JOINT FREQUENCY OFFSET AND CHANNEL ESTIMATION

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

MUHAMMET AVAN

IN PARTIAL FULLFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

NOVEMBER 2008

Approval of the thesis:

JOINT FREQUENCY OFFSET AND CHANNEL ESTIMATION

submitted by MUHAMMET AVAN in partial fulfillment of the requirements for the degree of Master of Science in Electrical and Electronics Engineering Department, Middle East Technical University by,

Prof. Dr. Canan Özgen _____
Dean, Graduate School of Natural and Applied Science

Prof. Dr. İsmet Erkmen _____
Head of Department, Electrical and Electronics Engineering Dept.

Assist. Prof. Dr. Çağatay Candan _____
Supervisor, Electrical and Electronics Engineering Dept.

Examining Committee Members:

Prof. Dr. T. Engin Tuncer _____
Electrical and Electronics Engineering Department, METU

Assist. Prof. Dr. Çağatay Candan _____
Electrical and Electronics Engineering Department, METU

Assist. Prof. Dr. A. Özgür Yılmaz _____
Electrical and Electronics Engineering Department, METU

Assist. Prof. Dr. İlkay Ulusoy _____
Electrical and Electronics Engineering Department, METU

Dr. B. Ahmet Doğrusöz _____
ASELSAN A.Ş.

Date: _____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct. I have fully cited and referenced all materials and results that are not original to this work.

Name, Lastname :Muhammet AVAN

Signature :

ABSTRACT

JOINT FREQUENCY OFFSET AND CHANNEL ESTIMATION

Avan, Muhammet

M. Sc., Department of Electrical and Electronics Engineering

Supervisor: Assistant Prof. Dr. Çağatay Candan

November 2008, 106 pages

In this thesis study, joint frequency offset and channel estimation methods for single-input single-output (SISO) systems are examined. The performance of maximum likelihood estimate of the parameters are studied for different training sequences. Conventionally training sequences are designed solely for the channel estimation purpose. We present a numerical comparison of different training sequences for the joint estimation problem. The performance comparisons are made in terms of mean square estimation error (MSE) versus SNR and MSE versus the total training energy metrics. A novel estimation scheme using complementary sequences have been proposed and compared with existing schemes. The proposed scheme presents a lower estimation error than the others in almost all numerical simulations. The thesis also includes an extension for the joint channel-frequency offset estimation problem to the multi-input multi-output systems and a brief discussion for multiple frequency offset case is also given.

Keywords: Frequency offset estimation, channel estimation, maximum likelihood (ML), least sum of square error (LSSE), complementary sequences, channel impulse response (CIR)

ÖZ

FREKANS KAYMASI VE KANAL KESTİRİMİNİN BİRLİKTE YAPILMASI

Avan, Muhammet

Yüksek Lisans, Elektrik ve Elektronik Mühendisliği Bölümü
Tez Yöneticisi: Doç. Dr. Çağatay Candan

Kasım 2008, 106 sayfa

Bu tez çalışmasında tek-girişli tek-çıkışlı (SISO) kanallarda frekans kayması ve kanal kestirimini birlikte yapan yöntemler incelenmiştir. Söz konusu parametrelerin en yüksek olasılık yöntemiyle kestirim performansları üzerinde farklı alıştırma serileri kullanılarak çalışılmıştır. Çalışılan alıştırma serileri yalnızca kanal kestirimi için tasarlanmış serilerdir. Bu tez kapsamında, farklı alıştırma serilerinin ortak kestirim problemi için sayısal karşılaştırmaları sunulmuştur. Performans karşılaştırmaları ortalama karesel hataya karşılık SNR ve ortalama karesel hataya karşılık toplam alıştırma serisi enerjisi şeklinde gerçekleştirilmiştir. Tamamlayıcı serileri kullanan yeni bir kestirim yöntemi önerilmiş ve bu yöntem var olan metotlarla karşılaştırılmıştır. Önerilen metot, sayısal simülasyonların hemen hepsinde diğer metotlardan daha düşük kestirim hatası sergilemiştir. Tez, çok-girişli çok-çıkışlı (MIMO) sistemlerde frekans kayması ve kanal kestiriminin birlikte yapılmasına yönelik bir ek bölüm içermekte ve kanaldaki frekans kayması sayısının birden çok olması durumu için kısa bir giriş yapmaktadır.

Anahtar sözcükler: Kanal kayması kestirimi, kanal kestirimi, en yüksek olasılık, en küçük hata kareleri toplamı, tamamlayıcı seriler, kanal karakteristiği

ACKNOWLEDGEMENT

I would like to express my deepest gratitude to my supervisor Assist. Prof. Dr. Çağatay Candan for his encouragement, guidance, advice, criticism and insight throughout the research.

I would like to express my infinite gratitude to my parents for their endless support in every phase of my life and their consolations in my distressed days.

