DECOUPLED MAXIMUM LIKELIHOOD CHANNEL ESTIMATOR FOR SPACE-TIME BLOCK CODED SYSTEM

SHENG JIANGUO

DEPARTMENT OF ELECTRICAL AND

COMPUTER ENGINEERING

A THESIS SUBMITTED

FOR THE MASTER OF ENGINEERING DEGREE

NATIONAL UNIVERSITY OF SINGAPORE

2003

ACKNOWLEDGEMENT

I wish to express my sincerest thanks to my supervisors, Dr. Arumugam Nallanathan and Professor Tjeng Thiang Tjhung, for the opportunity they provided me to study in the challenging field of channel estimation in STBC system. Their invaluable supports, guidance, encouragements, patience and creative advice throughout my research work are highly appreciated.

TABLE OF CONTENTS

LIST OF TABLE

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

M	number of transmit antennas
N	number of receive antennas
P	frame length
s_{mt}	transmitted signal
x_{nt}	received signal
w_{nt}	additional noise
h_{nm}	fading coefficient
d_i	transmitted symbol
$\tilde{d}_i$	combined transmitted symbol
$\hat{d}_i$	estimate of transmitted symbol
$\mathbf{S}$	transmitted signal matrix
$\mathbf{X}$	received signal matrix
$\mathbf{W}$	additional noise matrix
$\mathbf{H}$	channel coefficient matrix
$\mathbf{Q}$	spatial covariance matrix
$\mathbf{G}$	transmission matrix
s_k	source sequence
y_k	noiseless observation sequence
x_k	observation sequence
w_k	noise sequence
$\delta_{i,j}$	Kronecker delta function

Ψ	finite complex constellation
ρ	the cross-correlation coefficient of the Rayleigh faded envelopes
λ	squared magnitude of the cross-correlation coefficient
$E_i(\eta)$	complete elliptic integral of the second kind with modulus η
δ_x^2	desired signal power
$\mathbf{L}$	coloring matrix
MIMO	multiple input multiple output
STBC	space-time block codes
STTC	space-time trellis codes
LST	layered space-time
USTM	unitary space-time modulation
CSI	channel state information
DEML	decoupled maximum likelihood
ML	maximum likelihood
SS	spatial smoothing
CR	cross relation
LSS	least squares smoothing
ANMSE	asymptotic normalized mean square error
DML	deterministic ML
IQML	iterative quadratic maximum likelihood
TSML	two-step maximum likelihood
SML	statistical ML
EM	expectation-maximization
CRB	Cramer-Rao bound
BER	bit error rate

SNR	signal to noise ratio
BPSK	binary phase shit keying
i.i.d.	independent identically distributed

ABSTRACT

A computationally efficient channel estimation scheme based on the decoupled maximum likelihood (DEML) algorithm is introduced for space-time block coded (STBC) system. The BER performance of the STBC system with the DEML channel estimator is obtained under spatially uncorrelated and correlated flat Rayleigh fading channels. It is shown that the DEML channel estimator could perform well only under uncorrelated fading channels. When the fading channels are correlated, a decorrelation algorithm is applied on the correlated signals before the DEML channel estimator is used. A general procedure on the generation of correlated Rayleigh fading envelops is also introduced in such case. In addition, an iterative ML detector is introduced to improve the system performance with the DEML channel estimator, both under uncorrelated and correlated fading channels.

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

The next generation wireless communication systems are required to carry much higher data rates than those available today. Given a limited radio spectrum, the only way to support high data rates is to develop new spectrally efficient techniques. It has been shown recently that multiple input multiple output (MIMO) systems have great potential to increase the spectral efficiency significantly. MIMO systems can be realized with multi-element array antennas.

Space-time coding has been proposed recently to obtain coded diversity for communication systems with multiple transmit and receive antennas, which combines error control coding and transmit diversity to achieve diversity and coding gains over un-coded systems without expanding system bandwidth. There are various approaches in the literature, including space-time block codes (STBC) [1]–[3], space-time trellis codes (STTC) [4], space-time turbo trellis codes [5] and layered space-time (LST) architectures [6].

STBC, introduced in [1]-[3], is able to achieve full diversity made possible by the large number of transmit and receive antennas. A strong feature of STBC is its simple maximum likelihood decoding algorithm based only on linear receiver processing. The codes are constructed using orthogonal designs and exist only for few sporadic values of the number of transmit antennas. Recently, many new space-time techniques based on STBC have been explored. The differential STBC proposed in [7] has simple differential encoding and decoding algorithms, while the unitary space-time modulation (USTM) proposed in [8] can be applied when the CSI is not known at both the transmit and the receive antennas. However, this approach requires exponential encoding and decoding complexity.

The decoding of space-time codes requires the perfect channel state information (CSI) at the receiver. The space-time decoder will use them to extract symbol estimates. However, in practical scenarios, channel fading coefficients are not always known to transmitter and receiver. In the absence of perfect CSI at the receiver, a channel estimator must be used to estimate the channel coefficients. Then these channel estimates are used as if they were perfectly known at the receiver to extract symbol estimates.

1.2 CONTRIBUTION OF THIS THESIS

In this thesis, we have presented a computationally efficient channel estimation method for STBC system based on the DEML algorithm. The BER performances of the STBC systems with DEML channel estimator are given, both under spatially uncorrelated and correlated flat Rayleigh fading channels. The DEML channel estimator performs well when incident signals are uncorrelated. It can be directly applied to STBC system under spatially uncorrelated fading channel. When the incident signals are correlated, the DEML channel estimator has some performance degradation. Thus for STBC system under spatially correlated fading channels, the correlated signals have to be decorrelated before the DEML channel estimator is applied. A common decorrelation approach used for highly correlated sources is the spatial smoothing (SS) [27] algorithm. This technique resides in dividing the sequence of received signals into sub arrays and summing the estimated spatial correlation matrices obtained from each sub array to form a smoothed correlation matrix. Grenier has brought a significant improvement to the spatial smoothing technique by smoothing the estimated source space instead of the entire space. This approach is called the DEESE algorithm [28] and was later extended to the complexity reduced DEESE algorithm [29] by Grenier.

We have also obtained the BER performance of the STBC system under spatially correlated fading channels. To study the performance of STBC system under correlated fading channels, we have presented a general method on the generation of correlated Rayleigh fading sequences. In this method, independent fading processes with desired autocorrelations are first generated and then multiplied by a coloring matrix. Some selected envelope and phase plots for various correlation coefficients ρ are given and compared. And the BER performance of STBC system with different ρ is also shown and discussed.

In addition, an iterative ML detector is introduced in STBC systems both under the spatially uncorrelated and spatially fading channels to improve the system performance with DEML channel estimator. The iterative ML detector can obtain, after convergence, the performance of the exact ML detector in the case of unknown $\mathbf{H}$ and $\mathbf{Q}$, without significantly increasing computational complexity.

1.3 ORGANIZATION OF THESIS

The outline of the thesis is as follows. In Chapter 2, an overview of space-time coding is given. The space-time coding is based on combining error control coding and transmitter diversity techniques, which can provide spectral efficiency for wireless communications. A specific type of space-time codes, STBC is introduced. In Chapter 3, an overview of channel estimation methods is presented. From the moment-based methods to the ML approaches, we outline the basic ideas behind some new developments. The assumptions, identifiability conditions and their performances are given. The proposed DEML channel estimator is explained in Chapter 4. Its properties are also given in this chapter. In Chapter 5, the BER performance of STBC system with DEML channel estimator under spatially uncorrelated flat Rayleigh fading channels is shown. An iterative ML detector is introduced to improve the system BER performance with DEML channel estimator. In Chapter 6, the BER performance of STBC system with DEML channel estimator under spatially correlated flat Rayleigh fading channel is shown. A general procedure on the generation of correlated Rayleigh fading envelopes and a decorrelation algorithm are developed. Finally, conclusions and future works are given in Chapter 7.

CHAPTER 2

OVERVIEW OF SPACE-TIME CODING

In this chapter, we first introduce a brief background on diversity techniques. Space-time coding is based on combining error control coding and transmitter diversity techniques, which can provide spectral efficiency for wireless communications. The principle, system model, and some approaches of space-time coding are given. Lastly, a specific type of space-time codes, STBC, is introduced.

2.1 DIVERSITY TECHNIQUES

It is well known that significant degradations may occur in the performance of wireless communication system over Rayleigh fading channels. Such degradation in system performance will often requires the signals to be transmitted with an excessive power just to overcome the deleterious fading effects. However, this will cause more cost in design and application.

One method commonly employed to overcome the performance degradation in wireless communication system due to fading is diversity. The goal of diversity is to reduce the fade depth and/or the fade duration by supplying the receiver with multiple replicas of the transmitted signals that have passed over independent fading channels. Given that the channels are independent, the probability that all the channels will fade below a certain threshold at the same time is significantly lower than the probability that one channel fades below the threshold.

Several diversity techniques have been employed in wireless communication systems, including time diversity, frequency diversity, space diversity, and etc.

1) Time Diversity: Channel coding in combination with limited interleaving is used to provide time diversity. However, while channel coding is extremely effective in fast fading environments (high mobility), it offers very little protection under slow fading (low mobility and fixed wireless access) unless significant interleaving delays can be tolerated.

2) Frequency Diversity: The fact that signals transmitted over different frequencies induce different multipath structure and independent fading is exploited to provide frequency diversity (sometimes referred to as path diversity). In TDMA systems, frequency diversity is obtained by the use of equalizers when the multipath delay spread is a significant fraction of a symbol period. Global system for mobile communication (GSM) uses frequency hopping to provide frequency diversity. In DS-CDMA systems, RAKE receivers are used to obtain path diversity. When the multipath delay spread is small, compared to the symbol period, however, frequency or path diversity does not exist.

3) Space Diversity: Space diversity is achieved by using multiple antennas that are separated and/or differently polarized at the transmitter/receiver to create independent fading channels. It can be realized with transmitter diversity and/or receiver diversity. The obvious advantage of transmitter diversity is that the complexity of having multiple antennas is placed on the transmitter. The portable receivers can use just a single antenna and still benefit from the diversity gain.

Different diversity techniques can be combined together. For example, space and time diversity can be combined together by using space-time coding techniques. When possible, cellular systems should be designed to encompass all forms of diversity to ensure adequate performance. However, not all forms of diversity can be available at all times.

2.2 SPACE-TIME CODING

Space-time (ST) coding is based on combining error control coding and transmitter diversity techniques. It is an effective and practical way to approach the capacity of MIMO wireless channels. Coding is performed in both spatial and temporal domain to introduce spatial and temporal correlation into signals transmitted from different antennas and different time periods. The spatial-temporal correlation of the code is used to exploit the MIMO channel fading and to minimize transmission errors at the receiver. By doing so, space-time coding can achieve diversity and coding gain over un-coded systems without sacrificing the bandwidth.

Consider the space-time coded system with M transmit and N receive antennas. Usually it has three functions: encoding and transmitting signals at the transmitter; combining scheme at the receiver and the decision rule for maximum likelihood detection. In the absence of perfect CSI at the receiver, channel estimation should be done at the receiver. In the following, we will briefly introduce the ST transmitter, system transmission model and the ST receiver.

The transmitted data are encoded by a space-time encoder. The encoder chooses the symbols to transmit so that both the coding and the diversity gains at the receiver are maximized. The coded data sequence is applied to a serial-to-parallel (S/P) converter producing parallel data sequence. At each time instant the parallel output are simultaneously transmitted by different antennas. All transmitted signals have the same transmission duration T .

We assume that the frame length is P . An $M \times P$ space-time codeword matrix is obtained by arranging the transmitted sequence in an array as

$$\mathbf{S} = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1P} \\ s_{21} & s_{22} & \cdots & s_{2P} \\ \vdots & \vdots & \ddots & \vdots \\ s_{M1} & s_{M2} & \cdots & s_{MP} \end{bmatrix} \tag{2.2.1}$$

The m^{th} row of $\mathbf{S}$ is the signal sequence transmitted from the m^{th} transmit antenna over the $P \times T$ transmission periods. The p^{th} column of $\mathbf{S}$ is the signal sequence transmitted simultaneously at time t_p , over the M transmit antennas.

The received signals are arranged into an $N \times P$ matrix $\mathbf{X}$, given by

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1P} \\ x_{21} & x_{22} & \cdots & x_{2P} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{NP} \end{bmatrix} \tag{2.2.2}$$

The n^{th} row of $\mathbf{X}$ is the signal sequence received at the n^{th} transmit antenna over the $P \times T$ transmission periods. The p^{th} column of $\mathbf{X}$ is the signal sequence received simultaneously at time t_p, over the N receive antennas.

Signals arriving at different receive antennas undergo independent fading. The signal at each receive antenna is a noisy superposition of the faded versions of the transmitted signals. A flat Rayleigh fading channel is assumed. At time t, the received signal at receive antenna n is given by

$$x_{nt} = \sum_{m=1}^{M} h_{nm} s_{mt} + w_{nt}, \ t = t_1, \ ..., t_P, \ n = 1, \ ..., \ N \tag{2.2.3}$$

where h_{nm} is the fading attenuation for the path from transmit antenna m to receive antenna n at time t, which is a independent complex Gaussian random variable with zero mean and variance $1/2$ per dimension. w_{nt} is the additive noise component at receive antenna n at time t, which is an independent sample of the zero mean complex Gaussian random variable with variance σ^2.

According to (2.2.3), the received signal vector can be related to the transmitted signal vector by

$$\mathbf{X} = \mathbf{HS} + \mathbf{W} \tag{2.2.4}$$

where $\mathbf{S}$ is the $M \times P$ complex transmitted signal matrix as given in (2.2.1), $\mathbf{X}$ is the $N \times P$ complex received signal matrix as given in (2.2.2), $\mathbf{W}$ is the $N \times P$ additional noise matrix and $\mathbf{H}$ is the $N \times M$ channel coefficient matrix. In this notation, all signals and noise matrices are function of time.

The received signals are decoded by a space-time decoder. We assume that the space-time decoder is based on the maximum likelihood Viterbi algorithm. The Viterbi algorithm tracks valid space-time code sequences in the code trellis and selects one that is closet to the received sequence based on the Euclidean distance path metric.

Assuming perfect CSI, the branch metric of the Viterbi algorithm at time t is computed as

$$\sum_{n=1}^{N} \left| x_{nt} - \sum_{m=1}^{M} h_{nm} s_{mt} \right|^2 \tag{2.2.5}$$

The path metric is given by

$$\sum_{t=1}^{P} \sum_{n=1}^{N} \left| x_{nt} - \sum_{m=1}^{M} h_{nm} s_{mt} \right|^2 \tag{2.2.6}$$

The Viterbi algorithm selects the path with the lowest accumulated path metric as the decoded codeword.

In the absence of perfect CSI, a channel estimator must be applied to get the channel estimates and then these channel estimates are used for decoding.

There are various approaches of space-time codes in their coding structures, including ST block codes (STBC) [1]-[3], ST trellis codes (STTC) [4], ST turbo trellis coded modulation (TCM) [5] and layered ST (LST) architectures [6]. STTC offers the maximum possible diversity gain and the coding gain without any sacrifice in the transmission bandwidth. The decoding of these codes, however, would require the use of a vector form of the Viterbi decoder. When the number of transmit antennas is fixed, the decoding complexity of STTC increases exponentially with transmission rate. On the contrary, STBC can offer a much simple way of obtaining transmitter diversity without any sacrifice in bandwidth and without requiring huge decoding complexity.

2.3 SPACE-TIME BLOCK CODING

In addressing the issue of decoding complexity in space-time codes, Alamouti [1] discovered a remarkable space-time block coding scheme for transmission with two transmit antennas, which supports maximum-likelihood detection based only on linear processing at the receiver. This scheme was later generalized in [2]-[3] to an arbitrary number of antennas and is able to achieve the full diversity promised by the number of transmit and receive antennas.

In Alamouti's scheme, during any given transmission period two signals are transmitted simultaneously from two transmit antennas. The transmission matrix is given by

$$\mathbf{S_2} = \begin{bmatrix} d_1 & -d_2^* \\ d_2 & d_1^* \end{bmatrix} \qquad (2.3.1)$$

where d^* is the complex conjugate of d.

During the first transmission period, two signals, d_1 and d_2, are simultaneously transmitted from transmit antenna one and transmit antenna two, respectively. During the second transmission period, signal $-d_2^*$ is transmitted from transmit antenna one and signal d_1^* is transmitted from transmit antenna two, simultaneously. It is clear that the encoding is done in both space and time domain.

The key feature of Alamouti's encoding scheme is that

$$\mathbf{S_2} \bullet \mathbf{S_2}^* = \left(|d_1|^2 + |d_2|^2\right)\mathbf{I_2} \tag{2.3.2}$$

where $\mathbf{S_2}^*$ is the Hermitian (transpose conjugate) of $\mathbf{S_2}$ and $\mathbf{I_2}$ is the 2×2 identity matrix.

Let us assume that one receive antenna is used at the receiver. The channel fading coefficients from the first and second transmit antennas to the receive antenna are denoted by h_{11} and h_{12}, respectively. At the receive antenna, the received signals over two consecutive transmission periods, denoted by x_{11} and x_{12}, respectively, can be expressed using (2.2.3) as

$$\begin{aligned} x_{11} &= h_{11}d_1 + h_{12}d_2 + w_{11} \\ x_{12} &= -h_{11}d_2^* + h_{12}d_1^* + w_{12} \end{aligned} \tag{2.3.3}$$

where w_{11} and w_{12} are additive complex noise at the receive antenna at these two consecutive transmission periods, respectively.

