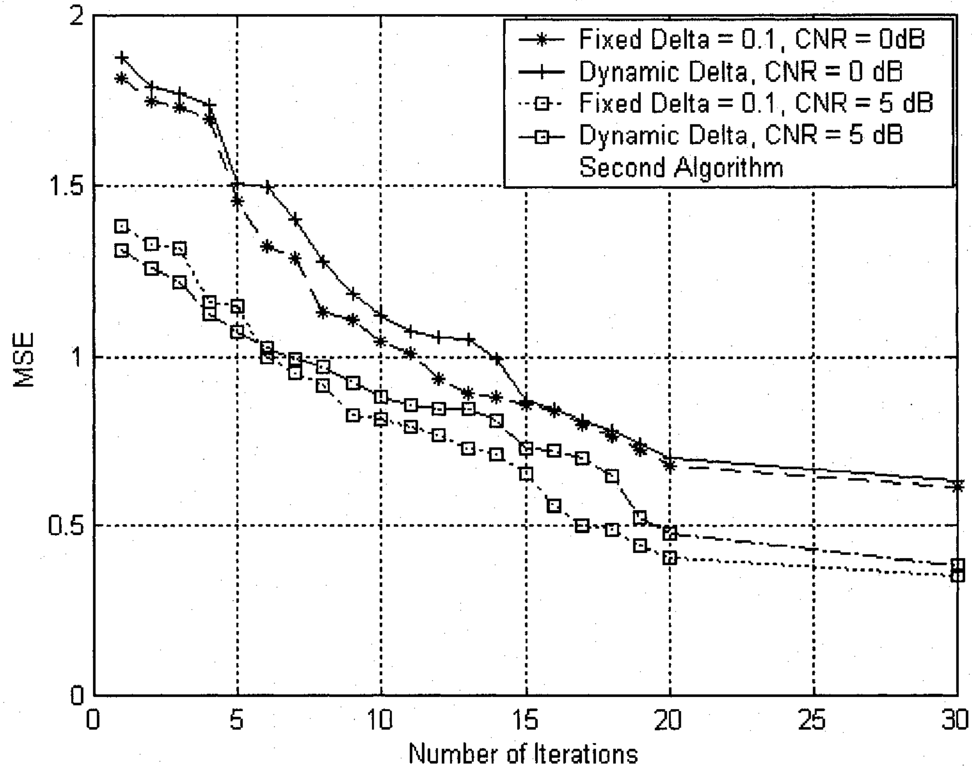

Figure 6.8: Average MSE performance comparison of the second algorithm with fixed and dynamic Δ for CNR values of 0 and 5 dB.

6.3.4 Performance Comparison of the two Algorithms and with Conventional Estimation Techniques

Comparison curves between the MSE results of the proposed algorithms with a variable CNR of 0 and 5 dB are presented in Fig. 6.9 for fixed Δ of 0.1. The first algorithm converges much faster (fewer iterations) towards the actual values of the channels coefficients. The second algorithm worsens the average MSE performance slightly, but the computational complexity decreases by a factor of at least 20. This reduction factor is because no averaging was performed in the second algorithm.

We conclude by comparing these results against the performance indexes of the least mean square (LMS) [102] and recursive least square (RLS) [103] algorithms. The LMS algorithm is a steepest-descent algorithm in, which the true gradient vector is approximated by an estimate obtained directly from the data [23], [104]. In order to achieve faster convergence, RLS (Kalman) algorithm that deals directly with the received data in minimizing the cumulative square error is used [105].

Figure 6.10 depicts the results of our two algorithms against the performance curves of the LMS and RLS techniques for the single antenna case provided in [77], [22]. Comparison shows that we have obtained a faster rate of convergence and produced an acceptable low and stable MSE values with fewer number of iterations, even with the challenging task of estimating increased number of parameters for our two transmit antennas case. Our first algorithm results in a comparatively good performance relative to that obtained with the LMS algorithm.