I would like to thank my colleagues in Communication ES/EA System Engineering Department of ASELSAN for their tolerance during the period of the preparation of this thesis.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Communication systems may suffer from intersymbol interference (ISI) which is due to the frequency selective nature of the channel. ISI is caused by the symbol duration being much smaller than the delay spread of the channel [4]. A second detrimental effect is the carrier frequency offset between transmitter and receiver. The main reason of frequency offset is the poor synchronization between the local oscillator in the transmitter and the local oscillator in the receiver. Loss of synchronization can be caused by frequency drifts in the local oscillators. However, a more significant effect is the frequency offset due to Doppler effect which is seen in the mobile communication systems. Furthermore the frequency offset due to Doppler-shift can be at different amounts for each path of the frequency selective channel [9].

In order to combat ISI one method is to recover the channel impulse response (CIR) using various channel estimation algorithms. These algorithms can be divided into three groups such as “blind”, “semiblind” and “training based” methods. Training based channel estimation algorithms are based on the usage of the knowledge of a known pilot symbol inserted into the channel. The pilot symbols are generally transmitted in the preamble phase of communication [1], [2]. Blind channel estimation algorithms use stochastic or other characteristics properties of the unknown user data to recover CIR. Semiblind algorithms use both knowledge of the known pilot symbol and properties of the unknown user data, [4]. Although blind and semiblind algorithms are more recent and have shown an improvement in

channel estimation, the training based algorithms are still preferred in use because of their reliability and ease of implementation.

Most of the frequency offset estimation algorithms are semiblind algorithms which use knowledge of known pilot symbols, unknown user data and noise in the channel. Some of the algorithms profit from pilot carriers to recover frequency offset with line spectral estimation methods like MUSIC and ESPRIT in OFDM communication systems, [21], [22]. In SISO (single input single output) and MIMO (multi input multi output) communication systems foremost frequency offset algorithms are methods which use maximum likelihood (ML) estimation schemes. These algorithms utilize deterministic characteristics of a known training sequence and stochastic characteristics of the noise in the channel [5], [8], [9], [10]. ML algorithms also determine the CIR jointly together with frequency offset.

Training sequence selection is an important issue in training based frequency offset and channel estimation algorithms. Most of the channel estimation methods need training sequences that show impulsive autocorrelation characteristics in order to minimize estimation error. Although frequency offset estimation methods have different requirement in training sequence selection, joint frequency offset and channel estimation algorithms generally prefer those have impulsive autocorrelation characteristics. Although, some of the estimation algorithms develop their optimum training sequences based on their estimation method [1], [2], [3], [8], for some of them the optimum sequences have to be determined [5].

The complementary sequences which are firstly introduced by Golay, [11], [12] have complementary autocorrelations. That is, the sum of autocorrelations sequence A and its mate sequence B is zero for all autocorrelation lags except the one at the origin. The sum of the autocorrelation samples of a complementary pair has impulsive autocorrelation characteristics. Complementary sequences recently have been begun to use in channel estimation and resulted in an improvement in estimation error [7], [14]. In frequency offset estimation, complementary sequences have not been applied yet in the literature.

In this thesis some of the frequency offset and channel estimation methods given in the literature have been reviewed and compared. In this scope, the basic channel estimation algorithms like least sum of square error (LSSE), and maximum

likelihood (ML) have been analyzed [1], [4]. In addition, some channel estimation schemes which use complementary training sequences have been examined [7], [14]. In the scope of this study, some maximum likelihood (ML) based frequency offset estimation methods for SISO and MIMO communication systems, in presence of single and multiple frequency offset, have been analyzed [5], [8], [9], [10]. Using the basic properties of ML estimation scheme in [5], ML estimatior for joint frequency offset and channel response using complementary sequences have been derived. New scheme has also been applied to some special cases of MIMO systems. The performances of the complementary sequences and different training sequences, which are known to be optimal for channel estimation methods, on joint frequency offset and channel estimation have been determined and compared using computer simulations.

This thesis is organized as follows: In Chapter 2, basic training based channel estimation methods and basic properties of the complementary sequences have been introduced. In Chapter 3, ML frequency offset and channel estimation algorithms for SISO systems are presented and ML joint frequency offset and channel estimation scheme using complementary training sequences is proposed. An ad-hoc technique is also examined. In Chapter 4, the results of numerical simulations of different techniques under different criterions are examined. and finally in Chapter 5, a conclusions derived from this study is presented along with a short summary

CHAPTER 2

BASIC CONCEPTS

In this chapter, we introduce signal model and review some channel estimation approaches given in the literature. This chapter establishes a base for the following chapters. In Section 2.1, the signal model for finite impulse response linear time-invariant communication channels is presented. In Section 2.2, conventional training based channel estimation methods are explained. In Section 2.3 the training sequence placement strategies have been examined. Finally, in section 2.4, complementary sequences and some of their mathematical properties are examined.

2.1 Signal Model for Finite Impulse Response Time-Invariant Communication Channel

A SISO (Single Input Single Output) communication channel employs a single transmit and a single receive antenna. During the transmission through a SISO channel the transmitted signal can propagate through different paths which have different complex channel gain. A SISO communication channel can be modeled as a finite impulse response (FIR) filter of length L.

$$\boldsymbol{h} = [h[0]\ h[1]\ \dots\ h[L-1]]^T \tag{2-1}$$

The communication channels which have more than one tap, that is channels with have multiple paths between transmitter and receiver, are called as frequency-selective channels. If the channel has only one propagation path, it is called as flat-

fading channel. It should be noted that channel characteristics also depends on the symbol duration in addition to the length of the channel impulse response, [19].