If the channel fading coefficients, h_{11} and h_{12}, can be perfectly recovered at the receiver, the receiver will use them as the CSI in the decoder. A combiner forms the following combined signals

$$\begin{aligned} \tilde{d}_1 &= h_{11}^* x_{11} + h_{12} x_{12}^* \\ \tilde{d}_2 &= h_{12}^* x_{11} - h_{12} x_{12}^* \end{aligned} \tag{2.3.4}$$

Substituting for x_{11} and x_{12} from (2.3.3), the combined signals can be written as

$$\tilde{d}_1 = \left(|h_{11}|^2 + |h_{12}|^2\right) d_1 + h_{11}^* w_{11} + h_{12} w_{12}^*$$
$$\tilde{d}_2 = \left(|h_{11}|^2 + |h_{12}|^2\right) d_2 - h_{11} w_{12}^* + h_{12}^* w_{11} \tag{2.3.5}$$

As the signal $\tilde{d}_1$ depends only on d_1 and the signal $\tilde{d}_2$ depends only on d_2, we can decide on d_1 and d_2 by applying the maximum likelihood rule on $\tilde{d}_1$ and $\tilde{d}_2$ separately. These combined signals are sent to a maximum likelihood detector which selects a symbol $\hat{d}_i, i = 1, 2$, for each transmitted symbol d_i, from the M-ary signals set, such that the Euclidean distance between the two symbols $\tilde{d}_i$ and $\hat{d}_i$ is minimum, where $\hat{d}_i$ is the estimate of the transmitted symbol d_i. The complexity of the decoder is linearly proportional to the number of antennas and the transmission rate. The distinguished feature of this type of space-time codes is a very simple maximum likelihood decoding algorithm based only on linear processing at the receiver.

In general, a space-time block code is defined by an $M \times P$ transmission matrix $\mathbf{G}$, here M represents the number of transmit antennas and P represents the number of time periods for transmission of one block of symbols. The K modulated symbols d_1, d_2, ..., d_K are encoded by a space-time block encoder to generate M parallel signal sequences of length P according to the transmission matrix $\mathbf{G}$. The entries of this matrix are linear combination of these K modulated symbols and their conjugates.

These coded sequences will be transmitted through M transmit antennas simultaneously in P transmission periods. The m^{th} row of $\mathbf{G}$ is the signal sequence transmitted from the m^{th} transmit antenna over the P transmission periods. The p^{th} column of $\mathbf{G}$ is the signal sequence transmitted simultaneously at time t_p, over the M transmit antennas.

In order to achieve full transmit diversity of M, the transmission matrix $\mathbf{G}$ is constructed based on orthogonal designs such that

$$\mathbf{G} \bullet \mathbf{G}^* = \left(|d_1|^2 + |d_2|^2 + \cdots + |d_K|^2\right)\mathbf{I_M} \tag{2.3.6}$$

where $\mathbf{G}^*$ is the Hermitian of $\mathbf{G}$ and $\mathbf{I_M}$ is the $M \times M$ identity matrix.

The rate of a space-time block code is defined as the ratio between the number of symbols the encoder takes as its input and the number of transmission periods. It is given by

$$R = K / P \tag{2.3.7}$$

The rate of a space-time block code with full transmitter diversity is less than or equal to one ($R \le 1$). The code with full rate ($R = 1$) requires no bandwidth expansion while the code with rate $R < 1$ will have the bandwidth expansion of $1/R$.

Note that orthogonal designs are applied to construct space-time block codes. The rows of the transmission matrix are orthogonal to each other. The orthogonality enables to achieve the full transmitter diversity for a given number of transmit antennas. In addition, it allows the receiver to decouple the signals transmitted from different antennas. Consequently, a simple maximum likelihood decoding, based only on linear processing at the receiver can be performed.

CHAPTER 3

OVERVIEW OF BLIND CHANNEL ESTIMATION

In this chapter, a review of recent blind channel estimation algorithms is presented. From the moment-based methods to the maximum likelihood (ML) methods, we outline basic ideas behind some new developments. The assumptions, identifiability conditions and their performance are given.

3.1 INTRODUCTION

There have been considerable interests in the so called "blind" problem. The impetus behind the increased research activities in blind techniques is perhaps their potential application in wireless communications, which are currently experiencing explosive growth.

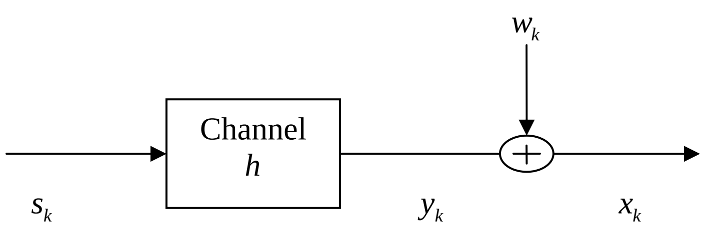

Figure 3-1: Schematic of blind channel estimation

The basic blind channel estimation problem involves a channel model shown in Figure 3-1, where only the observed signal is available for processing in the identification and estimation of channel. This is in contrast to the identification and estimation problem in classical input-output system where both input and observation are used.

The essence of blind channel estimation rests on the exploitation of channel structures and properties of inputs. Existing blind channel estimation algorithms are classified into the moment-based methods and the ML methods. We further divide these algorithms based on the modeling of the input signals. If input is assumed to be random with prescribed statistics (or distributions), the corresponding blind channel estimation schemes are considered to be statistical. On the other hand, if the input does not have a statistics description, or although the source is random but the statistical properties of the source are not exploited, the corresponding estimation algorithms are classified as deterministic. Figure 3-2 shows a map for different classes of algorithms.

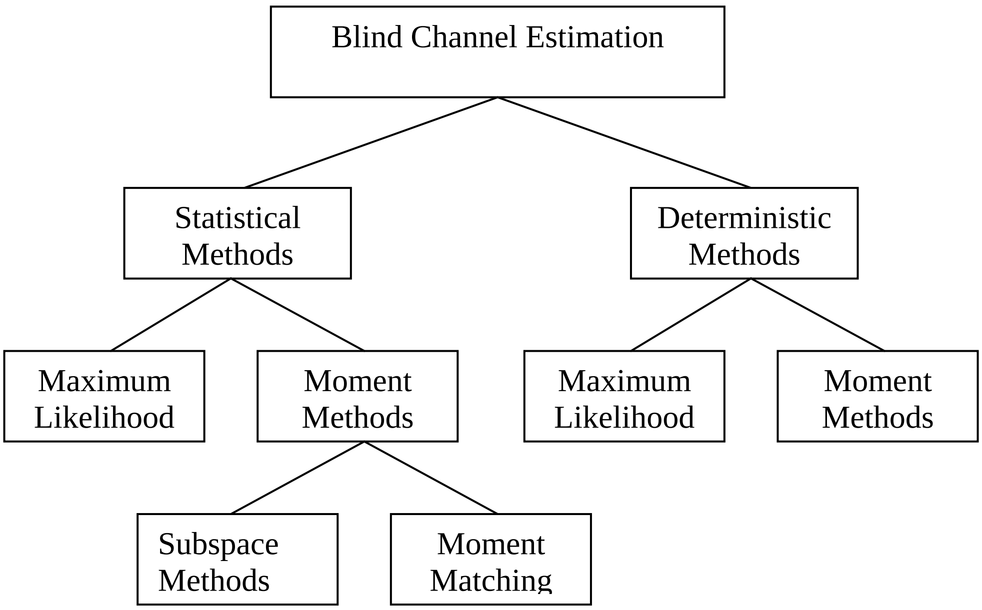

Figure 3-2: Classification of blind channel estimation methods.

3.2 THE SUBSPACE METHODS

Many recent blind channel estimation techniques exploit subspace structures of observations. The key idea is that the channel (or part of the channel) vector is in a one-dimensional subspace of either the observation statistics or a block of noiseless observations. These methods are often referred to as the subspace methods, which are considered as parts of the moment methods sometimes. They are attractive because of the closed form identification. On the other hand, as they rely on the property that the channel lies in a unique direction (subspace), they may not be robust against modelling errors, especially when the channel matrix is close to being singular. The second disadvantage is that they are often more computationally expensive.

3.2.1 DETERMINISTIC SUBSPACE METHODS

Deterministic subspace methods do not assume that the input source has a specific statistical structure. A more striking property of deterministic subspace methods is the so-called finite sample convergence property. Namely, when there is no noise, the estimator produces the exact channel using only a finite number of samples, provided that the identifiably condition is satisfied. Therefore, these methods are most effective at high SNR and for small data sample scenarios. On one hand, deterministic methods can be applied to a much wide range of source signals. On the other hand, not using the source statistics affects its asymptotic performance, especially when the identifiability condition is close to be violated.

1) Assumptions: The following conditions are assumed:

1.1) The noise sequence w_k is zero mean, white with known covariance σ^2;

1.2) The channel has known order L;

The assumption that the channel order L is known may not be practical. To address this problem, there are three kinds of approaches. First, channel order detection and parameter estimation can be performed separately. There are well known order detection schemes that can be used in practice. Second, some statistical subspace methods require only the upper bound of L. Third, channel order detection and parameter estimation can be performed jointly. Similarly, the noise variance σ^2 may be unknown in practice, but it can be estimated in many ways.

2) Identifiability: Under above assumptions, the channel coefficients can be uniquely identified up to a constant factor from the noiseless observation sequence y_k if:

2.1) The sub-channels are coprime;

2.2) The source sequence s_k has linear complexity greater than $2L$;

3) Examples: Some approaches of the deterministic subspace methods are described below.

The cross relation (CR) approach [10] wisely exploits the multi-channel structure. It is very efficient for small data sample applications at high SNR. The main problem of this approach is that the channel order L cannot be over estimated. For finite samples, this algorithm may also be biased.

The noise subspace approach [11] exploits the structure of the filtering matrix directly. There is a strong connection between the CR approach and the noise subspace approach. They are different only in their choices of parameterizing the signal or the noise subspace. Similar to the CR approach, the noise subspace approach also requires the knowledge of the channel order L and it is suitable for short data size applications. Although it is a bit more complex than the CR approach, it appears to offer improved performance in many cases.

Although deterministic approaches enjoy the advantage of having fast convergence, they share some common difficulties. For example, the determination of the channel order is required and often difficult. Second, the adaptive implementation of these algorithms is not straightforward. Recently, a new approach based on the least squares smoothing (LSS) of the observation process is proposed [12]. The key idea of LSS rests on the isomorphic relation between the input and the observation spaces. This approach has two attractive features. First, it converts a channel estimation problem to a linear LSS problem for which there are efficient adaptive implementations using lattice filters. Second, a joint channel order detection and channel estimation algorithm can be derived that determines the best channel order and channel coefficients to minimize the smoothing error.

3.2.2 SECOND-ORDER STATISTICAL SUBSPACE METHODS

In statistical subspace approaches, it is assumed that the source is a random sequence with known second-order statistics.

1) Assumptions: Although algorithms discussed here can be extended in many different ways, we shall assume the following assumptions in our discussion.

1.1) The source sequence s_k is zero mean, white with unit variance;

1.2) The noise sequence w_k, uncorrelated with s_k, is zero mean, white, with known covariance σ^2;

1.3) The channel order L is known;

Most algorithms of the statistical methods can be extended to cases where the noise is colored but with known correlations. Some statistical methods do not require knowledge of the channel order. Instead, they require the upper bound of the channel order.

2) Identifiability: Under above assumptions, the channel can be uniquely identified up to a constant factor from the autocorrelation matrix $\mathbf{R}_{\mathbf{xx}}$ if and only if the sub-channels are coprime.

3) Examples: Some approaches of the second-order subspace methods are described below.

3.1) Identification via Cyclic Spectra: This approach [13] exploits the complete cyclic statistics of the received and source signals, as well as the FIR structure of the channel model. The disadvantage of this algorithm is that it requires the convergence of the source statistics, which means that even when there is no noise, there is estimation error for any fixed sample size, although the algorithm is mean square consistent.

3.2) Identification via Filtering Transform: This approach [14] introduces a two-step closed form identification algorithm. It first finds the filtering matrix and then estimates the channel from the estimated filtering matrix. The implementation of this algorithm requires the channel order and the noise variance. While it is consistent, this approach may not perform well for two reasons. First, the algorithm fails to take advantage of the special structure of the filtering transform. Second, the performance of such a two-step procedure is often affected by the quality of the estimation in the first step.

3.3) Identification via Linear Prediction: This approach [15] uses all second-order statistics of the received signal and it is mean square consistent. It does not require the exact channel order, thus it is robust against over-determination of the channel order. Derived from the noiseless model, the linear prediction idea is no longer valid in the presence of noise. However, when channel parameters are estimated from the automation functions, the effect of noise can be lessened by subtracting the terms related to the noise correlation. The main disadvantage of this algorithm is that it is a two-step approach whose performance depends on the accuracy of the estimates from the first step.

3.2.3 OTHER RELATED SUBSPACE APPROACHES

Some related approaches have been developed recently which can be applied to the general subspace methods to improve performance. For example, the weighted subspace approach, successfully used in the direction of arrival estimation in array signal processing, employs an additional weighting matrix which is chosen optimally. The optimal selection of the weighted matrix is, however, nontrivial, and it is often a function of the true channel parameters. A practical solution is to use a consistent estimate of the channel to construct the optimal weighting matrix.

3.3 OPTIMAL MOMENT METHODS

When the source has a statistical model, most subspace methods are part of the moment methods. They all can be viewed as estimating channel parameters from the estimated second-order moments of the received signals. For the class of consistent estimators, asymptotic normalized mean square error (ANMSE) can be used as a performance measure. Small ANMSE is desired in blind channel estimators using the second-order moment methods. The optimal moment methods with the minimum ANMSE can be achieved with some certain conditions. The moment matching approach is motivated by the existence of a moment method that achieves the minimum ANMSE. While moment matching methods have a robust performance against channel order selection and the channel condition, they are unfortunately not easy to implement because of the existence of local minima in the optimization.

3.4 THE ML METHODS

One of the most popular parameter estimation algorithms is the ML method. Not only can such methods be derived in a systematic way, but more importantly, the class of ML estimators are usually optimal for large data records as they approximate the minimum variance unbiased estimators. Asymptotically, under certain regularity conditions, the variances of ML estimators approach the Cramer-Rao Bound (CRB), which is the lower bound on variances for all unbiased estimators. Unfortunately, unlike subspace based approaches, the ML methods usually cannot be obtained in closed form. Their implementations are further complicated by the existence of local minima. However, ML approaches can be made very effective by including the subspace and other suboptimal approaches as initialization procedures.

We will briefly introduce the general formulation of the ML estimation, which can be found in many textbooks. The problem at hand is to estimate the deterministic (vector) parameter θ given the probabilistic model of the observation. Specifically, let $f\left(y;\theta\right)$ be the probability density function of random variable Y parameterized by $\theta \in \Theta$. Given an observation $Y = y$, θ is estimated by maximizing

$$\hat{\theta} = \arg\max_{\theta\in\Theta} f\left(y;\theta\right) \tag{3.4.1}$$

where $f\left(y;\theta\right)$, when viewed as the function of θ, is referred to as the likelihood function.

3.4.1 DETERMINISTIC ML APPROACHES

The deterministic ML (DML) approach assumes no statistical model for the input sequence s_k. In other words, both the channel coefficient vector $\mathbf{H}$ and the input source vector $\mathbf{S}$ are parameters to be estimated.

Consider the channel model in Figure 3-1, the DML problem can be stated as follows: given $\mathbf{X}$, estimate $\mathbf{H}$ and $\mathbf{S}$ by

$$\left\{\hat{\mathbf{H}}_{\mathbf{DML}},\hat{\mathbf{S}}_{\mathbf{DML}}\right\}=\arg\max_{\mathbf{H},\mathbf{S}} f\left(\mathbf{X};\mathbf{H},\mathbf{S}\right) \qquad (3.4.1.1)$$

where $f\left(\mathbf{X};\mathbf{H},\mathbf{S}\right)$ is the density function of the observation vectors $\mathbf{X}$ parameterized by both the channel coefficients $\mathbf{H}$ and the input source $\mathbf{S}$.

1) Assumptions: In considering the deterministic model, we assume the following assumptions.

1.1) The noise sequence w_k is zero mean Gaussian with known covariance σ^2.

1.2) The channel has known order L.

The assumptions for DML are almost the same as those for the deterministic subspace methods, except that the noise in DML is assumed to be Gaussian. The noise variance can also be considered as part of the parameters to be estimated in some approaches.

2) Identifiability: It is not surprising that the identifiability condition for DML is the same as that for the deterministic second-order moment methods. Specifically, the channel is identifiable if the sub-channels are coprime and the source sequence has linear complexity greater than $2L+1$. The reason is that, when the noise is Gaussian, all information about the channel in the likelihood function resides in the second-order moments of the observations.

3) Examples: Some approaches of the DML methods are given below. The iterative quadratic ML (IQML) approach [16] transforms the DML problems into a sequence of quadratic optimization problems for which simple solutions can be obtained. The two-step maximum likelihood (TSML) approach [17] uses the CR methods to obtain an initial estimate of the channel and then this initial estimate is used for optimization.

3.4.2 STATISTICAL ML APPROACHES

In statistics ML (SML) approaches, we consider the statistical model where the source sequence s_k is random with known distribution. In such formulation, the only unknown parameter is the channel vector.