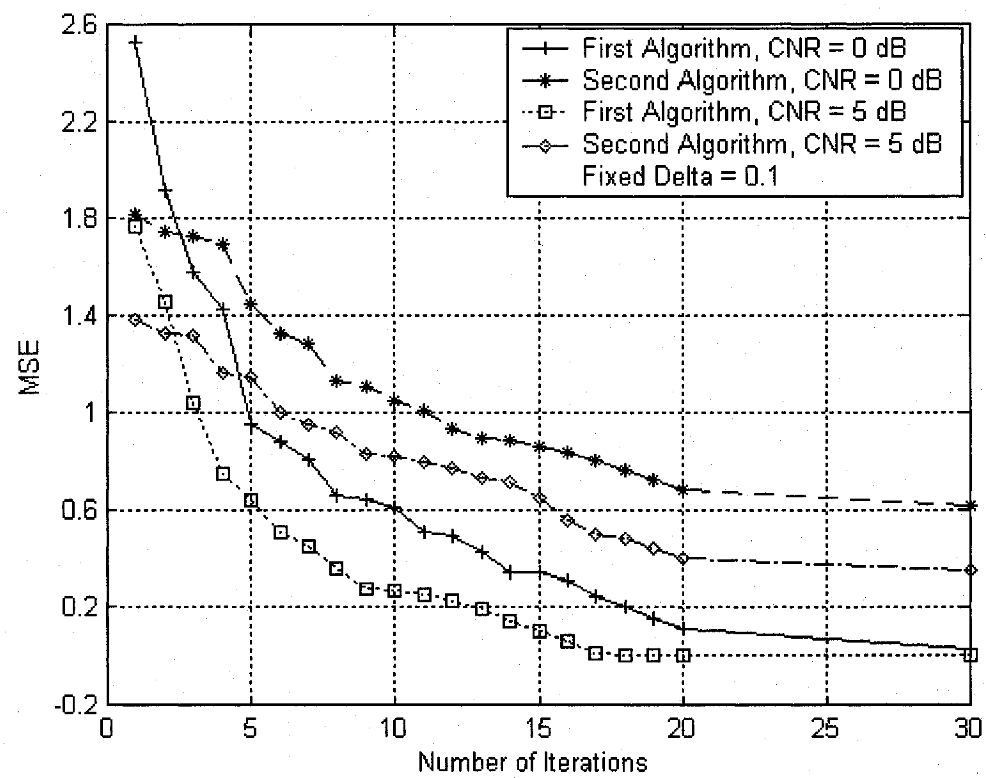

Figure 6.9: Performance comparison between the two algorithms with a variable CNR of 0 and 5 dB for a fixed Δ of 0.1.

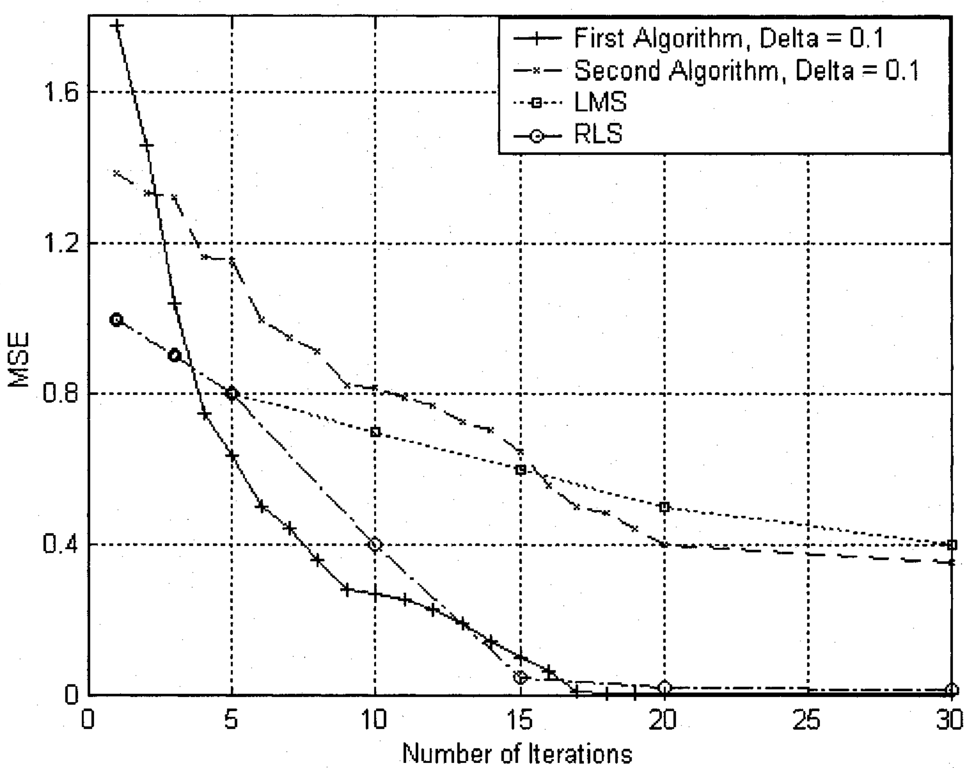

Figure 6.10: MSE Performance comparison of the two proposed algorithms against the LMS and RLS estimation techniques.