Assuming the transmission of a length of N sequence, $x[n]$, over the channel and the coherence time of the channel is larger than the duration of the transmitted signal, the observed signal at the output at the receiver can be modeled as in Eq. (2-2).

$$y[n] = \sum_{l=0}^{L-1} h[l]x[n-l] + w[n] \tag{2-2}$$

As can be observed from Eq. (2-2), the received signal samples, $y[n]$,are convolution of channel impulse response with transmitted signal samples, $x[n]$. The convolution of channel impulse response and transmitted signal is corrupted by an additive white Gaussian noise (AWGN) with zero-mean and variance σ_w^2.

2.2 Training Based Channel Estimation Methods

2.2.1 Correlation Method

Correlation method is the simplest method for channel estimation. It relies on matched filtering concept.

In correlation method time-reversed and conjugated version of training sequence is convolved with received signal from the channel.

$$\hat{h}[n] = y[n] * x^*[-n] \quad n = 0, 1, \dots, N-1 \tag{2-3}$$

$$\hat{h}[n] = \sum_{k=0}^{N-1} \left(\sum_{l=0}^{L-1} h[l]x[k-l] + w[n] \right) x^*[k-n] \tag{2-4}$$

$$\hat{h}[n] = \sum_{l=0}^{L-1} h[l] \sum_{k=0}^{N-1} x[k-l]x^*[k-n] + \sum_{n=0}^{N-1} w[n]x^*[k-n] \tag{2-5}$$

$$\hat{h}[n] = \sum_{l=0}^{L-1} h[l] r_{xx}[n-l] + \sum_{n=0}^{N-1} w[n] x^*[k-n] \tag{2-6}$$

For correct estimation of channel impulse response correlation method requires an noise-like training sequence, so that $r_{xx}(k) = \delta(k)$. If this type of training sequence is used the result becomes channel impulse response itself summed with second addition in right hand side of Eq. (2-6).

$$\hat{h}[n] = h[n] + \sum_{n=0}^{N-1} w[n] x^*[k-n] \quad n = 0, 1, \dots, N-1 \tag{2-7}$$

In Eq. (2-7), last $N - L$ elements of $h[n]$ are zero. Using Eq. (2-7) error for estimation is obtained as in Eq. (2-8) and Eq. (2-9).

$$e[n] = \hat{h}[n] - h[n] \tag{2-8}$$

$$e[n] = \sum_{n=0}^{N-1} w[n] x^*[k-n] \quad n = 0, 1, \dots, N-1 \tag{2-9}$$

2.2.2 Least Sum of Squared Error (LSSE) Method

Least Sum of Squared Errors (LSSE) algorithm is one of the basic methods in the training based channel estimation literature. It is used to estimate the impulse response of the channel using the short preamble training sequences. The LSSE method uses the deterministic autocorrelation matrix of the training sequence for estimation purposes. In order to estimate the channel, the LSSE requires a preamble which consists of the training sequence and a precursor of length of $L - 1$. The structure of the preamble is shown in Figure 2-1 and the mathematical expression is given in Eq. (2-10), [1].

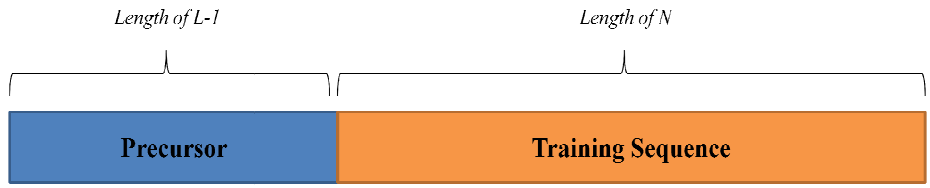


Figure 2-1: Preamble for LSSE Method

$$\boldsymbol{x} = [x[-L+1] \quad x[-L+2] \quad \dots \quad x[0] \quad x[1] \quad \dots \quad x[N-1]]^T \tag{2-10}$$

In Eq. (2-10) the preamble sequence, $\boldsymbol{x}$, consists of $L-1$ precursor and a training sequence length of N. The length of training sequence, N, is greater than the number of the channel taps, L. As a conclusion, the preamble has $N+L-1$ number of symbols.

$$N_p = N + L - 1 \tag{2-11}$$

After transmitting the preamble from the transmit antenna, the LSSE is applied on the observed signal samples in the receive side. The mathematical expression for observed received signal samples vector is expressed as in Eq. (2-12).

$$\boldsymbol{y} = [y[0] \quad y[1] \quad \dots \quad y[N-1]]^T \tag{2-12}$$

The vector given by Eq. (2-12) contains N observed samples which are the convolution of the preamble sequence and channel impulse response in the absence of AWGN. The form of each received sample is given by Eq. (2-13).