Consider the channel model in Figure 3-1, the SML problem can be stated as follows: given $\mathbf{X}$, estimate $\mathbf{H}$ by

$$\hat{\mathbf{H}}_{\mathbf{SML}} = \arg\max_{\mathbf{H}} f\left(\mathbf{X};\mathbf{H}\right) \tag{3.4.2.1}$$

where $f\left(\mathbf{X};\mathbf{H}\right)$ is the density function of the observation vectors $\mathbf{X}$ parameterized by $\mathbf{H}$.

1) Assumptions: The SML estimation hinges on the availability and the evaluation of the likelihood function. Although the SML methods can be applies to more general cases, we shall make the following assumptions in our discussion.

1.1) Components of the source $\mathbf{S}$ and the noise $\mathbf{W}$ are jointly independent;

1.2) The noise sequence w_k is zero mean Gaussian with covariance σ^2;

1.3) Components of the source $\mathbf{S}$ are independent, identically distributed (i.i.d.) with known probability density function.

2) Identifiability: Identifiability remains to be an important issue in the SML approach. The identifiability condition tells when the SML method can be applied. A main issue is whether the likelihood function provides sufficient information to distinguish different models. Under above assumptions, the channel parameter is identifiable by the likelihood function if and only if one of the following conditions is satisfied:

2.1) The source $\mathbf{S}$ is non-Gaussian;

2.2) The sub-channels are coprime;

Obviously, parameters identifiable by the moment methods are identifiable by the likelihood function. It is not surprised to see that the class of channels identifiable by the SML methods is larger than that by the moment methods.

3) Examples: The expectation-maximization (EM) algorithm was proposed in [18] to transform the complicated optimization in (3.4.2.1) to a sequence of quadratic optimizations. The performance of the EM algorithm depends on its initialization, which may be facilitated by the moment techniques such as those described in Section 3.2. When the EM algorithm converges globally, the estimate achieves asymptotically the CRB for the case of i.i.d. sequences.

CHAPTER 4

DEML CHANNEL ESTIMATOR

In this chapter, we will present a computationally efficient channel estimation method based on the decoupled maximum likelihood (DEML) algorithm. The DEML channel estimator decouples the multi-dimensional problem of the exact ML estimator into a set of one-dimensional problems and hence is computationally efficient. The properties of the DEML channel estimator are also given in this chapter.

4.1 PROBLEM FORMULATION

Space-time coding has been shown to be a promising technique for increasing the capacity of wireless systems. The decoding of space-time codes requires the perfect CSI at the receiver. In the absence of perfect CSI at the receiver, a channel estimator must be used to estimate the channel coefficients. Then these channel estimates are used as if they were perfect known at the receiver to extract symbol estimates.

Although many high-resolution estimation algorithms have been devised in the past few decades, these research efforts are mainly put on the areas, where a priori knowledge is not available to the receivers. These algorithms are developed without considering any knowledge of the input signals, except for some general statistical properties such as the second-order ergodicity. Several deterministic or statistical estimators are also devised for such applications. The deterministic estimators, such as the DML estimators, model the unknown signals as the unknown deterministic parameters. The statistical estimators, such as the SML estimators, model the unknown signals as random processes.

But in some applications especially in a mobile communication system, a priori knowledge is known to the receivers, although the actual transmitted symbol stream is unknown. In such a system, a known preamble is added to the message for training purposes. Such extra information may be exploited to enhance the accuracy of the estimates and may be used to simplify the computational complexity of the estimation algorithms.

Consider the wireless communication system with M transmit antennas and N receive antennas. The received data vector can be modelled as

$$\mathbf{X} = \mathbf{HS} + \mathbf{W} \tag{4.1.1}$$

where $\mathbf{X}$ is the $N \times T$ complex received signal vector, $\mathbf{S}$ is the $M \times T$ complex transmitted signal vector, $\mathbf{W}$ is the $N \times T$ additive noise vector and $\mathbf{H}$ is the $N \times M$ channel coefficient matrix. In this notation, all signal and noise vectors are function of time.

The waveforms of the transmitted signals are assumed to be known and the fading channel is assumed to be quasi-static. The noise vector is assumed to be a complex Gaussian random vector with zero-mean and arbitrary covariance matrix $\mathbf{Q}$ and is sampled to be temporally white, i.e.

$$E[w(t_i)w^*(t_j)] = \mathbf{Q}\delta_{i,j} \tag{4.1.2}$$

where $(\bullet)^*$ denotes the complex conjugate transpose, and $\delta_{i,j}$ is the Kronecker delta function. The unknown covariance matrix $\mathbf{Q}$ models both thermal noises caused by the sensor output receivers and all other outside radio interference and jamming. Finally the signal and the noise vectors are assumed to be uncorrelated, i.e.

$$\lim_{L\to\infty}\frac{1}{L}\sum_{l=1}^{L}\mathbf{S}(t_l)\mathbf{W}^*(t_l)=\lim_{L\to\infty}\frac{1}{L}\mathbf{SW}^*=0 \tag{4.1.3}$$

with probability 1.

The problem of interest herein is to determine the channel coefficients matrix $\mathbf{H}$ and the noise covariance matrix $\mathbf{Q}$ from the L independent data samples $\mathbf{X}(t_1),\mathbf{X}(t_2),\ldots,\mathbf{X}(t_L)$.

4.2 DEML CHANNEL ESTIMATOR

We consider below a large sample estimator based on the DEML algorithm for estimating channel coefficients matrix $\mathbf{H}$ and noise covariance matrix $\mathbf{Q}$. It is easy to see that an exact ML estimator requires a multi-dimensional search over the parameter space and is computationally burdensome. We shall show below that the DEML channel estimator decouples the K-dimensional search problem into K one-dimensional search problems for an arbitrary sensor array and hence it is computationally efficient.

The log-likelihood function of the received signals $\mathbf{X}(t_l), l=1,2,\ldots,L$ is proportional to (within an additive constant) [9]

$$-\ln|\mathbf{Q}|-\frac{1}{L}tr\left\{\mathbf{Q}^{-1}(\mathbf{X}-\mathbf{HS})(\mathbf{X}-\mathbf{HS})^*\right\} \tag{4.2.1}$$

where $|\bullet|$ denotes the determinant, $tr\{\cdot\}$ denotes the trace operation and $(\cdot)^*$ denotes the conjugate transpose.

It is easy to show that maximizing this likelihood function with respect to $\mathbf{Q}$ yields

$$\hat{\mathbf{Q}} = \frac{1}{L}(\mathbf{X}-\mathbf{HS})(\mathbf{X}-\mathbf{HS})^* \tag{4.2.2}$$

where $\hat{\mathbf{Q}}$ is the estimate of $\mathbf{Q}$.

Substituting (4.2.2) in (4.2.1), we can see that maximizing the log-likelihood function is equivalent to minimizing

$$\left|\frac{1}{L}(\mathbf{X}-\mathbf{HS})(\mathbf{X}-\mathbf{HS})^*\right| \tag{4.2.3}$$

Let the "covariance matrix" $\mathbf{R}_{\mathrm{SX}}$, $\mathbf{R}_{\mathrm{SS}}$ and $\mathbf{R}_{\mathrm{XX}}$ be defined as follows.

$$\mathbf{R}_{\mathrm{SX}} \triangleq \frac{1}{L}\sum_{l=1}^{L}\mathbf{S}(t_l)\mathbf{X}^*(t_l) = \frac{1}{L}\mathbf{SX}^* \tag{4.2.4}$$

$$\mathbf{R}_{\mathrm{SS}} \triangleq \frac{1}{L}\sum_{l=1}^{L}\mathbf{S}(t_l)\mathbf{S}^*(t_l) = \frac{1}{L}\mathbf{SS}^* \tag{4.2.5}$$

$$\mathbf{R}_{\mathrm{XX}} \triangleq \frac{1}{L}\sum_{l=1}^{L}\mathbf{X}(t_l)\mathbf{X}^*(t_l) = \frac{1}{L}\mathbf{XX}^* \tag{4.2.6}$$

Now we can calculate

$$\begin{aligned}\mathbf{F} &\triangleq \frac{1}{L}(\mathbf{X}-\mathbf{HS})(\mathbf{X}-\mathbf{HS})^* \\ &= \mathbf{R}_{\mathrm{XX}} - \mathbf{HR}_{\mathrm{SX}} - \mathbf{R}_{\mathrm{SX}}^*\mathbf{H}^* + \mathbf{HR}_{\mathrm{SS}}\mathbf{H}^* \qquad (4.2.7) \\ &= \left(\mathbf{H}-\mathbf{R}_{\mathrm{SX}}^*\mathbf{R}_{\mathrm{SS}}^{-1}\right)\mathbf{R}_{\mathrm{SS}}\left(\mathbf{H}-\mathbf{R}_{\mathrm{SX}}^*\mathbf{R}_{\mathrm{SS}}^{-1}\right)^* + \mathbf{R}_{\mathrm{XX}} - \mathbf{R}_{\mathrm{SX}}^*\mathbf{R}_{\mathrm{SS}}^{-1}\mathbf{R}_{\mathrm{SX}}\end{aligned}$$

Since the matrix $\mathbf{R}_{SS}$ is positive definite and the second and third terms in (4.2.7) do not depend on $\mathbf{H}$, it follows that

$$\mathbf{F} \geq \mathbf{F}\,|_{\mathbf{H}=\mathbf{R}_{SX}^{*}\mathbf{R}_{SS}^{-1}} \tag{4.2.8}$$

Since the whole sample covariance matrix $\mathbf{F}$ is minimized, the estimate $\hat{\mathbf{H}} = \mathbf{R}_{SX}^{*}\mathbf{R}_{SS}^{-1}$ of $\mathbf{H}$ will minimize any non-decreasing function of $\mathbf{F}$ including the determinant of $\mathbf{F}$, which is $|\mathbf{F}|$ in (4.2.3). Thus we get the estimate of $\mathbf{H}$ as:

$$\hat{\mathbf{H}} = \mathbf{R}_{SX}^{*}\mathbf{R}_{SS}^{-1} \tag{4.2.9}$$

It is easy to see that $\hat{\mathbf{H}}$ is a consistent estimate of $\mathbf{H}$.

Substituting (4.2.9) back into (4.2.2), the estimate of $\mathbf{Q}$ is given as

$$\hat{\mathbf{Q}} = \mathbf{R}_{XX} - \mathbf{R}_{SX}^{*}\mathbf{R}_{SS}^{-1}\mathbf{R}_{SX} \tag{4.2.10}$$

It is also easy to see that $\hat{\mathbf{Q}}$ is a consistent estimate of $\mathbf{Q}$.

In this way, we decouple the multi-dimensional problem of the exact maximum likelihood estimator in (4.2.1) into a set of one-dimensional problems as given by (4.2.9) and (4.2.10). A decoupled maximum likelihood (DEML) channel estimator is formed.

If the incident signals are uncorrelated with each other, all estimates mentioned above are consistent and large sample realizations of the ML estimates, it follows that the estimation method is asymptotically statistically efficient, according to the general properties of ML estimators.

If the incident signals are moderately correlated, the DEML estimator is no longer a large sample ML estimator. The performance of the DEML estimator has small degradation. But the asymptotic statistical performance is still close to that of CRB's [9].

If the incident signals are highly correlated, the performance of the DEML has obvious degradation. Thus the DEML channel estimator can not be used on the highly correlated incident signals directly. In such cases, a decorrelation algorithm must be applied on the correlated incident signals before the DEML channel estimator can be used on them.

4.3 PROPERTIES

We will now give some significant advantages of the DEML estimator for uncorrelated signals with known waveforms as compared with other standard ML estimators for uncorrelated signals with unknown waveforms.

First, the large sample and asymptotically statistically efficient DEML estimator is much more computationally efficient than any existing large sample ML estimators for unknown waveform signals. It has been shown that for the case of uncorrelated and unknown waveform signals, the K-dimensional estimation problem can also be asymptotically decoupled into K one-dimensional problems with the standard ML estimators. These ML estimators, however, require the eigen-decomposition of the array covariance matrix, which is computationally expensive.

On the contrary, the cost function associated with the DEML estimator does not require any eigen-decomposition. Moreover, on a parallel computer, the DEML estimator can be naturally implemented in a parallel fashion, i.e., by calculating the estimate of the channel coefficient matrix in (4.2.9) and the estimate of the noise covariance matrix in (4.2.10) in parallel.

Second, the accuracy provided by the DEML estimator for uncorrelated signals with known waveforms is superior to that of the best one provided by the estimators for unknown waveform signals. In fact, when unknown waveform signals are modelled as unknown deterministic parameters and the number of array sensors is finite, no estimator can achieve its CRB, which is bound to be greater than or equal to the CRB for signals with known waveforms due to the parsimony principle.

Third, the DEML estimator has no constraints on the number of incident signals at all, provided that the number of data samples is large enough, while the estimators for unknown waveform signals require that the number of signals be less than the number of array sensors.

Fourth, the DEML channel estimator can handle the case of unknown spatially colored noise with little additional difficulties. The estimators for unknown waveform signals, however, fail to handle this case. This advantage of the DEML channel estimator is particularly useful for estimating the incident signals with known waveforms in the presence of unknown interfering and jamming signals that are not completely correlated with any of these known waveform signals. This is especially true when the number of interfering and jamming signals is large and when some of the interfering and jamming signals are wideband. The unknown noise covariance matrix **Q** may be used to accommodate both the presence of these interfering and jamming signals and any other noise, including the thermal noise.

CHAPTER 5

PERFORMANCE OF DEML CHANNEL ESTIMATOR UNDER SPATIALLY UNCORRELATED FADING CHANNEL

In this chapter, we deal with the STBC system under spatially uncorrelated flat Rayleigh fading channel. First, the STBC system model is summarized. The DEML channel estimator performs well when the incident signals are uncorrelated with each other. Thus it can be applied directly to the uncorrelated STBC system. In addition, an iterative ML detector is introduced to improve the system performance with the DEML estimator. The system BER performances and some discussions are given at the last part of this chapter.

5.1 SYSTEM MODEL

Consider a STBC system with M transmit and N receive antennas. The K modulated symbols $d_1,\ d_2,\ ...,\ d_K$ are encoded by a space-time block encoder. The output of the encoder is arranged into M blocks, each containing P complex modulation signals, described by matrix $\mathbf{S}$ as follows

$$\mathbf{S} = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1P} \\ s_{21} & s_{22} & \cdots & s_{2P} \\ \vdots & \vdots & \ddots & \vdots \\ s_{M1} & s_{M2} & \cdots & s_{MP} \end{bmatrix} \tag{5.1.1}$$

The entries of the matrix are linear combinations of these K corresponding symbols and their conjugates, which belong to a finite complex constellation Ψ with

$|\Psi|$ elements. The m^{th} row of $\mathbf{S}$ is the signal sequence transmitted from the m^{th} transmit antenna over the P symbol periods. The p^{th} column of $\mathbf{S}$ is the signal sequence transmitted simultaneously at time t_p, over the M antennas.

The channel is assumed to be flat Rayleigh fading and quasi-static. The fading coefficient between the m^{th} transmit antenna and the n^{th} receive antenna is defined as h_{nm}, which is independent with respect to both m and n and is a complex Gaussian random variable with zero mean and variance $1/2$. It remains constant within, but changes to a new independent realization, every P symbol periods.

Let s_{mt} be the transmitted signal at the m^{th} transmitter and time t. The received signal at the n^{th} receiver and time t is given by

$$x_{nt} = \sum_{m=1}^{M} h_{nm} s_{mt} + w_{nt}, \ t = 1, \ ..., \ T, \ n = 1, \ ..., \ N \quad (5.1.2)$$

where w_{nt} denotes additive noise at the n^{th} receiver and time t, which is independent with respect to both n and t. It is the complex Gaussian noise with zero mean and variance σ^2.

The average energy of the transmitted symbols from each transmit antenna is normalized to be one. So the average energy of the received signal at each receive antenna is M. If we define the signal-to-noise ratio as SNR, we can get the noise variance $\sigma^2 = M/(2SNR)$.

Equation (5.1.2) can be re-written in matrix form as:

$$\mathbf{X} = \mathbf{HS} + \mathbf{W} \tag{5.1.3}$$

where $\mathbf{X}$ is the $N \times T$ complex received signal matrix, $\mathbf{S}$ is the $M \times T$ complex transmitted signal matrix, $\mathbf{W}$ is the $N \times T$ additive noise matrix and $\mathbf{H}$ is the $N \times M$ channel coefficient matrix. In this notation, all signal and noise matrices are function of time.

Although the spatial covariance of the additive noise $\mathbf{W}$ is difficult to determine, it can be written as $E[w(t_i)w^*(t_j)] = \mathbf{Q}\delta_{i,j}$, where $\mathbf{Q}$ denotes the unknown spatial covariance matrix and $\delta_{i,j}$ is the Kronecker delta function.