6.4 Conclusions

The analytical expressions derived in the previous chapter to evaluate the APPs for each coefficient involved a large amount of computation. Two low complexity algorithms were introduced instead in the Chapter. The two proposed algorithms differ in the way the first iteration is performed and the expressions used in computation. In the first algorithm, we reduced the number of multiple summations as we proceeded in the estimation process from one coefficient to the next. Thus, the last coefficients were evaluated using a simple multiplication formula.

The computational operations in the first iteration were further reduced in the second algorithm. We started the algorithm by randomly choosing for each coefficient a value selected from the set of amplitude levels being assumed. Thus, the second algorithm significantly reduces the complexity resulting in a simplified channel and practical channel estimation technique. To improve the reliability of the initial estimates of the coefficients, an iterative procedure was used for the two algorithms.

The iteration took place on the probabilities rather than the coefficients values. The computed APPs in one iteration were updated and fed as a priori values for the iteration. A scale factor with fixed and dynamic values was used in the update procedure for the two algorithms. Simulations were conducted for the two algorithms and the results were examined and investigated for the various cases. The performances of the two proposed algorithms were also compared against the performance of LMS and RLS conventional estimation techniques.

Chapter 7

Conclusions and Future Works

7.1 Conclusions

Multiple-Input Multiple-Output (MIMO) systems that use multiple transmit and receive antennas to provide high data rates and achieve significant gain in channel capacity emerged rapidly as the new frontier of wireless communications. Band-limited (frequency selective) channels are narrow pipes that do not accommodate rapid flow of data. Deploying multiple antennas at both transmitter and receiver broadened this data pipe by exploiting the spatial dimension.

In this dissertation, the effects of imperfect multipath channel estimation on the performance of MIMO diversity system employing STC was investigated. Then a new technique to estimate the channel parameters for a two branch transmit diversity scheme was derived. In addition, two practical low complexity algorithms for MIMO channels that are characterized as frequency selective fading were proposed.

An effective and practical way to attain the promised increase in capacity of MIMO channels was to employ STC technique performed in both spatial and temporal domains. As a result, numerous MIMO schemes were developed that included STTC and STBC to achieve high spectral efficiencies and performance gains. However, the main problem of all these schemes was that they were originally designed and later analyzed assuming known flat fading channels. In the real world, MIMO channels often undergo frequency selective fading that causes ISI in the received signal. In addition, the availability of accurate channel estimation at the receiver is not always justified. In practice, it is impossible to achieve the perfect channel estimation especially for frequency selective MIMO channels because a large number of channel parameters have to be estimated at reduced transmit power for each transmit antenna.

A simple transmit diversity (STD) scheme utilizing STC and using two transmit and multiple receive antennas had received recently a great deal of research attention. The scheme appeared as a simple way to achieve the same diversity gain as MRC scheme with low decoding complexity. However, the effectiveness of the proposed scheme and most STC schemes relied on the assumption of known flat fading channel. In reality, this will not be the case if the channel experience multipath fading and non-negligible delay spread causing ISI in the received signal. Moreover, the diversity benefits cannot be achieved unless the channel is accurately estimated at the receiver. Hence, we conducted an analysis and comparison of the error performance of the STD and MRC schemes with multipath channel estimation errors. Closed form expressions for the BER of the two schemes were derived and the BER curves were plotted.

Performance curves showed that the STD scheme is significantly more susceptible to errors than the MRC scheme. The deterioration in the STD performance increased rapidly relative to the MRC performance with $n = 4$ and 5% estimation error. Thus, the practical implementation of the STD system should be carefully considered in channels with large number of multipath components and channel estimation error exceeding 5% at the receiver.

Our results showed the deleterious effects of inaccurate channel estimation on the performance of MIMO systems and in particular the proposed STD scheme. Channel estimation for MIMO systems is a major challenge and requires additional effort for frequency selective MIMO channels because of the large number of channel parameters to be estimated. Hence, the development of novel and effective channel estimation techniques that can accurately estimate a large number of channel parameter were required for MIMO systems utilizing STC.

We unveiled a new MAP based channel estimation technique amenable for two-branch transmit diversity scheme employing STC. The multipath channel between each transmit and receive antenna pair was characterized as frequency selective that results in ISI. An equivalent discrete-time model of MIMO channel with ISI was developed. The channel parameters are the attenuation and delay incurred by the signal transversal along the propagation paths. Assuming the channel varies very slowly for the duration of the training sequence transmission, the parameters remained constant during that time. Discrete realization for the channel coefficients were used even though the original channel model is continuous in practice, this enabled us to apply the proposed MAP approach. Unlike classic estimation techniques, the presented technique iterated on the probabilities of channel coefficients, rather than on the values of the coefficients. This new MAP approach inherently used the principle of turbo decoding, which utilizes both the a priori probability and the APPs. The APPs expressions required to optimize the various coefficients were derived. For each coefficient, the derived APPs were computed for M possible amplitude levels and the level having the largest probability was selected as an estimate of the coefficient being optimized. The proposed technique required successive evaluation of M probabilities for each coefficient, which involved intensive computational operations. In particular, the averaging performed over all coefficients (except the one being currently optimized) if the number M of the amplitude levels was large.