$$y[n] = \sum_{l=0}^{L-1} h[l]x[n-l] + w[n]$$

$$= \boldsymbol{x}[n]^{\boldsymbol{T}}\boldsymbol{h} + \boldsymbol{w}[n] \qquad n = 0, 1, \dots, N-1 \tag{2-13}$$

In Eq. (2-13) subscript T denotes transpose of channel impulse response vector **h** provided by Eq. (2-14).

$$\boldsymbol{h} = [h[0] \quad h[1] \quad \dots \quad h[L-1]]^T \tag{2-14}$$

Using Eq. (2-10), and Eq. (2-14), the Eq. (2-13) can be expressed as

$$y[n] = [x[n] \quad x[n-1] \quad \dots \quad x[n-L+1]] \begin{bmatrix} h[0] \\ h[1] \\ \vdots \\ h[L-1] \end{bmatrix} + w[n] \tag{2-15}$$

Substituting Eq. (2-15) into Eq. (2-12), the observed signal samples vector can be defined as

$$\boldsymbol{y} = \begin{bmatrix} x[0] & x[-1] & \dots & x[-L+1] \\ x[1] & x[0] & \dots & x[-L+2] \\ \vdots & \vdots & \ddots & \vdots \\ x[N-1] & x[N-2] & \dots & x[N-L] \end{bmatrix} \begin{bmatrix} h[0] \\ h[1] \\ \vdots \\ h[L-1] \end{bmatrix} + \begin{bmatrix} w[0] \\ w[1] \\ \vdots \\ w[N-1] \end{bmatrix} \tag{2-16}$$

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{h} + \boldsymbol{w} \tag{2-17}$$

where $\boldsymbol{X}$ is N-by-L matrix containing symbols in the preamble.

The channel characteristics can be estimated by minimizing the sum of the squared error between the observed signal, $\boldsymbol{y}$, and the estimation of it, $\hat{\boldsymbol{y}}$. The sum of the squared error (sse) for a channel estimation process can be defined as

$$sse = |\boldsymbol{y} - \hat{\boldsymbol{y}}|^2 = \sum_{n=0}^{N-1} |\boldsymbol{y}[n] - \hat{\boldsymbol{y}}[n]|^2 \tag{2-18}$$

where

$$\hat{y}[n] = \sum_{l=0}^{L-1} \hat{h}[l] x[n-l]$$

$$= \boldsymbol{x}[n]^T \hat{\boldsymbol{h}} \quad n = 0, 1, \dots, N-1 \tag{2-19}$$

and $\hat{\boldsymbol{h}}$ is the vector of estimated channel impulse response.

Using Eq. (2-19), the estimated observation vector is expressed as

$$\hat{\boldsymbol{y}} = \begin{bmatrix} \hat{y}[0] \\ \hat{y}[1] \\ \vdots \\ \hat{y}[N-1] \end{bmatrix} = \begin{bmatrix} x[0] & x[-1] & \dots & x[-L+1] \\ x[1] & x[0] & \dots & x[-L+2] \\ \vdots & \vdots & \ddots & \vdots \\ x[N-1] & x[N-2] & \dots & x[N-L] \end{bmatrix} \begin{bmatrix} \hat{h}[0] \\ \hat{h}[1] \\ \vdots \\ \hat{h}[L-1] \end{bmatrix} \tag{2-20}$$

$$\hat{\boldsymbol{y}} = \boldsymbol{X}\hat{\boldsymbol{h}} \tag{2-21}$$

Substituting Eq. (2-21) into Eq. (2-18) and rearranging gives

$$sse = \boldsymbol{y}^H\boldsymbol{y} - 2\hat{\boldsymbol{h}}^H\boldsymbol{X}^H\boldsymbol{y} + \hat{\boldsymbol{h}}^H\boldsymbol{X}^H\boldsymbol{X}\hat{\boldsymbol{h}} \tag{2-22}$$

where the subscript H denotes the Hermitian of a matrix or vector.

Rearranging Eq. (2-22) the sse can be reexpressed as

$$sse = E_y - 2\hat{\boldsymbol{h}}^H\boldsymbol{r}_{yx} + \hat{\boldsymbol{h}}^H\boldsymbol{R}_{xx}\hat{\boldsymbol{h}} \tag{2-23}$$

In Eq. (2-23), E_y is the energy of observed signal, $\boldsymbol{y}$, given by Eq. (2-24), and r_{yx} is an L-dimensional column vector given by Eq. (2-25)

$$E_y = \boldsymbol{y}^H\boldsymbol{y} = |\boldsymbol{y}|^2 = \sum_{n=0}^{N-1} |y[n]|^2 \tag{2-24}$$

$$\boldsymbol{r}_{yx} = \boldsymbol{X}^H\boldsymbol{y} = \begin{bmatrix} x[0]^* & x[1]^* & \dots & x[N-1]^* \\ x[-1]^* & x[0]^* & \dots & x[N-2]^* \\ \vdots & \vdots & \ddots & \vdots \\ x[-L+1]^* & x[-L+2]^* & \dots & x[N-L]^* \end{bmatrix} \begin{bmatrix} y[0] \\ y[1] \\ \vdots \\ y[N-1] \end{bmatrix} \tag{2-25}$$