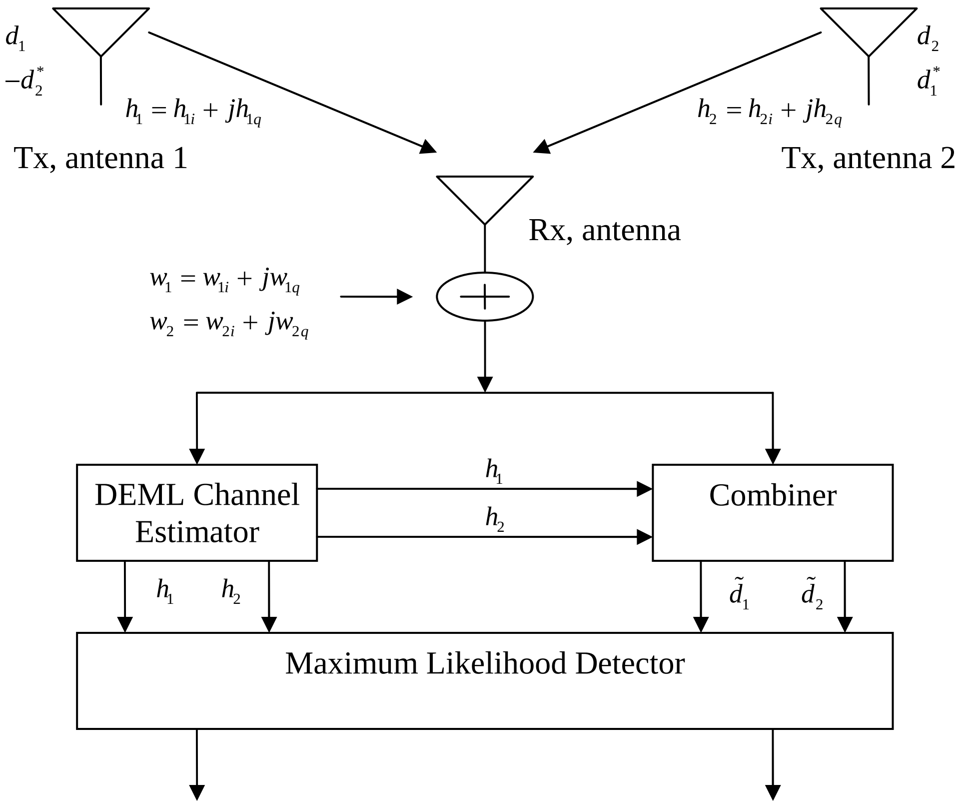

Figure 5-1: The STBC system with two transmit and one receive antennas

In Figure 5-1, the STBC system with two transmit antennas and one receive antenna is shown. It has four function parts: encoding and transmitting signals at the transmitter; channel estimation; combining scheme at the receiver and the decision rule for maximum likelihood detection. Here we use the same encoding scheme as that of Alamouti's scheme in [1] and we will introduce a DEML based channel estimator for the STBC system. Also an iterative ML detector is introduced to improve the system performance with the DEML estimator.

5.2 CHANNEL ESTIMATION

The channel estimation problem in STBC system is to determine the channel coefficients matrix **H** and the noise covariance matrix **Q** from the L independent data samples $\mathbf{X}(t_1), \mathbf{X}(t_2), \ldots, \mathbf{X}(t_L)$.

For the STBC system with uncorrelated fading channel, the incident signals are uncorrelated with each other. We can apply the DEML channel estimator directly to this kind of system. According to Chapter 4, the estimate of the channel coefficient matrix with DEML channel estimator is given by:

$$\hat{\mathbf{H}} = \mathbf{R}_{sx}^{*}\mathbf{R}_{ss}^{-1} \tag{5.2.1}$$

And the estimate of the noise variance matrix with DEML channel estimator is given by:

$$\hat{\mathbf{Q}} = \mathbf{R}_{xx} - \mathbf{R}_{sx}^{*}\mathbf{R}_{ss}^{-1}\mathbf{R}_{sx} \tag{5.2.2}$$

where

$$\mathbf{R}_{sx} = \frac{1}{L}\mathbf{S}\mathbf{X}^{*} \tag{5.2.3}$$

$$\mathbf{R}_{SS} = \frac{1}{L}\mathbf{SS}^{*} \tag{5.2.4}$$

$$\mathbf{R}_{XX} = \frac{1}{L}\mathbf{XX}^{*} \tag{5.2.5}$$

According to the properties of DEML channel estimator, all estimates mentioned above are consistent and large sample realizations of the ML estimates. It follows that the DEML channel estimator is asymptotically statistically efficient and computationally efficient in this kind of STBC system.

5.3 ML DETECTOR

The log-likelihood function of the received signals $\mathbf{X}$ can be written as:

$$L\left(\mathbf{X}\middle|\mathbf{H},\mathbf{Q},\mathbf{S}\right) = -\ln\left|\mathbf{Q}\right| - \frac{1}{L}tr\left\{\mathbf{Q}^{-1}\left(\mathbf{X}-\mathbf{HS}\right)\left(\mathbf{X}-\mathbf{HS}\right)^{*}\right\} \tag{5.3.1}$$

where $|\bullet|$ denotes the determinant, $Tr\{\bullet\}$ denotes the trace operation and $(\bullet)^{*}$ denotes the conjugate transpose.

5.3.1 COHERENT ML DETECTOR

If the channel coefficient matrix **H** and the noise covariance matrix **Q** are assumed to be known, the detection of the symbols $d_1,\ d_2,\ ...,\ d_K$ would amount to maximizing (5.3.1) with respect to D, where D is the set of transmitted symbols $\{d_k\}_{k=1}^{K}$, or equivalently to minimizing

$$\arg\min_{D}\sum_{t=1}^{L}\sum_{n=1}^{N}\left|x_{nt} - \sum_{m=1}^{M}h_{nm}s_{mt}\right|^{2} \tag{5.3.1.1}$$

This can be reduced to minimize

$$\arg\min_{d_i}\left(\left|\tilde{d}_i - d_i\right|^2 + (-1+\sum_{m,n}\left|h_{nm}\right|^2)\left|d_i\right|^2\right),\ i=1,\ \ldots,\ K \tag{5.3.1.2}$$

for detecting the symbols seperately [2] and where

$$\tilde{d}_i = \sum_{t=1}^{L}\sum_{n=1}^{N} x_{nt} h^*_{n,\varepsilon_t(i)} \delta_t(i),\ \ i=1,\ \ldots,\ K \tag{5.3.1.3}$$

The definition of $\varepsilon_t(i)$ and $\delta_t(i)$ is described below. Given an orthogonal design, the columns of the transmission matrix $\mathbf{G}$ are all permutations of the first column of $\mathbf{G}$ with possibly different signs. The sign of d_i in the t^{th} column of $\mathbf{G}$ is denoted as $\delta_t(i)$. Let ε_t denote the permutations corresponding to these columns. Then $\varepsilon_t(i) = j$ means that d_i is up to a sign change in the $\left(j^{th}, t^{th}\right)$ element of $\mathbf{G}$. More detail information about $\varepsilon_t(i)$ and $\delta_t(i)$ can be found in [2].

This is a very simple decoding strategy which decouples the multi-dimensional detection problems in (5.3.1.1) into K scalar detection problems in (5.3.1.2). The detector in (5.3.1.2) will be referred to as coherent ML detector. Note that the decisions in (5.3.1.2) do not depend on the training block $\mathbf{X}_T$. This is natural since $\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_L$ are the sufficient statistics for the detection problems when $\mathbf{H}$ and $\mathbf{Q}$ are known.

5.3.2 EXACT ML DETECTOR

In the more realistic case that the channel coefficient matrix $\mathbf{H}$ and the noise covariance matrix $\mathbf{Q}$ are unknown, the likelihood function in (5.3.1) needs to be maximized with respects to $\mathbf{H}$ and $\mathbf{Q}$.

It can be shown that the maximization of (5.3.1) with respect to $\mathbf{H}$ and $\mathbf{Q}$ yields

$$\begin{aligned} L\left(\mathbf{X}\middle|D\right) &= L\left(\mathbf{X}\middle|\mathbf{S}\right) \\ &\triangleq \arg\max_{\mathbf{H},\mathbf{Q}} L\left(\mathbf{X}\middle|\mathbf{H},\mathbf{Q},\mathbf{S}\right) \\ &= \arg\max_{\mathbf{H},\mathbf{Q}} L\left(\mathbf{X}\middle|\mathbf{H},\mathbf{Q},D\right) \end{aligned} \tag{5.3.2.1}$$

We will refer to the decision that follows from the maximization of (5.3.2.1) with respect to D as the exact ML detector. Note that, however, the maximization of (5.3.2.1) is not attractive since it requires a search over $|\Psi|^K$ possible sequences of D. In what follows, we will present an iterative approach to maximizing (5.3.1) which decouples the search into a sequence of simple detection problems similar to that in the coherent ML detector.

5.3.3 TRAINING-BASED ML DETECTOR

An approximation to the exact ML detector in (5.3.2.1) can be easily derived by using the received training block $\mathbf{X}_T$ to estimate the channel coefficient matrix $\mathbf{H}$ and the noise covariance matrix $\mathbf{Q}$ with the DEML channel estimator described in Section 5.2, and then these estimates are used as if they were known in the coherent ML detector in (5.3.1.2). The obtained detector will be referred to as training-based ML detector.

The training-based ML detector therefore consists of the following steps:

Step 1. Obtain initial estimates of $\mathbf{H}$ and $\mathbf{Q}$ based on the training block $\mathbf{X}_T$ with the DEML channel estimator.

Step 2. Use the estimates obtained in step 1 to detect the symbols with the coherent ML detector.

5.3.4 ITERATIVE ML DETECTOR

The symbols detected in the training-based ML detector can be used to re-estimate the channel coefficient matrix $\mathbf{H}$ and the noise covariance matrix $\mathbf{Q}$ with the DEML estimator. Proceeding in this way, we get the iterative ML detector.

The iterative ML detector consists of the following steps:

Step 1. Obtain the initial estimates of $\mathbf{H}$ and $\mathbf{Q}$, using either estimates from previous block of data, or estimates from the training block (if this is the first part of transmission).

Step 2. Use the estimates of $\mathbf{H}$ and $\mathbf{Q}$ to detect the symbols with the coherent ML detector.

Step 3. Re-estimate $\mathbf{H}$ and $\mathbf{Q}$ using the DEML estimator described in Section 5.2 with the detected symbols in step 2.

Step 4. Repeat step 2 and step 3 until convergence or until a pre-imposed iteration number.

Some remarks on the iterative ML detector:

1. If only step 1 and step 2 are taken, the iterative ML detector is referred to as the training-based ML detector.

2. The training-based initialization in step 1 is somewhat ad-hoc, yet the remaining part of the algorithm is nothing but the cyclic maximization of the likelihood function. Hence the above algorithm obtains, after convergence, the exact ML detector in the case of unknown $\mathbf{H}$ and $\mathbf{Q}$.

3. The maximum of (5.3.1) is unique with probability 1, so the iterative ML detector will converge in no more than $|\Psi|^K$ steps.

4. Each step has a computational complexity of the same order as that of the training-based detector in Section 5.3.3. The increase in computational complexity induced by our iterative scheme compared to the training-based ML detector is therefore proportional to the number of iterations.

5.4 PERFORMANCES AND DISCUSSIONS

Some simulation results are presented to demonstrate the BER performance of the STBC system with DEML channel estimator under uncorrelated flat Rayleigh fading channel. These simulations are done for different number of transmitters and receivers, and for different encoding, decoding and modulation schemes [19]. Also the simulation results of the STBC system with perfect CSI are shown and compared.

Firstly we consider a STBC system with two transmitters ($M = 2$) and different number of receivers ($N = 1, 2$). The simulations are done for BPSK modulation scheme under flat Rayleigh fading channel. The symbols are encoded into the 2×2 complex orthogonal design as

$$\mathbf{S_2} = \begin{bmatrix} d_1 & -d_2^* \\ d_2 & d_1^* \end{bmatrix} \tag{5.4.1}$$

which corresponds to the encoding scheme proposed in [1] and $K = 2, P = 2$.

We consider the detection for every ten consecutive transmission blocks of which the first one is used as the training block. The training overhead is therefore $1/10 = 10\%$. Each sequence of these ten transmission blocks contains $2 \times 10 = 20$ samples and carries 18 information bits.

In Figure 5-2 and Figure 5-3, we show the BER performance of STBC system with DEML channel estimator under uncorrelated flat Rayleigh fading channel. For comparison, we also show BER performance of the STBC system with perfect channel state information. We can see that the DEML channel estimator performs well in the STBC system under uncorrelated flat Rayleigh fading channel. There is small degradation in the BER performance of the STBC system with DEML channel estimator. This degradation, however, is partially because of the training block introduced in this system, which is treated as the noise signal in estimation problem. We note that increasing the iteration number of the iterative ML detector can improve system BER performance. Three iterations are sufficient to make the system BER performance to converge to within 2 dB of that of the system with perfect channel state information.

Secondly we consider a STBC system with four transmitters ($M = 4$) and different number of receivers ($N = 1, 2$). The simulations are done for QPSK modulation scheme under uncorrelated flat Rayleigh fading channel. The symbols are encoded into complex orthogonal design as:

$$\mathbf{S_4} = \begin{bmatrix} d_1 & -d_2 & -d_3 & -d_4 & d_1^* & -d_2^* & -d_3^* & -d_4^* \\ d_2 & d_1 & d_4 & -d_3 & d_2^* & d_1^* & d_4^* & -d_3^* \\ d_3 & -d_4 & d_1 & d_2 & d_3^* & -d_4^* & d_1^* & d_2^* \\ d_4 & d_3 & -d_2 & d_1 & d_4^* & d_3^* & -d_2^* & d_1^* \end{bmatrix} \tag{5.4.2}$$

which is the same as the encoding scheme in [2] with $1/2$ rate and $K = 4, P = 8$.

We consider the detection for every ten consecutive transmission blocks of which the first one is used as the training block. The training overhead is therefore $1/10 = 10\%$. Each sequence of these ten transmission blocks contains $4 \times 10 = 40$ samples and carries 36 information bits.

In Figure 5-4 and Figure 5-5, we show BER performance of STBC system with DEML channel estimator under uncorrelated flat Rayleigh fading channel. For comparison, we also show the BER performance of STBC system with perfect channel state information. Same results can be found as that of the STBC system with two transmitters. The DEML channel estimator performs well in the STBC system under uncorrelated flat Rayleigh fading channel. There is small degradation in the BER performance of the STBC system with DEML channel estimator. This degradation, however, is partially because of the training block introduced in this system, which is treated as the noise signal in estimation problem. Increasing the iteration number of the iterative ML detector can improve system BER performance. Only three iterations are needed for convergence of the system BER performance to within 2 dB of the system with perfect channel state information.

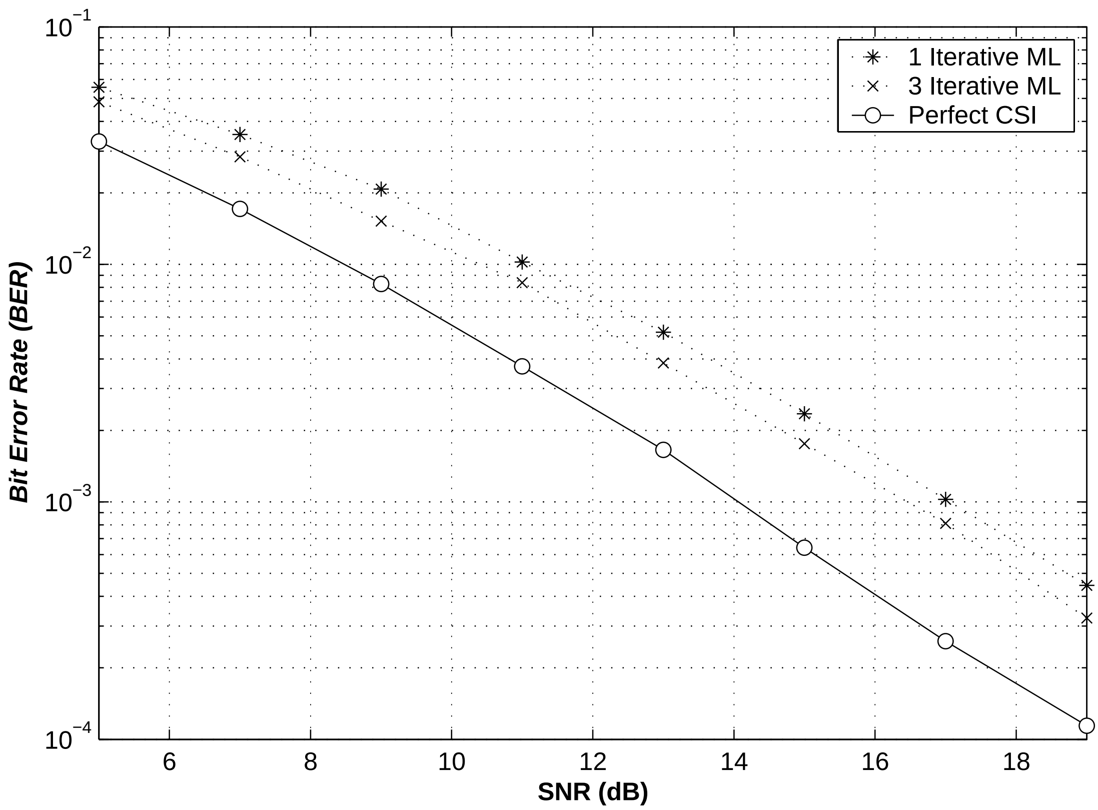

Figure 5-2: BER performance of STBC system with DEML channel estimator, two transmitters and one receiver

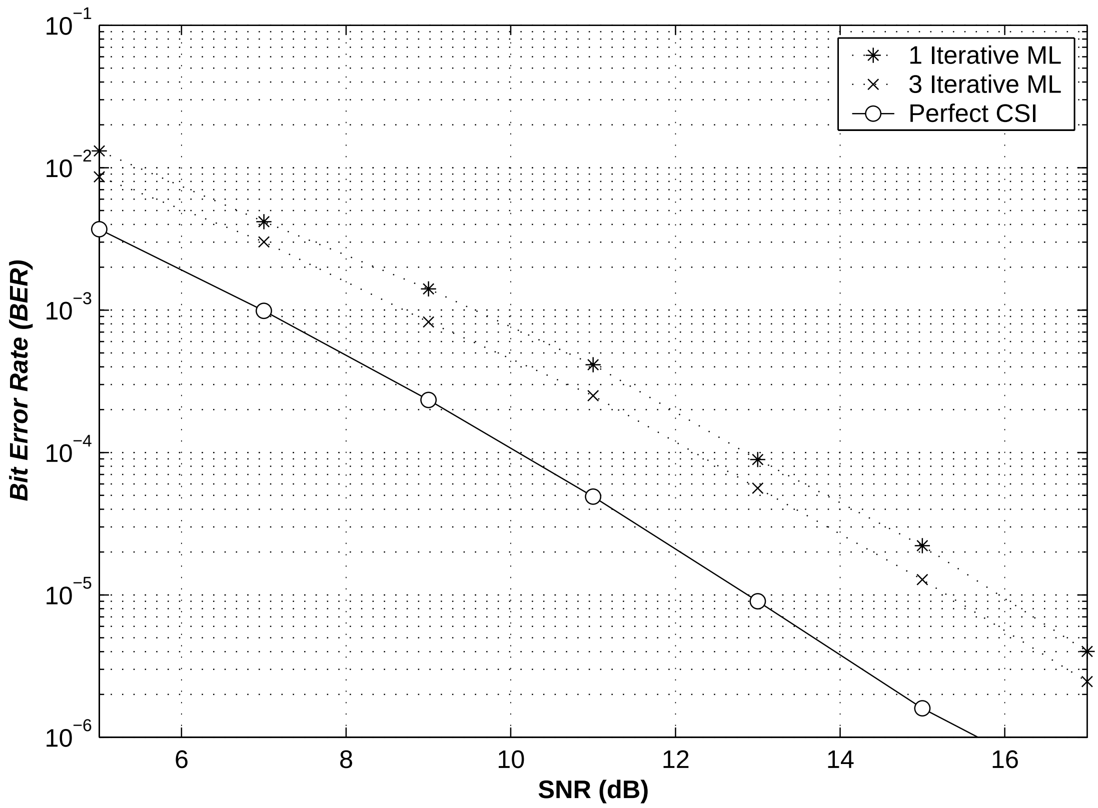

Figure 5-3: BER performance of STBC system with DEML channel estimator, two transmitters and two receivers.