To alleviate computational complexity, two low complexity algorithms were introduced instead. In the first algorithm, we reduced the number of multiple summations as we proceeded in the estimation process from one coefficient to the next. Thus, the last coefficient was evaluated using a simple multiplication formula. The computational operations in the first iteration were further reduced in the second algorithm. We started the algorithm by randomly choosing for each coefficient a value selected from the set of amplitude levels being assumed. To improve the reliability of the initial estimates of the coefficients, an iterative procedure was used for the two algorithms. Simulations were conducted for the two algorithms and the results were examined and compared against the performance of LMS and RLS conventional estimation techniques.

7.2 Future Research Directions

We believe the estimation algorithms proposed in this thesis only scratch the tip of the iceberg and thus many issues are still to be explored. As a continuation of the work in this dissertation, the following research topics can be identified:

1. *Max-Log-MAP Channel Estimation*

We proposed a new MAP based channel estimation technique for two-branch transmit diversity scheme in chapter 4. The APPs expressions required to optimize the various channel coefficients were also derived. The next step would be to extend our proposed technique to an arbitrary number of transmit antennas. Then, by taking the logarithm of the expressions derived to compute the APPs for each channel coefficient, the computational operations would be reduced to a simple logarithmic addition. Thus, a reduced complexity channel estimation technique based on max-log-MAP approach for MIMO systems would be developed.

2. *Simplified Channel Estimation Algorithms*

We introduced two low complexity algorithms that were derived from the general technique to implement the computed expressions to optimize the channel coefficients in chapter 6. The first and second algorithms were implemented using four and eight amplitude levels, respectively. Following our approach, practical and low complexity algorithms could be developed for max-log-MAP based channel estimation technique. Then, the developed algorithms could be implemented using a larger number of amplitude levels for the estimation of channel coefficients.

3. *Joint Channel and data detection Method*

An approach that is receiving increasing interest is the investigation of joint channel estimation and data detection methods. Where data decision obtained from decoding is used as additional training to refine the channel estimate. Combining joint data detection and channel estimation with an iterative algorithm is a method that is gaining popularity. Iterative channel estimation for MIMO systems use strategies that were described above could be an interesting research topic to pursue.

4. *Blind Channel Estimation*

Blind channel estimation methods are of great interest because they avoid training sequence and thus make efficient use of the available bandwidth. The challenge is to develop an algorithm for blind channel estimation of MIMO systems utilizing STC and operating in multipath environment. The algorithm based on the max-log-MAP criterion described above with a focus on complexity issues could be pursued.

References

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

[2] R. Prasad, *Universal Wireless Personal Communications*, Boston: Artech house, 1998.

[3] M. Shafi, S.Ogose, and T. Hattori, *Wireless Communications in the 21st Century*, New York: Wiley, 2002.

[4] E. Telatar, "Capacity of multi-antenna Gaussian channels," *AT&T Bell Labs Technical Memorandum*, June 1995. Also, *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585-595, December 1999.

[5] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311-335, March 1998.

[6] A. Paulraj, R. Naber, and D. Gore, *Introduction to Space-Time Wireless Communications*, Cambridge: Cambridge University Press, 2003.

[7] J. Ventura-Traveset, G. Caire, E. Biglieri, and G. Taricco, "Impact of diversity reception on fading channels with coded modulation-Part I: Coherent detection," *IEEE Transactions on Communications*, vol. 45, no. 5, pp. 563-572, May 1997.

[8] J. Ventura-Traveset, G. Caire, E. Biglieri, and G. Taricco, "Impact of diversity reception on fading channels with coded modulation-Part II: Differential block detection," *IEEE Transactions on Communications*, vol. 45, no. 6, pp. 676-696, June 1997.

[9] J. Ventura-Traveset, G. Caire, E. Biglieri, and G. Taricco, "Impact of diversity reception on fading channels with coded modulation-Part III: Co-Channel interferences," *IEEE Transactions on Communications*, vol. 45, no. 7, pp. 809-818, July 1997.

[10] A. Witteneben, "A new bandwidth efficient transmit antenna modulation diversity scheme for linear digital modulation," *Proc. IEEE International Conference on Communication (ICC)*, vol. 3, pp. 1630-1634, Geneva, Switzerland, 23-26 May, 1993.