$\boldsymbol{R}_{xx}$ is a correlation matrix produced from the preamble sequence. It is assumed to be positive definite and invertible. Also $\boldsymbol{R}_{xx}$ is Hermitian i.e $\boldsymbol{R}_{xx}^T = \boldsymbol{R}_{xx}^*$.

$$\boldsymbol{R}_{xx} = \boldsymbol{X}^H\boldsymbol{X} \tag{2-26}$$

The optimum channel impulse response estimation can be achieved by taking the derivative of the sum of the squared error (sse) given by Eq. (2-23) with respect to $\hat{\boldsymbol{h}}^H$ and equating the derivative to zero.

$$\frac{\partial sse}{\partial \hat{\boldsymbol{h}}^H} = -2\boldsymbol{r}_{yx} + 2\boldsymbol{R}_{xx}\hat{\boldsymbol{h}} = 0 \tag{2-27}$$

$$\widehat{\boldsymbol{h}} = \boldsymbol{R}_{xx}^{-1}\boldsymbol{r}_{yx} \tag{2-28}$$

Inserting Eq. (2-25) and Eq. (2-26) into the Eq. (2-28), the optimum channel estimation given by Eq. (2-29) is achieved.

$$\widehat{\boldsymbol{h}} = (\boldsymbol{X}^H\boldsymbol{X})^{-1}\boldsymbol{X}^H\boldsymbol{y} \tag{2-29}$$

Inserting Eq. (2-29) into Eq. (2-23) the least sum of squared error (LSSE) can be given as in Eq. (2-30)

$$lsse = E_y - \boldsymbol{r}_{yx}^H\boldsymbol{R}_{xx}^{-1}\boldsymbol{r}_{yx} \tag{2-30}$$

Eq. (2-29) gives the solution for the channel estimation using the least sum of squared error (LSSE) method. Inserting the definition of observation given in Eq. (2-13) into Eq. (2-29) and rearranging gives Eq. (2-31)

$$\begin{aligned}\widehat{\boldsymbol{h}} &= (\boldsymbol{X}^H\boldsymbol{X})^{-1}\boldsymbol{X}^H(\boldsymbol{X}\boldsymbol{h} + \boldsymbol{w}) \\ &= (\boldsymbol{X}^H\boldsymbol{X})^{-1}(\boldsymbol{X}^H\boldsymbol{X})h + (\boldsymbol{X}^H\boldsymbol{X})^{-1}\boldsymbol{X}^H\boldsymbol{w} \\ &= \boldsymbol{h} + \boldsymbol{R}_{xx}^{-1}\boldsymbol{X}^H\boldsymbol{w}\end{aligned} \tag{2-31}$$

Using Eq. (2-31) the estimation error is expressed as in Eq. (2-32).

$$\begin{aligned}\boldsymbol{e}_h &= \widehat{\boldsymbol{h}} - \boldsymbol{h} \\ &= \boldsymbol{R}_{xx}^{-1}\boldsymbol{X}^H\boldsymbol{w}\end{aligned} \tag{2-32}$$

The covariance of the channel estimation vector is an important measure for the performance of the estimation method. The covariance of the channel estimation for LSSE method is given by Eq. (2-33).

$$cov\{\hat{h}\} = E\{(\hat{\boldsymbol{h}} - \boldsymbol{h})(\hat{\boldsymbol{h}} - \boldsymbol{h})^H\}$$

$$= E\{\boldsymbol{e}_h \boldsymbol{e}_h^H\}$$

$$= E\{\boldsymbol{R}_{xx}^{-1} \boldsymbol{X}^H \boldsymbol{w}\boldsymbol{w}^H \boldsymbol{X} \boldsymbol{R}_{xx}^{-1}\}$$

$$= \boldsymbol{R}_{xx}^{-1} \boldsymbol{X}^H E\{\boldsymbol{w}\boldsymbol{w}^H\} \boldsymbol{X} \boldsymbol{R}_{xx}^{-1} \tag{2-33}$$

In Eq. (2-33), the operator $E\{.\}$ denotes the expectation. The statement $E\{\boldsymbol{w}\boldsymbol{w}^H\}$ is a N-by-N covariance matrix of the noise vector. This matrix has a diagonal matrix structure and diagonal elements are equal to variance of the channel noise.