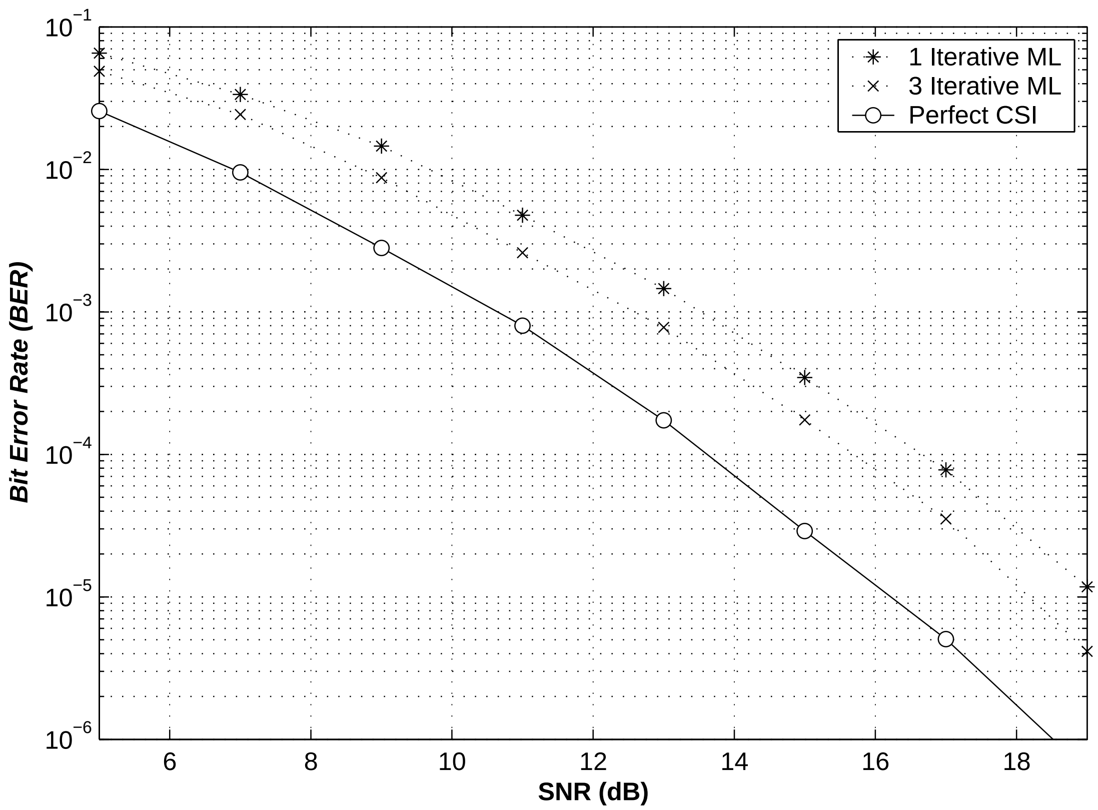

Figure 5-4: BER performance of STBC system with DEML channel estimator, four transmitters and one receiver.

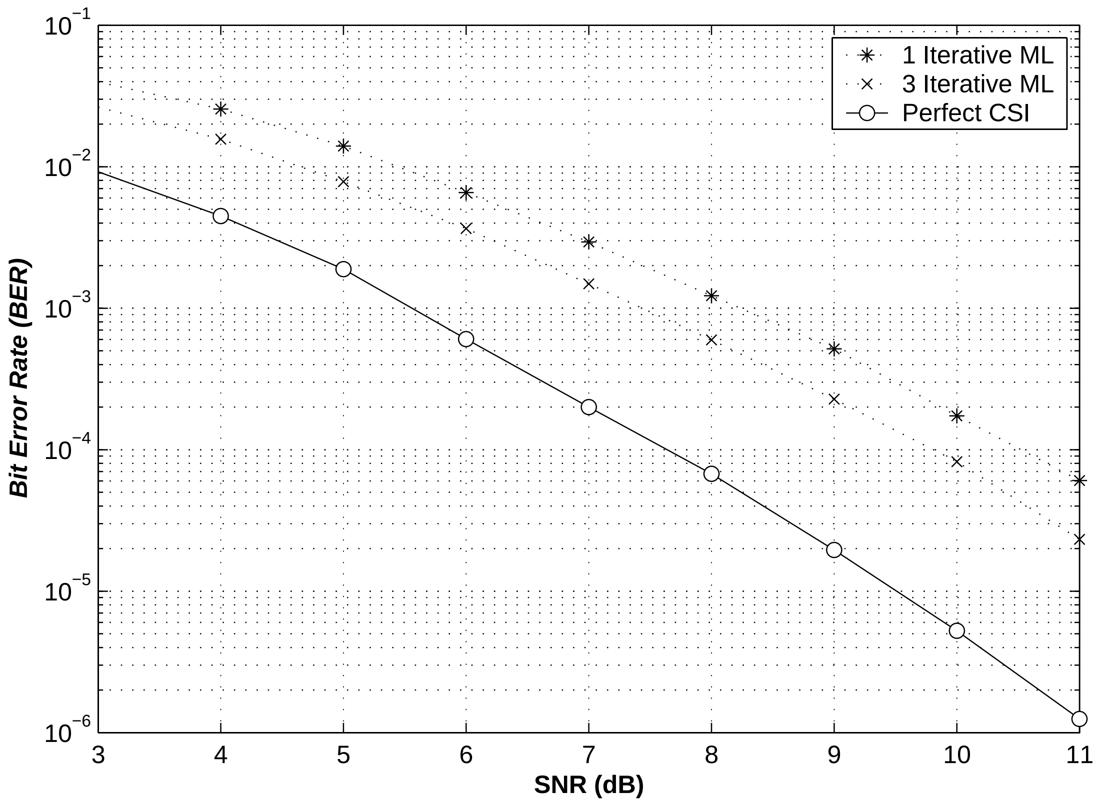

Figure 5-5: BER performance of STBC system with DEML channel estimator, four transmitters and two receivers.

CHAPTER 6

PERFORMANCE OF DEML CHANNEL ESTIMATOR UNDER SPATIALLY CORRELATED FADING CHANNEL

In this chapter, we deal with the STBC system under spatially correlated flat Rayleigh fading channel. A general procedure on the generation of correlated Rayleigh fading sequence is presented. The DEML estimator can not be applied to correlated fading channel directly. A decorrelation algorithm is introduced to this kind of STBC system before the DEML channel estimator is used. The BER performance of this kind of STBC system and some discussions are given at the last part of this chapter.

6.1 SYSTEM MODEL

The STBC system under correlated fading channel has the same system model as that of the STBC system under uncorrelated fading channel, which is described in Chapter 5. The only difference is that the fading channel is correlated. In the following, we will present a general procedure on the generation of correlated Rayleigh fading sequences.

Computer simulation of cross-correlated fading processes has become an important research topic due to the increased interest in using antenna arrays to improve cellular mobile communications. Simulators which can accurately capture the characteristics of correlated diversity channels are needed to enable realistic performance assessments of multiple antenna systems. The simulation of narrowband fading channels, in particular, requires the generation of cross-correlated Rayleigh fading sequences. Typically, the sequences must have specified auto-correlation and cross-correlation statistics. Since the desired fading coefficients are complex Gaussian variables, they can be generated in principle by factorization of the desired correlation matrix, followed by linear transformation of sequences of un-correlated variables [20, pp. 254-256]. Unfortunately, the expensive computational requirements of this direct method makes it impractical to implement.

Recently, several authors have published efficient methods of generation two [22], [23] or any number [24], [25] of cross-correlated Rayleigh fading channels. In all these approaches, independent fading processes with desired autocorrelations are first generated and then multiplied by a coloring matrix. The method was first proposed by Ertel and Reed [22] for generating two Rayleigh sequences with desired cross-correlation from two uncorrelated Rayleigh sequences each having a required autocorrelation. It was generalized and physically interpreted to model specified delay spread and frequency separation in [23]. Later on, it was extended to generate any number of cross-correlated sequences from un-correlated Rayleigh sequences by Natarajan [24] and Beaulieu [25] separately.

Let s_1 and s_2 denote the complex Gaussian samples of the Rayleigh fading signals. They can be expressed in complex format as:

$$\begin{aligned} s_1 &= s_{1i} + js_{1q} \\ s_2 &= s_{2i} + js_{2q} \end{aligned} \tag{6.1.1}$$

The envelopes of the received signals are given by:

$$r_1 = |s_1| = \sqrt{s_{1i}^2 + s_{1q}^2}$$
$$r_2 = |s_2| = \sqrt{s_{2i}^2 + s_{2q}^2} \tag{6.1.2}$$

Correlation values between s_1 and s_2 are:

$$\begin{aligned}
&E\left\{s_{1i}^2\right\} = E\left\{s_{1q}^2\right\} = E\left\{s_{2i}^2\right\} = E\left\{s_{2q}^2\right\} = \mu \\
&E\left\{s_{1i}s_{1q}\right\} = E\left\{s_{2i}s_{2q}\right\} = 0 \\
&E\left\{s_{1i}s_{2i}\right\} = E\left\{s_{1q}s_{2q}\right\} = \mu_1 \\
&E\left\{s_{1i}s_{2q}\right\} = -E\left\{s_{1q}s_{2i}\right\} = \mu_2
\end{aligned} \tag{6.1.3}$$

The normalized cross-correlation coefficient between r_1 and r_2 is expressed as [21],

$$\rho = \frac{(1+\lambda)E_i(\frac{2\sqrt{\lambda}}{1+\lambda}) - \frac{\pi}{2}}{2 - \frac{\pi}{2}} \tag{6.1.4}$$

where

$$\lambda^2 = \frac{\mu_1^2 + \mu_2^2}{\mu^2} \tag{6.1.5}$$

is the squared magnitude of the cross-correlation coefficient between s_1 and s_2, and $E_i(\eta)$ denotes the complete elliptic integral of the second kind with modulus η.

Equation (6.1.4) gives us an expression for the cross-correlation coefficient ρ of the Rayleigh faded envelopes in terms of λ, which itself is a function of the correlation properties of s_1 and s_2. We will use this relationship to determine the correlation properties of the complex Gaussian random variables that are needed to obtain the desired value of ρ.

Unfortunately, given ρ, it is not possible to solve λ from (6.1.4) in a closed form. Rather a root-finding algorithm, such as finite difference Newton's method, must be applied. The relation between ρ and λ is given in [21, Table II], and is reproduced here as Table 6-1.

ρ	λ	ρ	λ
0.00	0.00000	0.50	0.72543
0.05	0.23337	0.55	0.75922
0.10	0.32945	0.60	0.79123
0.15	0.40277	0.65	0.82168
0.20	0.46424	0.70	0.85070
0.25	0.51807	0.75	0.87842
0.30	0.56644	0.80	0.90494
0.35	0.61065	0.85	0.93033
0.40	0.65152	0.90	0.95463
0.45	0.68964	0.95	0.97787

Table 6-1: Values of ρ vs. λ

Choosing $\mu_1 = \mu_2$, the correlation matrix of $\mathbf{S} = [s_1,\ s_2]^T$ can be calculated as

$$\mathbf{R}_{\text{SS}} = \begin{bmatrix} \delta_x^2 & \frac{1}{\sqrt{2}}\lambda\delta_x^2(1-j) \\ \frac{1}{\sqrt{2}}\lambda\delta_x^2(1+j) & \delta_x^2 \end{bmatrix} \tag{6.1.6}$$

where $\delta_x^2 = 2\mu$ is the desired signal power.

Performing Cholesky decomposition on $\mathbf{R}_{\text{SS}}$, we find a lower triangular matrix $\mathbf{L}$ such that $\mathbf{R}_{\text{SS}} = \mathbf{L}\mathbf{L}^*$, where

$$\mathbf{L} = \begin{bmatrix} \delta_x & 0 \\ \frac{1}{\sqrt{2}}\lambda\delta_x(1+j) & \delta_x\sqrt{1-\lambda^2} \end{bmatrix} \tag{6.1.7}$$

is called the coloring matrix.

Assume u_1 and u_2 are two unit power uncorrelated Rayleigh fading signals. The correlation matrix for $\mathbf{U} = [u_1, u_2]^T$ is

$$\mathbf{R_{UU}} = E\{\mathbf{UU^*}\} = \mathbf{I_2} \tag{6.1.8}$$

where $\mathbf{I_2}$ denotes the 2×2 identity matrix.

Calculating $\mathbf{S = LU}$ gives the desired correlation matrix, since

$$E\{\mathbf{SS^*}\} = E\{\mathbf{LUU^*L^*}\} = \mathbf{LL^*} = \mathbf{R_{SS}} \tag{6.1.9}$$

The components of $\mathbf{U}$ are Gaussian and the components of $\mathbf{S}$ are weighted sums of $\mathbf{U}$, then $\mathbf{S}$ still has a bivariate Gaussian distribution as needed.

In summary, the procedure for generating the correlated Rayleigh fading signals is as follows:

1. From the desired correlation coefficient ρ find the appropriate value of λ using Table 6-1;
2. Specify the desired signal power δ_x^2;
3. Generate two unit power uncorrelated Rayleigh fading signals u_1 and u_2, and let $\mathbf{U} = [u_1, u_2]^T$;
4. Calculate the coloring matrix $\mathbf{L}$ using (6.1.7);
5. Calculate $\mathbf{S = LU}$, the envelopes of $\mathbf{S}$ are the desired Rayleigh faded samples.

After discussing the theoretical aspect on how to generate two correlated Rayleigh fading envelopes, some simulations are performed as follows.

Firstly, two sets of correlated Rayleigh fading envelopes are generated from the independent fading processes with desired autocorrelations. All the parameters, like ρ, λ and δ_x^2, are set to desired values. In order to obtain the relatively smooth envelop plots, a pre-designed digital Doppler filter with $f_d/T = 1/12$ is used to filter the sequence.

Secondly, these correlated Rayleigh fading sequences are used in the STBC system. The BER performance of the STBC system with different correlation coefficients ρ is shown and compared in [26].

Some selected envelope and phase plots for various ρ are given as follows. Figure 6-1a shows two cross-correlated Rayleigh distributed sequences with $\rho = 0.0$. The corresponding phase sequences for Figure 6-1a are presented in Figure 6-1b. The cross-correlated Rayleigh distributed sequences with $\rho = 0.3, 0.6, 0.9$ are shown in Figure 6-2a, Figure 6-3a and Figure 6-4a separately. The corresponding phase sequences for these cross-correlated Rayleigh distributed sequences are presented in Figure 6-2b, Figure 6-3b, and Figure 6-4b respectively. From all this diagrams we can see that the sequences with small value of ρ are less correlated both in the envelope and in the phase than those with large value of ρ.