[11] N. Seshadri and J. H. Winters, "Two schemes for improving the performance of frequency-division duplex (FDD) transmission systems using transmitter antenna diversity," *International Journal of Wireless Information Networks*, vol. 1, no. 1, pp. 49-60, January 1994.

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

[13] Z. Chen, B. S. Vucette, J. Yuan, and L. Leong, "Space-time trellis codes for 4-PSK with three and four transmit antennas in quasi-static flat fading channels,"

IEEE Communications Letters, vol. 6, no. 2, pp. 67-69, February 2002.

[14] B. Abdool-Rassool, F. Heliot, L. Revelly, R. Nakhai, and H. Aghvami, "4-PSK space-time trellis codes with five and six transmit antennas for slow Rayleigh fading channels," *IEE Electronics Letters*, vol. 39, no.3, pp. 297-299, February 2003.

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

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

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

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

[19] Y. Liu, M. P. Fitz, and O. Y. Takeshita, "Full rate space-time turbo codes," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 5, pp. 969-890, May 2001.

[20] C. Dongzhe and A. Haimovich, "Performance of parallel concatenated space-time codes," *IEEE Communications Letters*, vol. 5, no. 6, pp. 236-238, June

2001.

[21] W. Firmanto, B. Vucetic, J. Yuan, and Z. Chen, "Space-time turbo trellis coded modulation for wireless data communications, *EURASIP Journal on Applied Signal Processing*, vol. 2002, no. 5, pp. 459-470, May 2002.

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

[23] S. Qureshi, "Adaptive Equalization," *Proceedings of IEEE*, vol. 73, no. 9, pp. 1349-1387, September 1985.

[24] P. K. Shukla and L. F. Turner, "Channel-estimation-based adaptive DFE for fading multipath radio channels," *IEE Proceeding I Communications, Speech and Vision*, vol. 138, no.6, pp. 525-543, December 1991.

[25] R. A. Ziegler, M. W. Al-Dhahir, and J. M. Cioffi, "Nonrecursive adaptive DFE from channel estimates," *Proc. IEEE Vehicular Technology Conference (VTC)*, vol. 2, pp. 600-603, Denver, CO, 10-13 May, 1992.

[26] D. Dzung, "Error probability of MLSE equalization using imperfect channel measurements," *Proc. IEEE International Conference on Communications (ICC)*, vol. 2, pp. 558-562, Denver, CO, 23-26 June, 1991.

[27] N. Seshadri, "Joint data and channel estimation using blind trellis search techniques," *IEEE Transactions on Communications*, vol. 42, no. 234, pp. 1000-1011, February-April 1994.

[28] M. Stojanovic, J. G. Proakis, and J. A. Catipovic, "Analysis of the impact of channel estimation errors on the performance of a decision-feedback equalizer in fading multipath channels," *IEEE Transactions on Communications*, vol. 43, no.

234, pp. 877-886, February-April 1995.

[29] A. Gorokhov, "On the performance of the viterbi equalizer in the presence of channel estimation errors," *IEEE Signal Processing Letters*, vol. 5, no. 12, pp. 321-324, December 1998.

[30] L. Yumin and C. Jyh-Hau, "The effect of channel estimator memory mismatch on the performance of MLSE in wireless data communications," *Proc. IEEE International Conference on Communications (ICC)*, vol. 1, pp. 134-138, New York, NY, 28 April-2 May, 2002.

[31] A. F. Naguib, V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time coding modem for high-data-rate wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1459-1478, August 1998.

[32] V. Tarokh, A. F. Naguib, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communications: Performance criteria in the presence of channel estimation errors, mobility, and multipath paths," *IEEE Transactions on Communications*, vol. 47, no. 2, pp. 199-207, February 1999.

[33] L. Ling, Y. Yu-Dong, and L. Hongbin, "Channel estimation and equalization for space-time block coded systems in frequency selective fading channels," *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 1, pp. 300-304, San Antonio, TX, 25-29 November, 2001.

[34] A. Grant, "Joint decoding and channel estimation for linear MIMO channels," *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, vol. 3, pp. 1009-1012, Chicago, IL, 23-28 September, 2000.

[35] J. Boutros, F. Boixadera, and C. Lamy, "Bit-interleaved coded modulations for multiple-input multiple-output channels," *Proc. IEEE 6th International Symposium on Spread Spectrum Techniques and Applications (ISSSTA)*, vol. 1, pp. 123-126, Parsippany, NJ, 6-8 September, 2000.

[36] Y. Li, C. N. Georghiades, and G. Huang, "Iterative maximum-likelihood sequence estimation for space-time coded systems," *IEEE Transactions on Communications*, vol. 49, no. 6, pp. 948-951, June 2001.