$$E\{\boldsymbol{w}\boldsymbol{w}^H\} = \sigma_w^2 \boldsymbol{I} \tag{2-34}$$

Using Eq. (2-34), the covariance of channel estimation can be rearranged as

$$cov\{\hat{\boldsymbol{h}}\} = \sigma_w^2 \boldsymbol{R}_{xx}^{-1} \boldsymbol{R}_{xx} \boldsymbol{R}_{xx}^{-1}$$

$$= \sigma_w^2 \boldsymbol{R}_{xx}^{-1} \tag{2-35}$$

By Eq. (2-35) it can be said that the mean square error (MSE) of lth channel tap estimate is noise variance times lth diagonal element of inverse correlation matrix, $\boldsymbol{R}_{xx}^{-1}$, Therefore total mean square error of estimation is the sum of the diagonal elements (trace) of inverse correlation matrix multiplied by noise variance [1]. Hence, the minimum mean square error is achieved in LSSE channel estimation when the correlation matrix of the training sequence is perfectly diagonal. The reason is that the maximum for trace of correlation matrix is achieved when the matrix is perfectly diagonal [1]. (See Appendix-A)

$$mse(\hat{\boldsymbol{h}}) = \sigma_w^2 tr(\boldsymbol{R}_{xx}^{-1}) \tag{2-36}$$

2.2.3 Maximum Likelihood (ML) Method [4]

A Maximum Likelihood (ML) estimation approach for channel estimation is proposed by [4]. The first method proposed in [4] treats training sequence symbols

as samples of a Gaussian process. The second method in [4] treats training sequence as a deterministic sequence. In this section we examine both of these approaches.

Maximum Likelihood (ML) channel estimation is based on the maximization of the probability function of the communication signal over the channel impulse response vector, $\boldsymbol{h}$. It requires transmitting the training sequence more than one time, unlike the LSSE method requiring the transmission of training sequence only one time.

A total number of K training sequences separated by unknown data blocks is inserted into the channel in one signal burst. Although sending the different training sequences after each data block is possible, generally same training sequence is repeated. Similar to the signal model given by Eq. (2-2), the received signal samples after sending kth training sequence is expressed by

$$y_k[n] = \sum_{l=0}^{L-1} h[l] z_k[n-l] + w_k[n] \quad n = n_{k,} n_k + 1, \dots, n_k + N + L - 2 \qquad (2\text{-}46)$$

where $z_k[n_k - L + 1], \dots, z_k[n_k - 1]$ and $z_k[n_k - N], \dots, z_k[n_k - N + L - 2]$ are (L-1)-element data blocks, $z_k[n_k], \dots, z_k[n_k + N - 1] = x[0], \dots, x[N-1]$ is N-element training sequence and $w[n]$ is AWGN. Eq. (2-46) can be written in matrix form as sum of a deterministic and a stochastic vector.

$$\boldsymbol{y}_k = \boldsymbol{X}\boldsymbol{h} + \boldsymbol{\eta}_k \qquad (2\text{-}47)$$

In Eq. (2-47), $\boldsymbol{X}$ is Toeplitz matrix containing the elements of the training sequence and $\boldsymbol{h}$ is channel impulse response vector. The stochastic term $\boldsymbol{\eta}_k$ is given by

$$\boldsymbol{\eta}_k = \left[\begin{array}{ccc|ccc} h[L-1] & \dots & h[1] & & & \mathbf{0} \\ & \ddots & \vdots & h[0] & & \\ & & h[L-1] & \vdots & \ddots & \\ \mathbf{0} & & & h[L-2] & \dots & h[L-2] \end{array}\right] \boldsymbol{d}_k + \boldsymbol{w}_k \qquad (2\text{-}48)$$

$$\boldsymbol{\eta}_k = \boldsymbol{H}\boldsymbol{d}_k + \boldsymbol{w}_k \qquad (2\text{-}49)$$

where $\boldsymbol{H}$ is (N+L-1)-by-2L matrix, $\boldsymbol{d}_k$ is 2L-by-1 data symbol vector and $\boldsymbol{w}_k$ is (N+L-1)-by-1 AWGN vector.

$$\boldsymbol{d}_k = [z_k[n_k - L + 1], \dots, z_k[n_k - 1]\ z_k[n_k - N], \dots, z_k[n_k - N + L - 2]]^T \qquad (2\text{-}50)$$

Assuming both data and noise is zero mean and Gaussian, the expectation and covariance of Eq. (2-47) is described as

$$E\{\boldsymbol{y}_k\} = \boldsymbol{X}_k\boldsymbol{h} \tag{2-51}$$

$$cov\{\boldsymbol{y}_k\} = \boldsymbol{C}_d = \lambda^2\boldsymbol{H}\boldsymbol{H}^H + \sigma_w^2\boldsymbol{I}_{(N+L-1)} \tag{2-52}$$

Using Eq. (2-51) and Eq. (2-52), exponential and negative log likelihood functions are given by Eq. (2-53) and Eq. (2-54) respectively.

$$\Lambda(\boldsymbol{y};\boldsymbol{h},\boldsymbol{C}_d) = |\boldsymbol{C}_d|^{-K}.\exp\left\{-\sum_{k=1}^{K}[\boldsymbol{y}_k - \boldsymbol{X}_k\boldsymbol{h}]^H\boldsymbol{C}_d{}^{-1}[\boldsymbol{y}_k - \boldsymbol{X}_k\boldsymbol{h}]\right\} \tag{2-53}$$

$$-\mathcal{L} = \mathrm{K}\ln|\boldsymbol{C}_d| + \sum_{k=1}^{K}[\boldsymbol{y}_k - \boldsymbol{X}_k\boldsymbol{h}]^H\boldsymbol{C}_d{}^{-1}[\boldsymbol{y}_k - \boldsymbol{X}_k\boldsymbol{h}] \tag{2-54}$$

According to [4], because the matrix $\boldsymbol{C}_d$ can be any symmetric positive definite matrix, regardless of $\boldsymbol{h}$ and σ_w^2, Eq. (2-54) presents a new optimization approach separable in its two variables $\boldsymbol{h}$ and $\boldsymbol{C}_d$. By this way, the optimum channel response estimation is obtained with less complexity than natural search process of ML estimation.