The simulation results of STBC system (two transmitters and one receiver) with different correlation coefficients are shown in Figure 6-5. The BER performance of the STBC system under uncorrelated fading channels is shown in Figure 6-6 for comparison. From Figure 6-5 we can see that the curve marked with $\rho = 0.0$ shows BER performance for the un-correlated flat Rayleigh fading channels. It is the same as that of [1] with two transmitters and one receiver, which is shown in Figure 6-6. The curve marked with $\rho = 1.0$ shows the BER performance for full correlation. It is equivalent to a STBC system with one transmitter and one receiver, which is shown in Figure 6-6 too. For $\rho \le 0.6$, the BER performance curves are still very close to that of the system with un-correlated channels, which means they can still be treated as low correlation. Even for the deep correlation, like $\rho = 0.9$, when SNR is large enough, the BER performance is still not far away from un-correlated one.

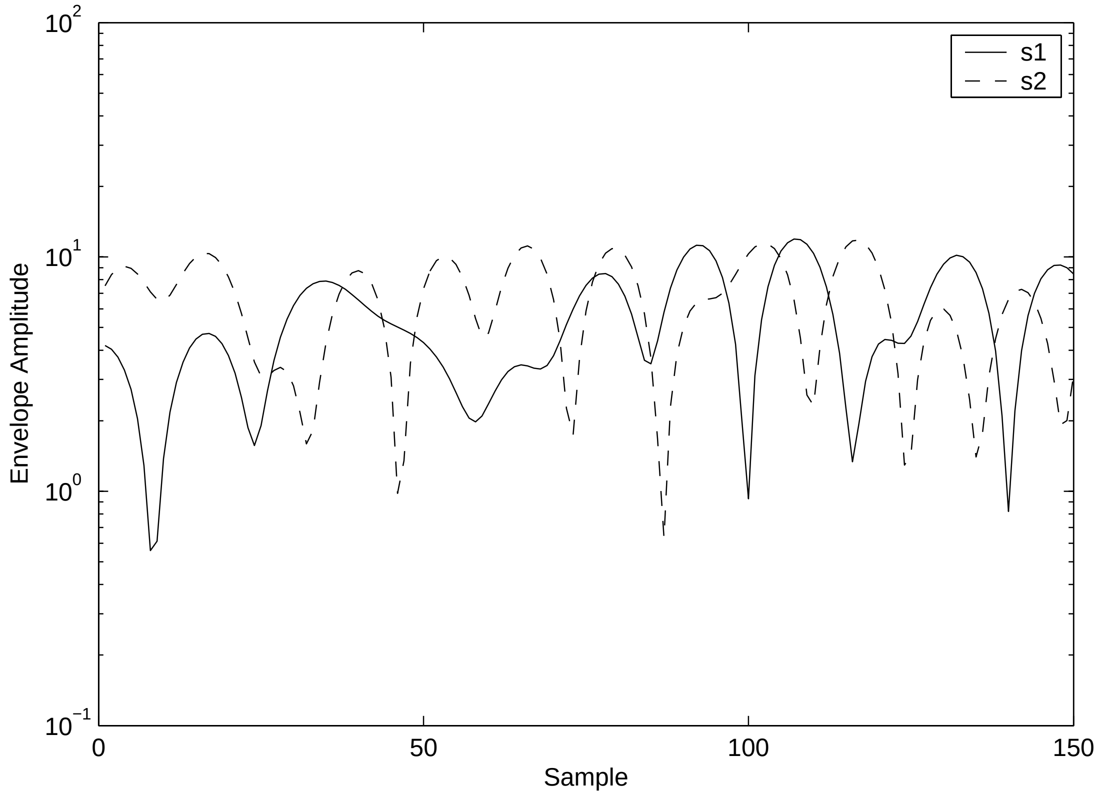

Figure 6-1a: Correlated Rayleigh Fading Envelopes (ρ = 0.0)

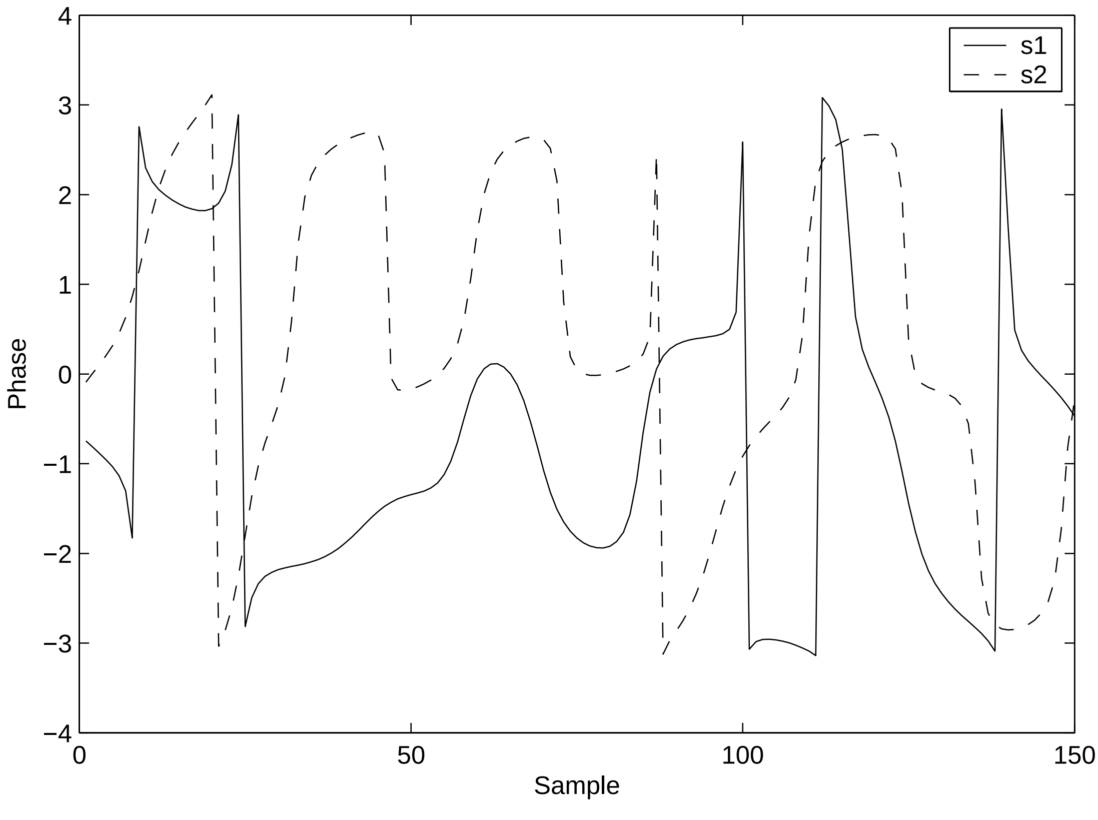

Figure 6-1b: Phases of the corresponding sample sequences (ρ = 0.0)

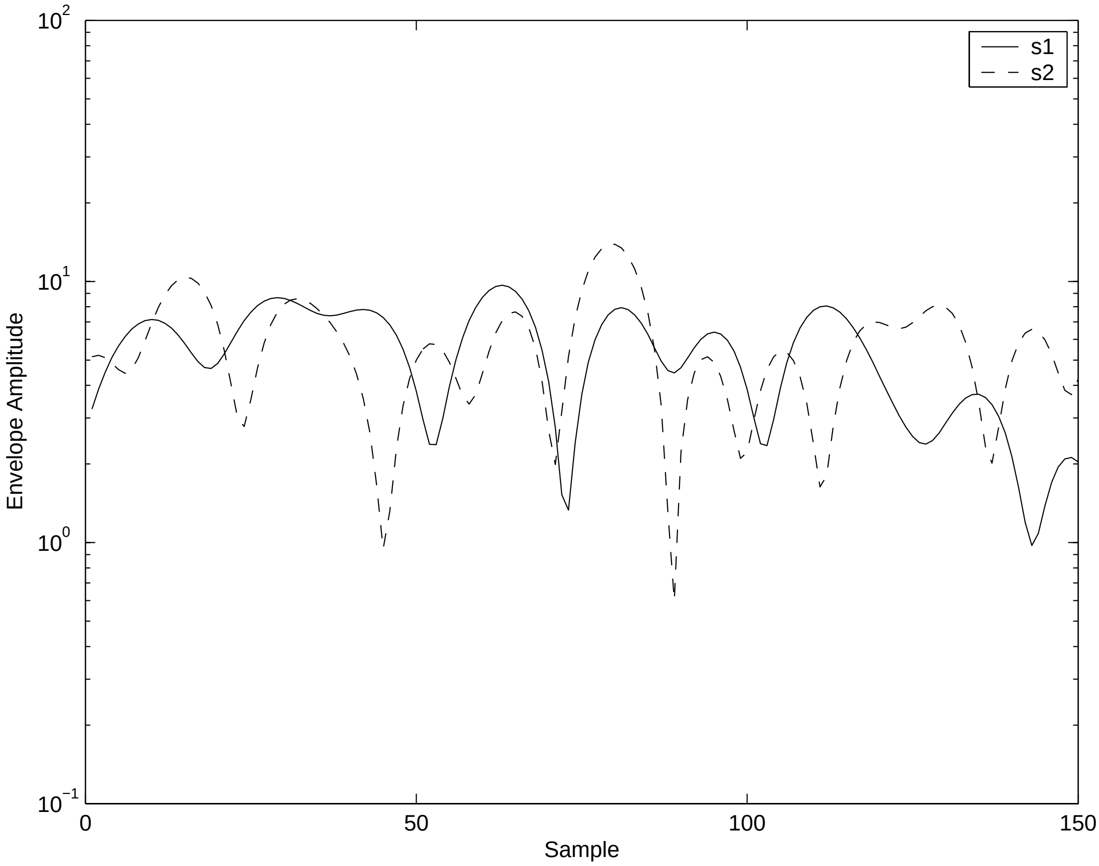

Figure 6-2a: Correlated Rayleigh Fading Envelopes (ρ = 0.3)

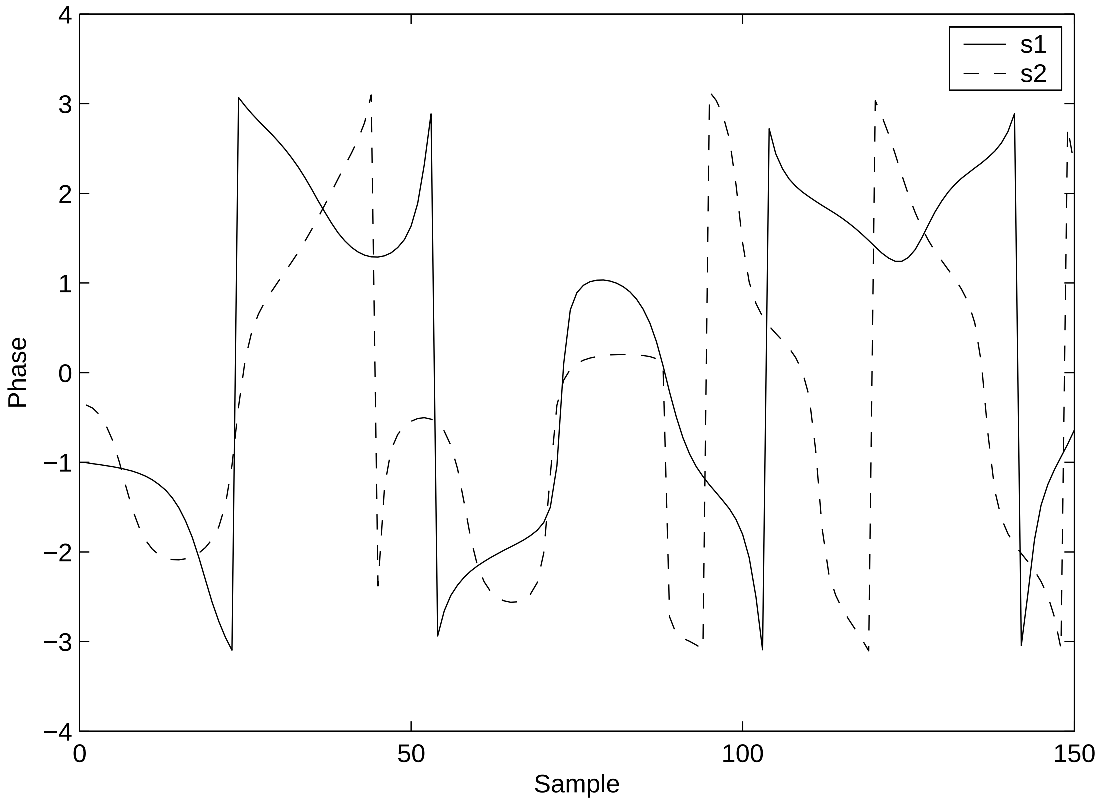

Figure 6-2b: Phases of the corresponding sample sequences (ρ = 0.3)

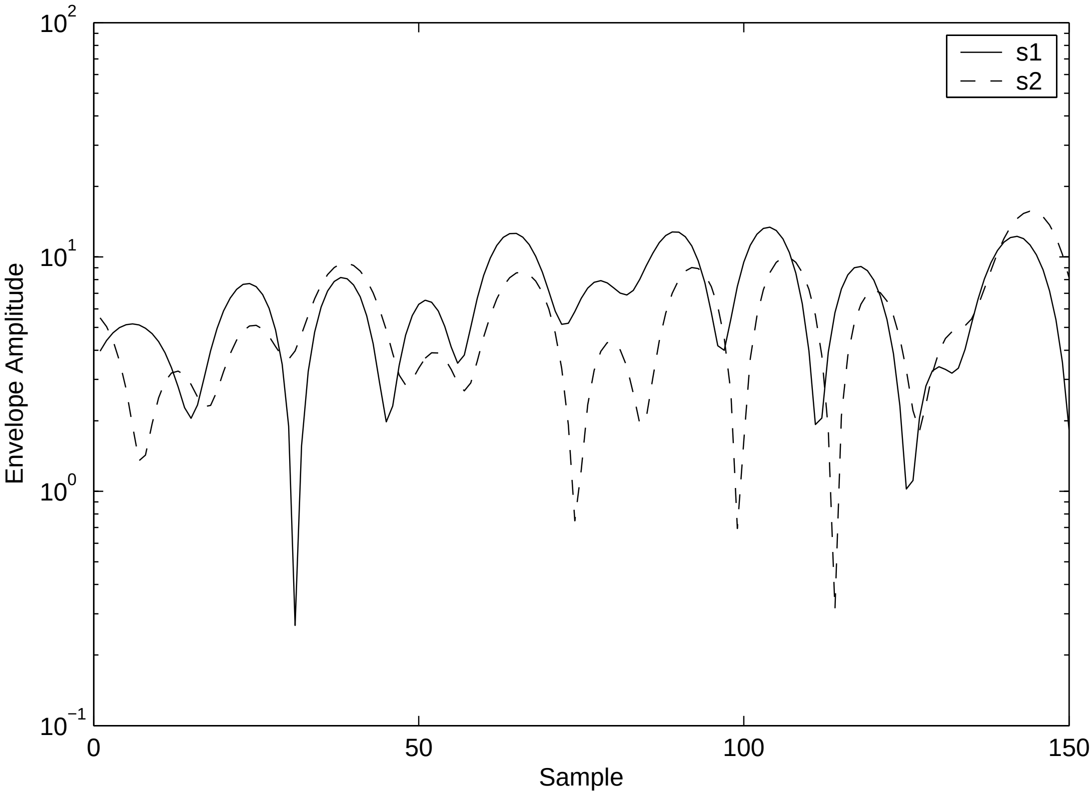

Figure 6-3a: Correlated Rayleigh Fading Envelopes (ρ = 0.6)

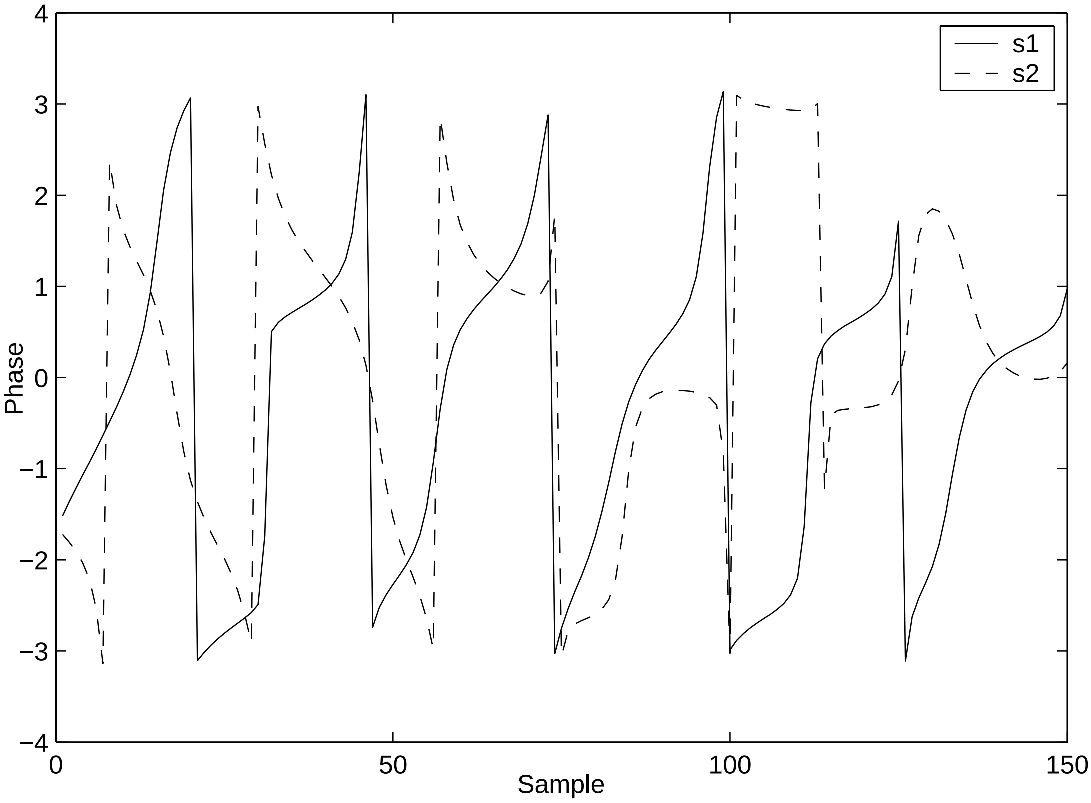

Figure 6-3b: Phases of the corresponding sample sequences (ρ = 0.6)