[37] E. Chiavaccini and G. M. Vitetta, "MAP symbol estimation on frequency flat Rayleigh fading channels via a Bayesian EM algorithm," *IEEE Transactions on Communications*, vol. 49, no. 11, pp. 1869-1872, November 2001.

[38] Z. B. Baranski, A. M. Haimovich, and J. Garcia-Frias, "EM-based iterative receiver for space-time coded modulation with noise covariance estimation," *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 1, pp. 355-359, Taipei, Taiwan, 17-21 November, 2002.

[39] C. Cozzo, and B. L. Hughes, "Joint channel estimation and data detection in space-time communications, *IEEE Transactions on Communications*, vol. 51, no. 8, pp. 1266-1270, August 2003.

[40] V. Jungnickel, T. Haustein, E. Jorswieck, V. Pohl, and C. von Helmolt, "Performance of a MIMO systems with overlay pilots," *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 1, pp. 594-598, San Antonio, TX, 25-29 November, 2001.

[41] M. C. Valenti and B. D. Woerner, " Iterative channel estimation and decoding of

pilot symbol assisted turbo codes over flat-fading channels," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 9, pp. 1697-1705, September 2001.

[42] H. –J. Su and E. Geraniotis, "Low-complexity joint iterative channel estimation and decoding for pilot symbol-assisted modulation and multiple differential detection systems with correlated Rayleigh fading," *IEEE Transactions on Communications*, vol. 50, no. 2, pp. 249-261, February 2002.

[43] Z. Haidong, B. Farhang-Boroujeng, and C. Schlegel, "Pilot embedding for joint channel estimation and data detection in MIMO communication systems," *IEEE Communications Letters*, vol. 7, no. 1, pp. 30-32, January 2003.

[44] X. Deng, A. M. Haimovich, and J. Garcia-Frias, "Decision directed iterative channel estimation for MIMO systems," *Proc. IEEE International Conference on Communications (ICC)*, vol. 4, pp. 2326-2329, Anchorage, AL, 11-15 May, 2003.

[45] E. Onggosanusi, A. Gatherer, A. Dabak, and S. Hosur, "Performance analysis of closed-loop transmit diversity in presence of feedback delay," *IEEE Transactions on Communications*, vol. 49, no. 9, pp. 1618-1630, September 2001.

[46] P. Dighe, R. Mallik, and S. Jamuar, "Analysis of transmit-receive diversity in Rayleigh fading," *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 2, pp. 1132-1136, San Antonio, TX, 25-29 November, 2001.

[47] J. Choi, "Performance analysis for transmit antenna diversity with/without channel information, *IEEE Transactions on Vehicular Technology*, vol. 51, no. 1, pp. 101-113, pp. 767-771, pp. 771-774, January 2002.

[48] T. Marzetta, "BLAST training: estimation channel characteristics for high-capacity space-time wireless," *Proc. 37th Annual Allerton Conference on Communications, Control and Computing*, pp. 958-966, Monticello, IL, September 1999.

[49] B. Hassibi and B. Hochwald, "Optimal training in space-time systems," *Proc. IEEE 34th Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 743-747, Pacific Grove, CA, October/November 2000.

[50] L. Tong, "Channel estimation for space-time orthogonal block codes," *Proc. IEEE International Conference on Communications (ICC)*, vol. 4, pp. 1127-1131, Helsinki, Finland, June 2001.

[51] W. C. Jakes, Ed., *Microwave Mobile Communications*, New York: Wiley, 1974.

[52] M. D. Yacoub, *Foundations of Mobile Radio Engineering*, Boca Raton, FL: CRC Press, 1993.

[53] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Transactions on Communications*, vol. 11, no. 4, pp. 360-393, December 1963.

[54] H. Suzuki, "A statistical model for urban radio propagation," *IEEE Transactions on communications*, vol. 25, no. 7, pp. 673-680, July 1977.

[55] G. L. Stuber, *Principles of Mobile Communications*, New York: IEEE Press, 1995.

[56] P. A. Bello and B. D. Nelin, "The effects of frequency selective fading on the binary error probabilities of incoherent and differentially coherent matched filter receivers," *IEEE Transactions on Communication Systems*, vol. 11, no. 2, pp. 170-186, June 1963.

[57] T. S. Rappaport, *Wireless Communications: Principle and Practice*, Upper Saddle River, New Jersey: Prentice Hall PTR, 1996.

[58] B. Sklar, *Digital Communications: Fundamentals and Applications*, 2nd ed., Upper Saddle River, New Jersey: Prentice Hall PTR, 2000.

[59] G. L. Turin, "Introduction to spread spectrum antimultipath techniques and their application to urban digital radio," *Proceedings of the IEEE*, vol. 68, no. 3, pp. 328-353, March 1980.