In order to solve the channel estimation problem, [4] proposes an iterative search process. The aim of iterations is minimizing the cost function with respect to one of the variable as other variable is fixed. In every iteration, the cost function differs in accordance with variable used for minimization. If cost function has a convex structure then the optimum estimation is the point where the function achieves its global minimum. On the other hand, if cost function is not convex, estimation is global minimum near to the starting point of iteration.

The iteration process begins by keeping the variable $\boldsymbol{C}_d$ as fixed firstly and searching the negative log likelihood function with respect to $\boldsymbol{h}$. The closed form for result of this search is obtained as

$$\widehat{\boldsymbol{h}}^1 = \boldsymbol{h}_{ML}\left(\boldsymbol{C}_d{}^0\right) = \left(\sum_{k=1}^{K} \boldsymbol{X}_k^H\, \boldsymbol{C}_d{}^{-1} \boldsymbol{X}_k\right)^{-1} \sum_{k=1}^{K} \boldsymbol{X}_k^H\, \boldsymbol{C}_d{}^{-1} \boldsymbol{y}_k \tag{2-55}$$

In the second iteration the covariance matrix which minimizes the cost function is searched by keeping $\boldsymbol{h}$ as fixed to the value found in the first iteration.

$$\widehat{\boldsymbol{C}}_d{}^2 = \boldsymbol{C}_{d_{ML}}\left(\widehat{\boldsymbol{h}}^1\right) = \frac{1}{K}\sum_{k=1}^{K} (\boldsymbol{y}_k - \boldsymbol{X}_k h)\,(\boldsymbol{y}_k - \boldsymbol{X}_k h)^H \tag{2-56}$$

The search for optimal values of $\boldsymbol{h}$ and $\boldsymbol{C}_d$ continuous by keeping $\widehat{\boldsymbol{C}}_d{}^2$ as fixed for the third iteration and $\widehat{\boldsymbol{h}}^3$ for the forth iteration. The iterations are carried on until no significant difference between two adjacent iteration is seen. Because of the nonlinear characteristics of ML search process, a closed form formulation for estimation error cannot be obtained.

As discussed at the beginning of this section, the case for deterministic data sequences can also be examined. In the presence of deterministic data sequences and AWGN, the first and the second order statistics of the received signal take the form in Eq. (2-57) and Eq. (2-58).

$$E\{\boldsymbol{y}_k\} = \boldsymbol{X}_k \boldsymbol{h} \tag{2-57}$$

$$cov\{\boldsymbol{y}_k\} = \sigma_w^2 \boldsymbol{I}_{(N+L-1)} \tag{2-58}$$

Inserting Eq. (2-57) and Eq. (2-58) into Eq. (2-54), the negative log likelihood function can be expressed as

$$-\mathcal{L} = \mathrm{Kln}\left|\sigma_w^2 \boldsymbol{I}_{(N+L-1)}\right| + \frac{1}{\sigma_w^2}\sum_{k=1}^{K} [\boldsymbol{y}_k - \boldsymbol{X}_k \boldsymbol{h}]^H [\boldsymbol{y}_k - \boldsymbol{X}_k \boldsymbol{h}] \tag{2-59}$$

Because the variance of noise, σ_w^2, is a scalar value, the minimization of the negative log likelihood function becomes equivalent to the minimization of the function in Eq. (2-60)

$$\ell = \sum_{k=1}^{K} [\boldsymbol{y}_k - \boldsymbol{X}_k \boldsymbol{h}]^H [\boldsymbol{y}_k - \boldsymbol{X}_k \boldsymbol{h}] \tag{2-60}$$

The point where the function ℓ has minimum is found by taking the derivative of Eq. (2-60) over $\boldsymbol{h}^H$, equating the derivative to zero and solving the equation for $\boldsymbol{h}$. Following this process, the optimum channel estimation with ML method is found as in Eq. (2-62)

$$\widehat{\boldsymbol{h}}_{ML} = \left(\sum_{k=1}^{K} \boldsymbol{X}_k^H \boldsymbol{X}_k \right)^{-1} \sum_{k=1}^{K} \boldsymbol{X}_k^H \boldsymbol{y}_k \tag{2-61}$$

Investigating the Eq. (2-61) carefully, it can be said that the ML estimate for deterministic data sequence case is equal to the optimal solution found in LSSE method. This means that in the conditions that only random variable is AWGN, the ML channel estimation approach reduces to the LSSE channel estimation method.