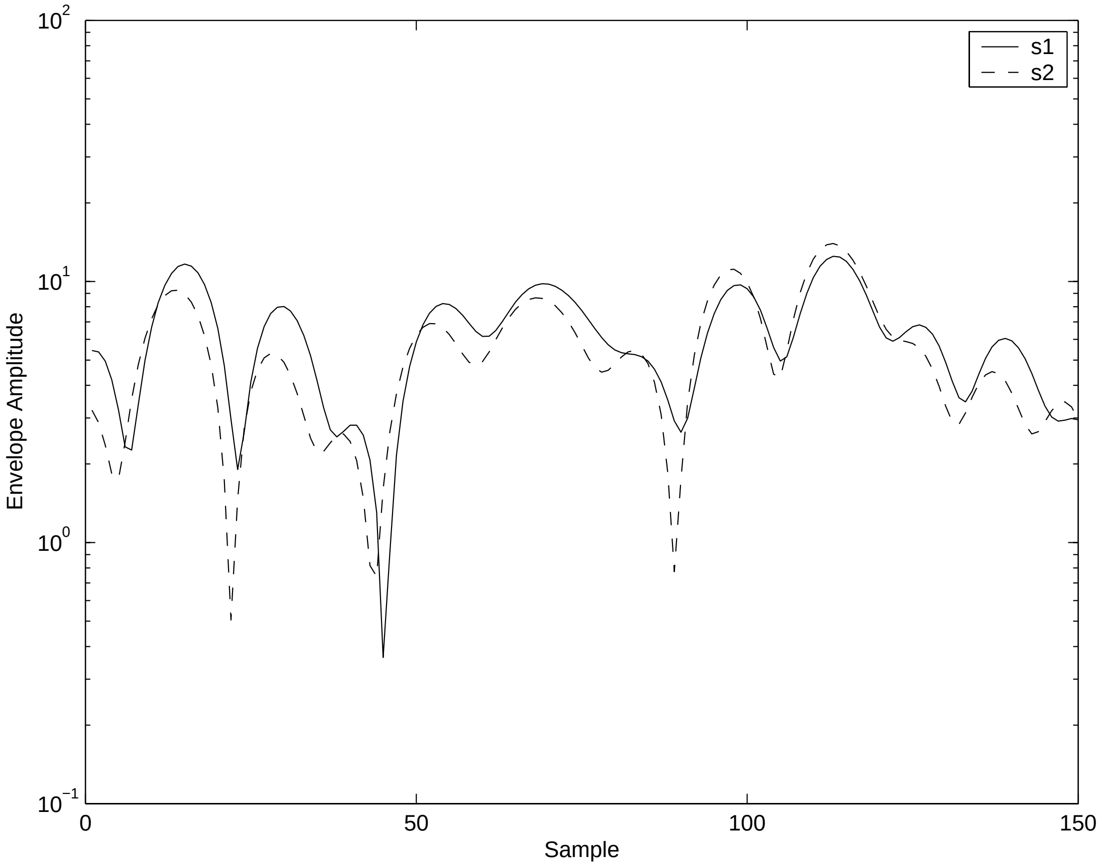

Figure 6-4a: Correlated Rayleigh Fading Envelopes (ρ = 0.9)

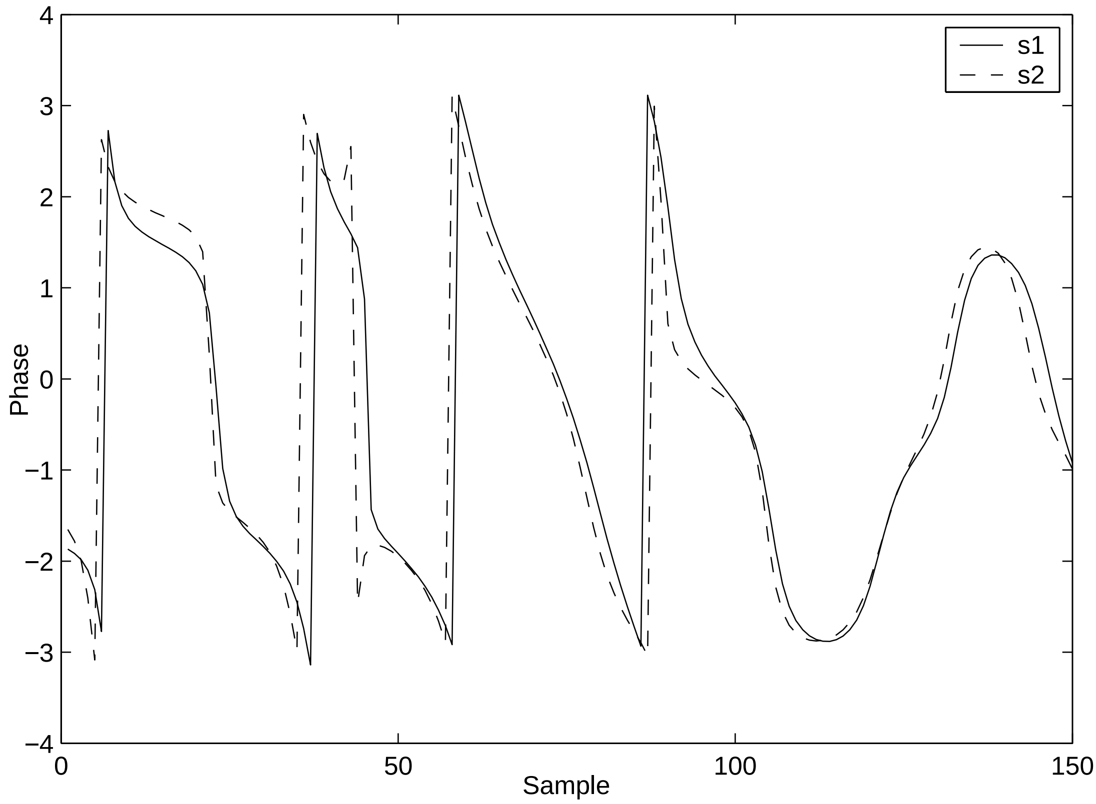

Figure 6-4b: Phases of the corresponding sample sequences (ρ = 0.9)

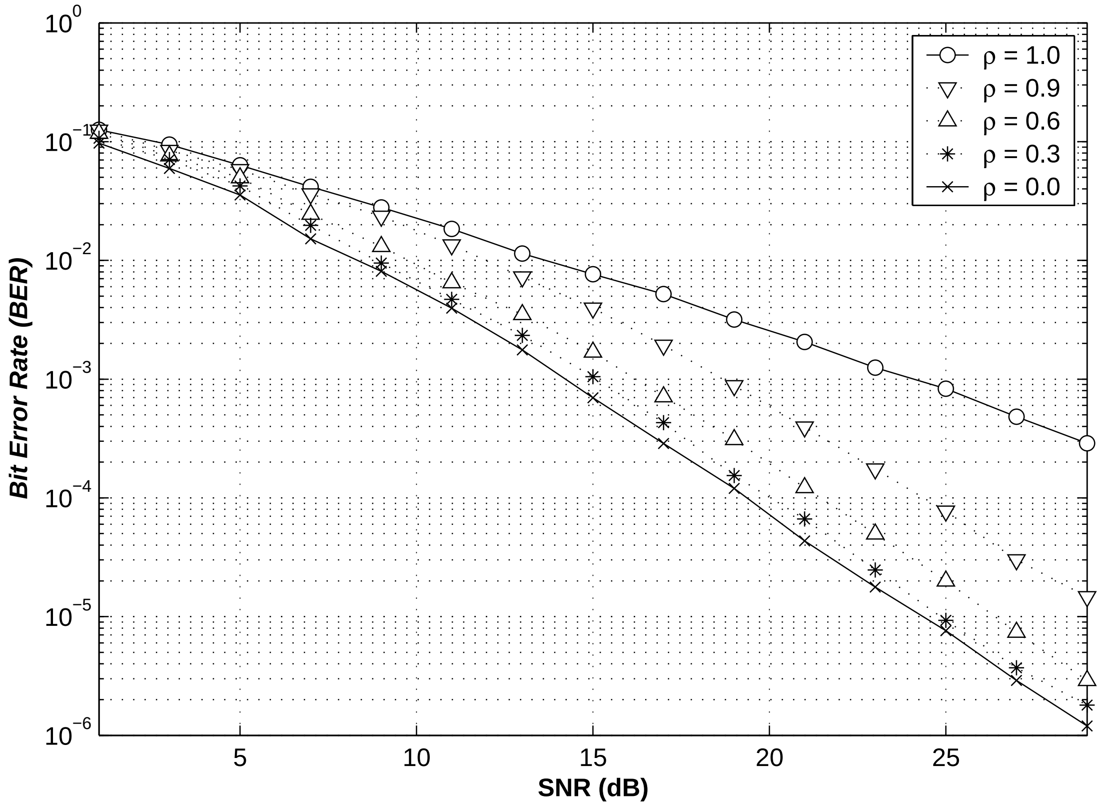

Figure 6-5: BER performance of correlated flat Rayleigh fading STBC system (two transmitters and one receiver) with different correlation coefficients.

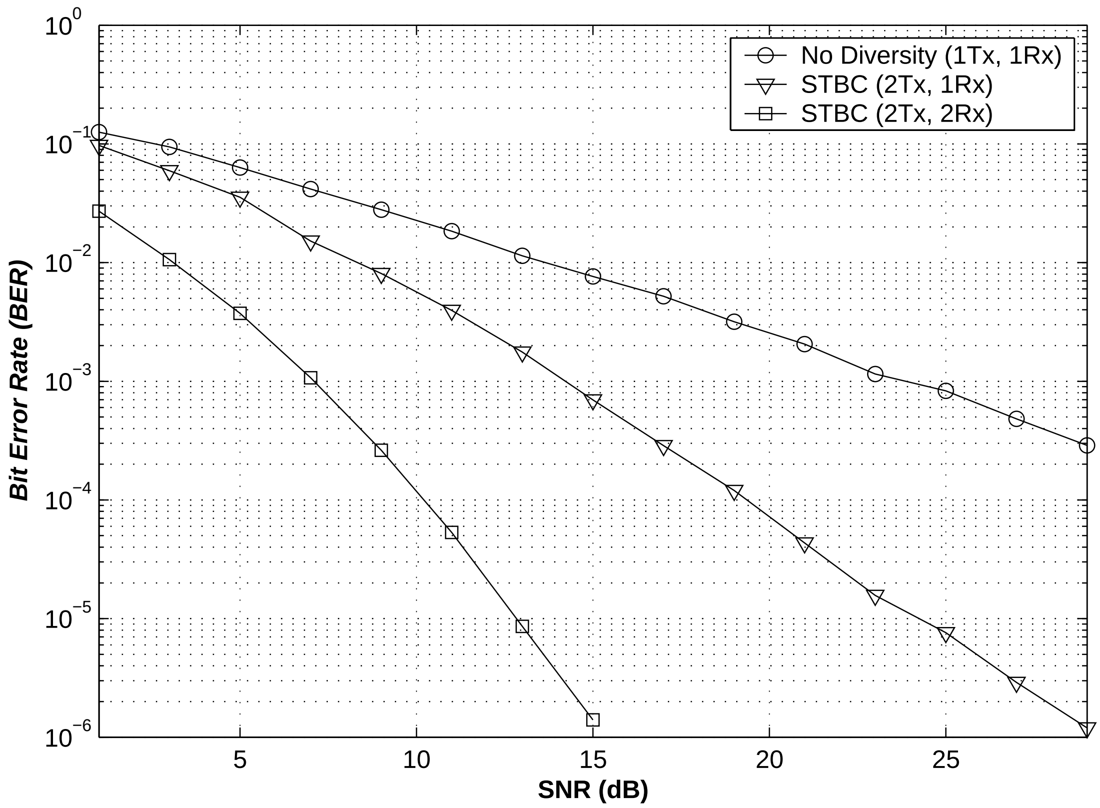

Figure 6-6: BER performance of uncorrelated flat Rayleigh fading STBC system with different number of antennas.

6.2 CHANNEL ESTIMATION

The channel estimation problem in STBC system is to determine the channel coefficients matrix **H** and the noise covariance matrix **Q** from the L independent data samples $\mathbf{X}(t_1),\mathbf{X}(t_2),...,\mathbf{X}(t_L)$.

The covariance matrix of the received signals can be calculated as

$$\mathbf{R}_{\mathrm{XX}} = E\left\{\mathbf{X}\mathbf{X}^*\right\} = \mathbf{H}\mathbf{R}_{\mathrm{SS}}\mathbf{H}^* + \sigma^2\mathbf{I} \tag{6.2.1}$$

where $\mathbf{R}_{\mathrm{SS}}$ is the covariance matrix of the transmitted signals, σ^2 is the noise covariance and $\mathbf{I}$ is the $M \times M$ unitary matrix.

We notice that $\mathbf{R}_{\mathrm{SS}}$ is diagonal when the transmitted signals are uncorrelated, non-diagonal and non-singular when the transmitted signals are partially correlated, and non-diagonal but singular when some of the transmitted signals are fully correlated (or coherent).

According to Chapter 4, the estimate of the channel coefficient matrix with DEML channel estimator is given by:

$$\hat{\mathbf{H}} = \mathbf{R}^*_{\mathrm{SX}}\mathbf{R}^{-1}_{\mathrm{SS}} \tag{6.2.2}$$

And the estimate of the noise variance matrix with DEML channel estimator is given by:

$$\hat{\mathbf{Q}} = \mathbf{R}_{\mathrm{XX}} - \mathbf{R}^*_{\mathrm{SX}}\mathbf{R}^{-1}_{\mathrm{SS}}\mathbf{R}_{\mathrm{SX}} \tag{6.2.3}$$

where

$$\mathbf{R}_{SX} = \frac{1}{L}\mathbf{S}\mathbf{X}^{*} \tag{6.2.4}$$

$$\mathbf{R}_{SS} = \frac{1}{L}\mathbf{S}\mathbf{S}^{*} \tag{6.2.5}$$

$$\mathbf{R}_{XX} = \frac{1}{L}\mathbf{X}\mathbf{X}^{*} \tag{6.2.6}$$

For the STBC system with correlated fading channel, $\mathbf{R}_{SS}$ is singular or close to be singular. It can not be used in the DEML estimator in (6.2.2) and (6.2.3) directly. Thus the DEML estimator can not be used in STBC system with correlated fading channel directly. A decorrelation algorithm must be applied to the correlated STBC system to get the modified covariance matrix before the DEML estimator is used.

6.3 DECORRELATION ALGORITHM

Received signal sequence of size L is divided into overlapping sub array signal sequences of size N , where N is the number of receive antennas, i.e. signal sequences $\{1,...,N\}$ form the first sub array, signal sequences $\{2,...,N+1\}$ form the second sub array, etc.

Let $\mathbf{X_k}$ denote the vector of received signals at the k^{th} sub array. Following the notation of (5.1.3), we can write

$$\mathbf{X_k} = \mathbf{H}\mathbf{D}^{(\mathbf{k-1})}\mathbf{S} + \mathbf{W_k} \tag{6.3.1}$$

where $\mathbf{D}^{(\mathbf{k})}$ denotes the k^{th} power of the $M \times M$ diagonal matrix and is given by [27]

$$\mathbf{D}^{(\mathbf{k})} = diag\left\{e^{-j\omega_0\tau_1},...,e^{-j\omega_0\tau_M}\right\} \tag{6.3.2}$$

The covariance matrix of the k^{th} sub array is therefore given by

$$\mathbf{R}_{\mathbf{x}_k\mathbf{x}_k} = \mathbf{H}\mathbf{D}^{(\mathbf{k}\text{-}\mathbf{1})}\mathbf{R}_{\mathrm{SS}}\left(\mathbf{D}^{(\mathbf{k}\text{-}\mathbf{1})}\right)^{*}\mathbf{H}^{*} + \sigma^2\mathbf{I} \tag{6.3.3}$$

The spatial smoothed covariance matrix is defined as the sample means of the sub array covariance:

$$\bar{\mathbf{R}}_{\mathrm{XX}} = \frac{1}{K}\sum_{k=1}^{K}\mathbf{R}_{\mathbf{x}_\mathbf{k}\mathbf{x}_\mathbf{k}} \tag{6.3.4}$$

where $K = L - N + 1$ is the number of sub arrays. Using (6.3.3), we can rewrite (6.3.4) as

$$\bar{\mathbf{R}}_{\mathrm{XX}} = \mathbf{H}\left(\frac{1}{K}\sum_{k=1}^{K}\mathbf{D}^{(\mathbf{k}\text{-}\mathbf{1})}\mathbf{R}_{\mathrm{SS}}\left(\mathbf{D}^{(\mathbf{k}\text{-}\mathbf{1})}\right)^{*}\right)\mathbf{H}^{*} + \sigma^2\mathbf{I} \tag{6.3.5}$$

or more compactly as

$$\bar{\mathbf{R}}_{\mathrm{XX}} = \mathbf{H}\bar{\mathbf{R}}_{\mathrm{SS}}\mathbf{H}^{*} + \sigma^2\mathbf{I} \tag{6.3.6}$$

where

$$\bar{\mathbf{R}}_{\mathrm{SS}} = \frac{1}{K}\sum_{k=1}^{K}\mathbf{D}^{(\mathbf{k}\text{-}\mathbf{1})}\mathbf{R}_{\mathrm{SS}}\left(\mathbf{D}^{(\mathbf{k}\text{-}\mathbf{1})}\right)^{*} \tag{6.3.7}$$

is the modified covariance matrix of the transmitted signals.