[60] S. H. Van Wanbeck and A. H. Ross, "Performance of diversity receiving systems," *Proceedings of IRE*, vol. 39, pp. 256-264, 1951.

[61] L. R. Kahn, "Ratio squarer," *Proceedings of IRE*, vol. 42, pp. 1704, Nov. 1954.

[62] M. K. Simon and M. S. Alouini, "A unified approach to the performance analysis of digital communications over generalized fading channels," *Proceedings of the IEEE*, vol. 86, no. 9, pp.1860-1877, September 1998.

[63] N. Balaban and J. Salz, "Dual diversity combining and equalization in digital cellular mobile radio," *IEEE Transactions on Vehicular Technology*, vol. 40, no. 2, pp. 342-354, May 1991.

[64] J. Winters, J. Salz, and R. D. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Transactions on*

Communications, vol. 42, nos. 2/3/4, pp. 1740-1751, Feb./Mar./ Apr. 1994.

[65] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41-59, Autumn 1996.

[66] J.-C. Guey, M. P. Fitz, M. R. Bell, and W.-Y. Kuo, "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels," *IEEE Transactions on Communications*, vol. 47, no. 4, pp. 527-537, April 1999.

[67] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*, Cambridge: Cambridge University Press, 2003.

[68] C. Shannon, "A mathematical theory of communication," *Bell Labs Technical Journal*, vol. 27, pp. 379-423, pp. 623-656, July and October 1948.

[69] A. Lapidoth and S. Moser, "On the fading number of multi-antenna systems over flat fading channels with memory and incomplete side information," *Proc. IEEE International Symposium on Information Theory*, p. 478, Lausanne, Switzerland, 30 June-5 July, 2002.

[70] R. Etkin and D. Tse, "Degrees of freedom in underspread MIMO fading channels," *Proc. IEEE International Symposium on Information Theory*, p. 323, Yokohama, Japan, 29 June-4 July, 2003.

[71] T. Koch and A. Lapidoth, "The fading number and degrees of freedom in non-coherent MIMO fading channels: A peace pipe," *Proc. IEEE International Symposium on Information Theory*, pp. 661-665, Adelaide, Australia, 4-9

September, 2005.

[72] L. Zheng and D. N. Tse, "Diversity and multiplexing: A fundamental trade-off in multiple antenna channels," *IEEE Transactions on Information Theory*, vol. 49, no. 5, May 2003.

[73] H. Yao and G. Wornell, "Structured space-time block codes with optimal diversity-multiplexing trade-off and minimum delay," *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 4, pp. 1941-1945, San Francisco, CA, 1-5 December, 2003.

[74] H. Gamal, G. Caire, and M. Damon, "Lattice coding and decoding achieve the optimal diversity-multiplexing trade-off of MIMO channels," *IEEE Transactions on Information Theory*, vol. 50, no. 6, June 2004.

[75] A. Goldsmith, *Wireless Communications*, New York: Cambridge University Press, 2005.

[76] H. W. Sorenson, "Least-squares estimation: from Gauss to Kalman," *IEEE Spectrum*, pp. 63-68, July 1970.

[77] A. Giordano and F. Hsu, *Least Square Estimation with Application to Digital Signal Processing*, New York: John Wiley and Sons, 1985.

[78] X. Feng and C. Leung, "Performance sensitivity comparison of two diversity schemes," *IEE Electronics Letters*, vol. 36, no. 9, pp. 838-839, April 2000.

[79] J. Zhou, H. Kikuchi, S. Sasaki, S. Muramatsu, and Y. Onozato, "Performance Investigation of Two Transmit Diversity Schemes with Perfect/Imperfect Channel Evaluation in Wireless Communications," *IEICE Transaction on*

Communications, vol. E85-B, no. 10, pp. 1957-1965, October 2002.

[80] D. Gu and C. Leung, "Performance analysis of transmit diversity scheme with imperfect channel estimation," *IEE Electronics Letters*, vol. 39, no. 4, pp. 402-403, February 2003.

[81] W.C. Ang and C.S. Ng, "Analysis of a Transmit Diversity Scheme in Fading Channels and Imperfect Channel Estimation," *Proc. IEEE International Conference on Communication Systems* (*ICCS*), vol. 1, pp. 188-192, Singapore, 25-28 November 2002.

[82] C. Shan, P.Y. Kam, and A. Nallanathan, "On the performance of space-time block coded systems with channel estimation," *Proc. IEEE Vehicular Technology Conference* (*VTC*), vol. 4, pp. 2439-2443, Los Angeles, CA, 26-29 September 2004.

[83] W. Hoteit, Y.R. Shayan, and A.K. Elhakeem, "Effects of imperfect channel estimation on STC performance," *IEE Proceedings Communications*, vol. 152, no. 3, pp. 277-281, June 2005.

[84] Y. Akyildiz and B.D. Rao, "Maximum ratio combining performance with imperfect channel estimates," *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing* (*ICASSP*), vol.3, pp. 2485-2488, Orlando, FL, 13-17 May 2002.

[85] G. Bauch and N. Al-Dhahir, "Reduced-Complexity Space-Time Turbo-Equalization for Frequency-Selective MIMO Channels," *IEEE Transaction on Wireless Communications*, vol. 1, no. 4, pp. 819-829, October 2002.

[86] A. Vielmon, Ye Li, and J. R. Barry, "Performance of transmit diversity over time-varying Rayleigh fading channels," *Proc. IEEE Global Telecommunications Conference* (*GLOBECOM*), vol. 5, pp. 3242-3246, San Antonio, TX, 25-29 November, 2001.

[87] N. Al-Dhahir, C. Fragouli, A. Stamoulis, W. Younis, and A. R. Calderbank, " Space-time processing for broadband wireless access," *IEEE Communication Magazine*, vol. 40, no. 9, pp. 136-142, September 2002.

[88] A. J. Paulraj and B. Chong Ng, "Space-time modems for wireless personal communications," *IEEE Personal Communications*, vol. 5, no. 1, pp. 36-48, February 1998.

[89] Z. Liu, X. Ma, and G. B. Giannakis, "Space-time coding and Kalman filter for time selective fading channels," *IEEE Transactions on Communications*, vol. 50, n0. 2, pp. 183-186, February 2002.

[90] E. Lindskog and A. Paulraj, " A transmit diversity scheme for channels with intersymbol interference," *Proc. IEEE International Conference on Communications* (*ICC*), vol. 1, pp. 307-311, New Orleans, LA, 18-22 June, 2000.

[91] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, New Jersey: Prentice Hall PTR, 1993.

[92] J. J. Van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Borjesson, "On channel estimation in OFDM systems," *Proc. IEEE Vehicular Technology Conference* (*VTC*), vol. 2, pp. 815-819, Chicago, IL, 25-28 July, 1995.

[93] Ye Li, L. J. Cimini, and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid Dispersive Fading Channels," *IEEE Transactions on Communications*, vol. 46, no. 7, pp. 902-915, July 1998.

[94] Ye Li, N. Seshadri, S. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 3, pp. 461-471, March 1999.

[95] Ye (Geoffrey) Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Transactions on Wireless Communications*, vol. 1, no. 1, pp. 67-75, January 2002.

[96] M. Cedervall and A. Paulraj, "Joint channel and space-time parameter estimation," *Proc. IEEE 30th Asilomar Conference On Signals, Systems and Computer*, vol. 1, pp. 375-379, Pacific Grove, CA, 3-6 November, 1996.

[97] G. G. Raleigh and T. Boros, "Joint space-time parameter estimation for wireless communication channels," *IEEE Transactions on Signal Processing*, vol. 46, no. 5, pp. 1333-1343, May 1998.

[98] M. Yan and B. D. Rao, "Soft decision-directed Map estimate of fast Rayleigh flat fading channels," *IEEE Transactions on Communications*, vol. 51, no. 12, pp. 1965-1969, December 2003.

[99] R. E. Kamel and Y. Bar-Ness, "Blind maximum likelihood sequence estimation of digital sequence in presence of intersymbol interference," *IEE Electronics Letters*, vol. 30, no. 7, pp. 537-539, March 1994.

[100] R. Schober, W. H. Gerstacker, L. H.-J. Lampe, "Performance analysis and

design of STBC'S for fading ISI channels," *Proc. IEEE International Conference on Communications* (*ICC*), vol. 3, pp. 1451-1455, New York, NY, 28 April-02 May, 2002.

[101] A. Narula, M. D. Trott, and G. W. Wornell, "Performance limits of coded diversity methods for transmitter antenna arrays," *IEEE Transactions on Information Theory*, vol. 45, no. 7, pp. 2418-2433, November 1999.

[102] J. G. Proakis and J. H. Miller, "An adaptive receiver for digital signaling through channels with intersymbol interference," *IEEE Transactions on Information Theory*, vol. 15, no. 4, pp. 484-497, July 1969.

[103] D. D. Falconer and L. Ljumg, "Application of fast Kalman estimation to adaptive equalization," *IEEE Transactions on Communications*, vol. 26, no. 10, pp. 1439-1446, October 1978.

[104] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, New Jersey: Prentice Hall PTR, 1985.

[105] F. M. Hsu, "Square root Kalman filtering for high speed data received over fading dispersive HF channels," *IEEE Transactions on Information Theory*, vol. 28, no. 5, September 1982.