It was shown in [27] that when $K \geq M$, the modified covariance matrix $\bar{\mathbf{R}}_{\mathrm{SS}}$ will be non-singular regardless of the coherence of the transmitted signals. In this way, we get the smoothed covariance matrix. Then the DEML channel estimator as given by (6.2.2) and (6.2.3) is applied on this smoothed covariance matrix.

6.4 PERFORMANCES AND DISCUSSIONS

We consider a STBC system with two transmitters and one receiver. Simulations are done for a BPSK modulation scheme under correlated flat Rayleigh fading channel, using DEML channel estimator [26]. The transmitted signals are encoded into the 2×2 complex orthogonal design with the same encoder as given by [1]. An iterative ML detector discussed in Chapter 5 is used to improve the system performance.

We consider the detection for every ten consecutive transmission blocks of which the first one is used as the training block. The training overhead is therefore $1/10 = 10\%$. Each sequence of these ten transmission blocks contains $2 \times 10 = 20$ samples and carries 18 information bits.

In Figure 6-7, we show the BER performance of STBC system with coherent BPSK under moderately correlated flat Rayleigh fading channel. The correlation coefficient is chosen as $\rho = 0.3$. The performances of the DEML channel estimator with and without the decorrelation algorithm are shown. As we mentioned in Section 6.1, it can be treated as moderate correlation for correlation coefficient $\rho = 0.3$. We can see that the DEML channel estimator can still perform reasonably well without the decorrelation algorithm. With the decorrelation algorithm, the performance can be improved, but not that much.

In Figure 6-8, we show BER performance of STBC system with coherent BPSK under highly correlated flat Rayleigh fading channel. The correlation coefficient is chosen as $\rho = 0.9$. The performances of the DEML channel estimator with and without the decorrelation algorithm are shown. It is treated as highly correlated fading for correlation coefficient $\rho = 0.9$. We can see that the channel estimation errors can result in a significant performance loss without the decorrelation algorithm. The DEML channel estimator can not be directly used without the decorrelation algorithm. With the decorrelation algorithm, the performance of the DEML channel estimator can be greatly improved, and the channel estimation errors are almost the same as those of moderately correlated fading.

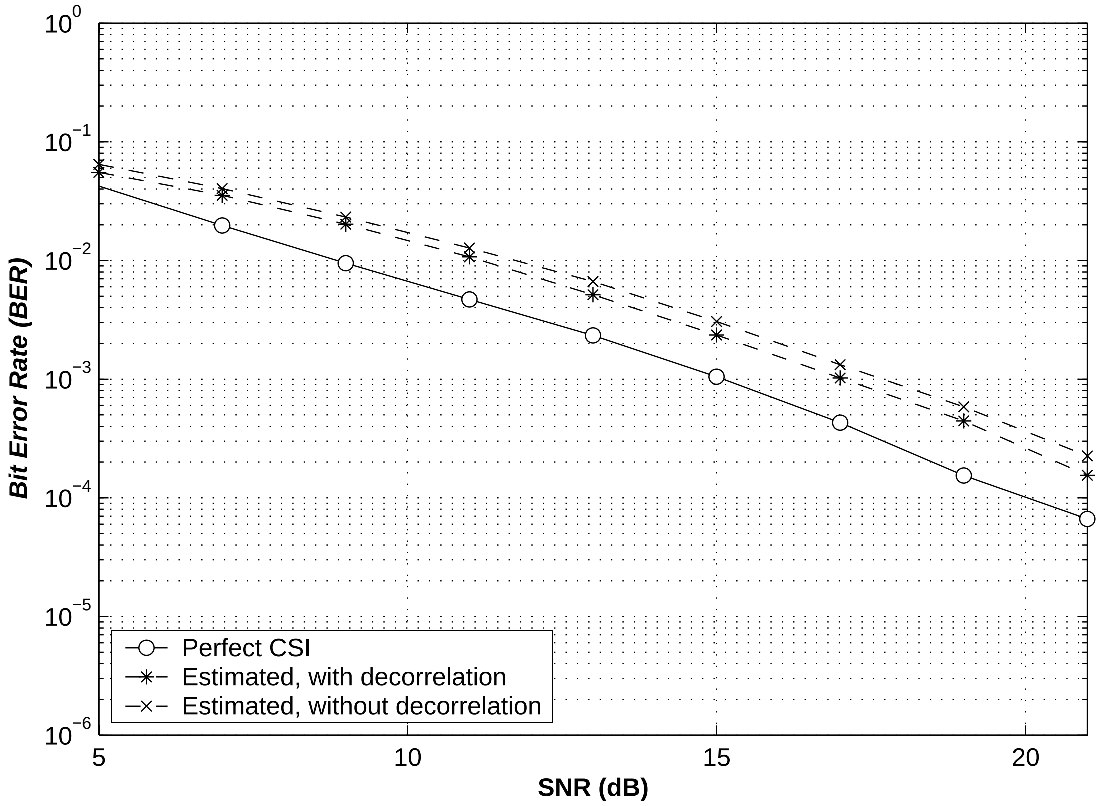

Figure 6-7: BER performance of STBC system with DEML estimator, under moderately correlated fading (ρ = 0.3).

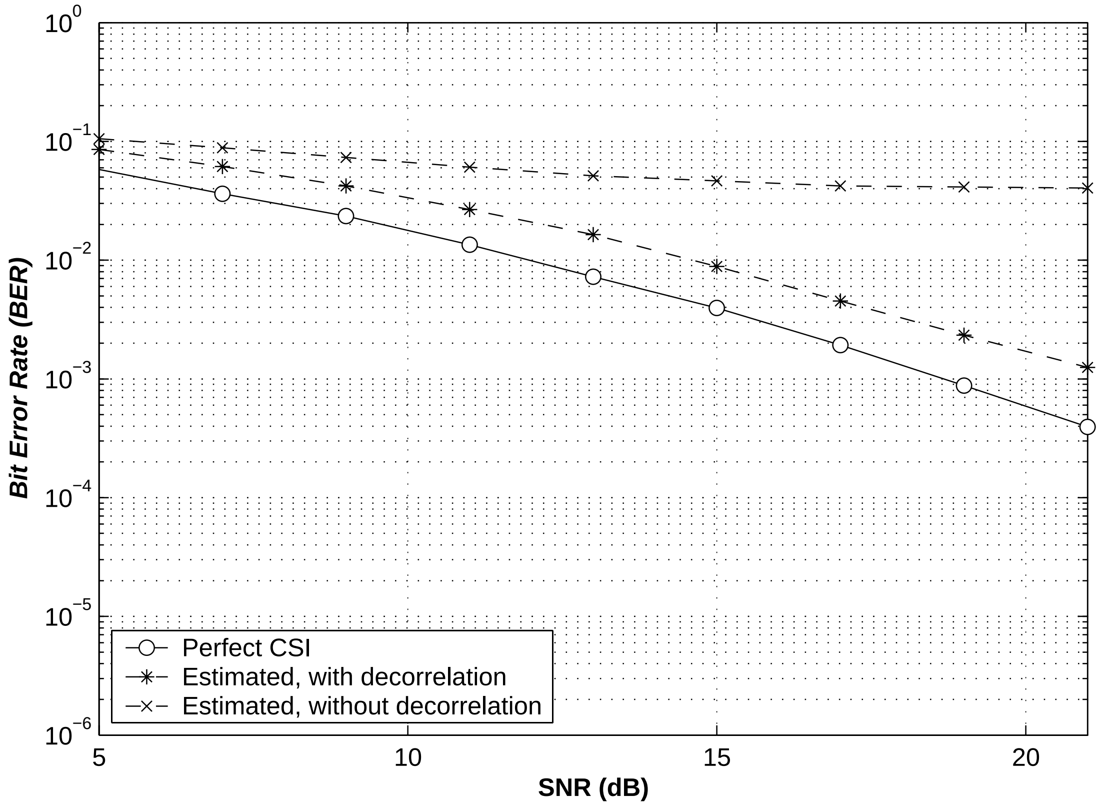

Figure 6-8: BER performance of STBC system with DEML estimator, under highly correlated fading (ρ = 0.9).

CHAPTER 7

CONCLUSION AND FUTURE WORKS

In this thesis, we have presented a computationally efficient channel estimation method for STBC system based on the decoupled maximum likelihood (DEML) algorithm. The DEML channel estimator decouples the multi-dimensional problem of the exact ML estimator into a set of one-dimensional problems and hence is computationally efficient. The BER performances of the STBC system with the DEML channel estimator both under spatially uncorrelated and correlated flat Rayleigh fading channels are shown.

If the incident signals are uncorrelated with each other, all estimates of the DEML channel estimator are consistent and large sample realizations of the ML estimates, it follows that the estimation method is asymptotically statistically efficient, according to the general properties of ML estimators.

If the incident signals are moderately correlated, the DEML estimator is no longer a large sample ML estimator. The performance of the DEML channel estimator has small degradation. But the asymptotic statistical performance is still close to that of CRB's.

If the incident signals are highly correlated, the performance of the DEML channel estimator has obvious degradation. Thus the DEML channel estimator can not be applied to the correlated STBC system directly. In such cases, a decorrelation algorithm must be applied on the correlated incident signals before the DEML channel estimator is used.

We have also obtained the BER performance of the STBC system under correlated fading channels. To study the performance of STBC system under correlated fading channels, we have presented a general method on generation of correlated Rayleigh fading sequences. In this method, independent fading processes with desired autocorrelations are first generated and then multiplied by a coloring matrix. Some selected envelope and phase plots for various ρ are given and compared. The sequences with small value of ρ are less correlated both in the envelope and in the phase than those with great value of ρ. The BER performance of STBC system with different correlation coefficients is also shown. For $\rho \leq 0.6$, the BER performance curves are still very close to that of the system with uncorrelated channels, which means they can still be treated as low correlation. Even for the deep correlation, like $\rho = 0.9$, when SNR is large enough, the BER performance is still not far away from uncorrelated one.

In addition, we have presented an iterative ML detector to improve the system BER performance with the DEML channel estimator. The iterative ML detector can obtain, after convergence, the exact ML detector in the case of unknown $\mathbf{H}$ and $\mathbf{Q}$, without increasing much more computational complexity. From the simulation results, we can see that the iterative ML detector can improve the system BER performance with the DEML channel estimator. Only few iteration numbers is required to make the system BER performance curve to converge enough.

The current work can be easily modified to accommodate cases that are more complicated. In future, more fading channel models can be considered. The performance of the DEML channel estimator can be evaluated in these fading channels and is made for comparison. Most analysis and simulations given in this thesis are based on a very simple STBC system model, which uses two transmit antennas and different number of receive antennas. In future, more complicated STBC system model can be used. The performance of the DEML channel estimator under certain type of STBC system can be analysed.

Furthermore, different space-time codes, as descried in Chapter 2, and different channel estimation methods, as described in Chapter 3, can be used. The system performance of different types of space-time coded systems with different types of channel estimators can be analysed and compared with our results.

REFERENCES

[1] S. M. Alamouti, “A simple transmit diversity scheme for wireless communications,” *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.

[2] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, “Space–time block codes from orthogonal designs,” *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456-1467, Jul. 1999.

[3] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, “Space-time block coding for wireless communications: performance results,” *IEEE J. Select. Areas Commun.*, vol. 17, pp. 451-460, Mar. 1999.

[4] V. Tarokh, N. Seshadri, and A. R. Calderbank, “Space–time codes for high data rate wireless communication: performance analysis and code construction,” *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1999.

[5] W. Firmanto, Z. Chen, B. Vucetic, J. Yuan, “Design of space-time turbo trellis coded modulation for fading channels,” *IEEE GLOBECOM ’01*, vol. 2, pp. 1093-1097, Dec. 2001.

[6] G. J. Foschini, Jr., “Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas,” *Bell Labs Tech. J.*, pp. 311-335, 1999.

[7] V. Tarokh, H. Jafarkhani, “A differential detection scheme for transmit diversity,” *IEEE, J. Select. Areas Commun.*, vol. 18, pp. 1169-1174, July 2000.

[8] B. M. Hochwald and T. L. Marzetta, “Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading,” *IEEE Trans. Inform. Theory*, vol. 46, pp. 543-564, Mar. 2000.

[9] J. Li, B. Halder, P. Stoica, and M. Viberg, “Computationally efficient angle estimation for signals with known waveforms,” *IEEE Trans. Signal Processing*, vol. 43, pp. 1995-2163, Sep. 1995.

[10] Y. Hua, “Fast maximum likelihood for blind identification of multiple FIR channels,” *IEEE Trans. Signal Processing*, vol. 44, pp. 661-672, Mar. 1996.

[11] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, “Subspace-methods for the blind identification of multichannel FIR filters,” *IEEE Trans. Signal Processing*, vol. 43, pp. 516-525, Feb. 1995.

[12] L. Tong and Q. Zhao, “Joint order detection and blind channel estimation by least squares smoothing,” *IEEE Trans. Signal Processing*, vol. 47, pp. 2345-2355, Sep. 1999.

[13] L. Tong, G. Xu, B. Hassibi, and T. Kailath, "Blind identification and equalization of multipath channels: a frequency domain approach," *IEEE Trans. Inform. Theory,* vol. 41, pp. 329-334, Jan. 1995.

[14] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: a time domain approach," *IEEE Trans. Inform. Theory,* vol. 40, pp. 340-349, Mar. 1994.

[15] K. Abed-Meraim, E. Moulines, and P. Loubaton, "Prediction error method for second-order blind identification," *IEEE Trans. Signal Processing,* vol. 45, pp. 694-705, Mar. 1997.

[16] Y. Bresler and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Trans. Audio, Speech, Signal Processing*, vol. ASSP-34, pp. 1081-1089, Oct. 1986.

[17] Y. Hua, "Fast maximum likelihood for blind identification of multiple FIR channels," *IEEE Trans. Signal Processing*, vol. 44, pp. 661-672, Mar. 1996.

[18] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via EM algorithm," *J. Royal Statist. Soci.*, vol. 39, ser. B, 1977.

[19] J. G. Sheng, A. Nallanathan and T. T. Tjhung, "Computationally efficient channel estimation for space-time block coded system," in *Proc. IEEE VTC'03*, Apr. 2003.

[20] A. Leon-Garcia, *Probability and Random Processes for Electrical Engineering*, Addison-Wesley, 1990.

[21] W.C Jakes, Ed., *Microwave Mobile Communications*, New York: IEEE Press, 1974.

[22] R. B. Ertel and J. H. Reed, “Generation of two equal power correlated Rayleigh fading envelops,” *IEEE Commun. Lett.*, vol. 2, pp. 276-278, Oct. 1998.

[23] N. C. Beaulieu, “Generation of Correlated Rayleigh Fading Envelopes,” *IEEE Commun. Lett.*, vol. 3, pp. 172-174, Jun. 1999.

[24] B. Natarajan, C. Nassar and V. Chandrasekhar, “Generation of correlated Rayleigh fading envelops for spread spectrum applications,” *IEEE Commun. Lett.*, vol. 4, pp. 9-11, Jan. 2000.

[25] N. C. Beaulieu, and M. L. Merani, “Efficient simulation of correlated diversity channels,” *IEEE WCNC 2000*, vol. 1, pp. 207-210, 2000.

[26] J. G. Sheng, A. Nallanathan and T. T. Tjhung, “Channel Estimation for Space-Time Block Coded System under Spatially Correlated Rayleigh Fading Channel,” to appear in *Proc. IEEE PIMRC’03*, Sep. 2003.

[27] T. Shan, M. Wax, and T. Kailath, “On spatial smoothing for direction-of-arrival estimation of coherent signals,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 806-811, Aug. 1985.

[28] D. Grenier and E. Bosse, “Decorrelation performance of DEESE and spatial smoothing techniques for direction-of-arrival problems,” *IEEE Trans. Signal Processing*, vol. 44, no. 6, pp. 1579-1584, Jun. 1996.

[29] C. Lagarde and D. Grenier, “Complexity-reduced direction-of-arrival estimation method for highly correlated sources,” in *Proc. IEE Radar, Sonar and Navig.*, vol. 147, no.4, pp. 157-161, Aug. 2000.

AUTHOR'S PUBLICATIONS

[1] J. G. Sheng, A. Nallanathan, and T. T. Tjhung, "Computationally efficient channel estimation for space-time block coded system," in *Proc. IEEE VTC'03*, Apr. 2003.

[2] J. G. Sheng, A. Nallanathan, and T. T. Tjhung, "Channel estimation for space-time block coded system under spatially correlated Rayleigh fading channel," to appear in *Proc. IEEE PIMRC'03*, Sep. 2003.