2.3 Placement of Training Sequences into the Channel

Placement of the training sequences into the communication signal is an important issue for the training based estimation schemes, [20]. Places of the sequences among unknown user data and placement method changes accordance with the used estimation scheme.

One of the basic training sequence placement scheme is illustrated in Figure 2-2. In this scheme, a single training sequence is placed into the channel before the transmission of the user data. Generally, prior to the training sequence, a precursor sequence is transmitted. The precursor is used to protect the autocorrelation structure of the training sequence and ignored during the processing of the preamble. In this case, communication channel is assumed as stable during the transmission of the preamble and the user data.

Figure 2-2: Training Sequence Placement Scheme-1

A second placement method for training sequences is expressed in Figure 2-3. In this method, a fixed symbol block is repeated to construct a periodic training sequence. Before the transmission of the periodic sequence, some part of the fixed data block can be used as precursor to prevent defecting the autocorrelation structure of the training sequence. This type of placement scheme is realized with estimation algorithms using benefit from the periodic autocorrelation structure of the training sequences.

Figure 2-3: Training Sequences Placement Scheme-2

A third placement method is illustrated in Figure 2-4. In this method, the training sequence blocks separated by unknown user data blocks are placed into the channel. According to the applied estimation scheme, the same training sequence block is repeated or the different training sequence blocks are transmitted in the preamble. This type of placement scheme is generally used with semiblind estimation methods using the statistical characteristics of the user data.

Figure 2-4: Training Sequences Placement Scheme-3

A fourth placement method applicable for frequency offset and channel estimation method proposed in this thesis study is shown in Figure 2-5. This placement method is applied in the case of using complementary sequences pairs, which are described in Section 2.4, as training purposes. In this method the sequences in the complementary pair are placed into the channel successively. In order to protect the complementary autocorrelation characteristics of the complementary sequences pair, communication channel left as idle along enough time duration between the transmission of the complementary sequences and the transmission of the second complementary sequence and the user data.

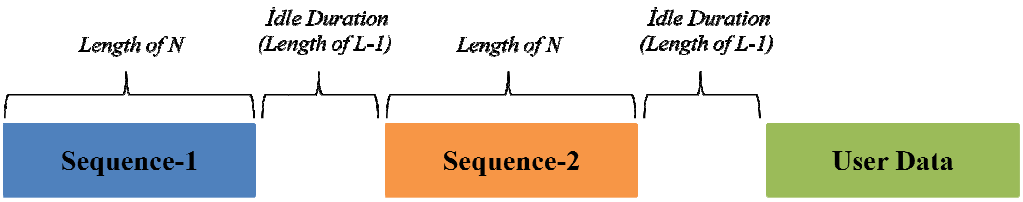


Figure 2-5: Training Sequences Placement Scheme-4

2.4 Complementary Sequences

Complementary sequences firstly were introduced by Marcel J. E. Golay in 1949, [11]. The sequences introduced by Golay were binary based. In the following years some other authors generalized the complementary sequences concept to polyphase complementary sequences, multilevel complementary sequences and arbitrary complex complementary sequences, [18].

Golay complementary sequences have found application in physics, spread spectrum communication, OFDM, radar and biomedical signal processing. In scope of this thesis study, complementary pairs have been used for frequency offset and channel estimation.

2.4.1 Definition

A complementary sequences pair are defined as pairs of sequences with the property that their aperiodic autocorrelation coefficients sum to zero [12].

Let $x_i = \{x_i[0], x_i[1], \dots, x_i[N-1]\}$ be a sequence of elements such that $x[n] \epsilon \{+1, -1\}$. The aperiodic autocorrelation function of x is defined by Eq. (2-62).

$$R_{x_i x_i}[k] = \sum_{n=0}^{N-1} x_i[n] x_i[n-k] \quad k = 0, 1, \dots, N-1 \tag{2-62}$$

The pair of sequences (x_i, x_j) is called as complementary sequences pair if the condition in Eq. (2-63) is satisfied.

$$\begin{cases} R_{x_i x_i}[k] + R_{x_j x_j}[k] = 2N & if\ k = 0 \\ R_{x_i x_i}[k] + R_{x_j x_j}[k] = 0 & if\ k \neq 0 \end{cases} \tag{2-63}$$

An example of complementary sequences pair introduced by Golay is:

$$x_1 = \{-1 + 1 - 1 - 1 + 1 - 1 - 1 - 1\}$$

$$x_2 = \{-1 - 1 - 1 + 1 + 1 + 1 - 1 + 1\}$$

The autocorrelation sequences for $k = 0, 1, \dots, 7$ are:

$$R_{x_1x_1} = \{8, -1, 0, 3, 0, 1, 0, 1\}$$

$$R_{x_2x_2} = \{8, 1, 0, -3, 0, -1, 0, -1\}$$

The individual autocorrelation sequences of x_1 and x_2 and sum of the autocorrelation sequences are illustrated in Figure 2-2. As it is seen from the figure the sum of the non-zero lags of autocorrelation sequences is zero while the the zeroth lag sums to $2N$.

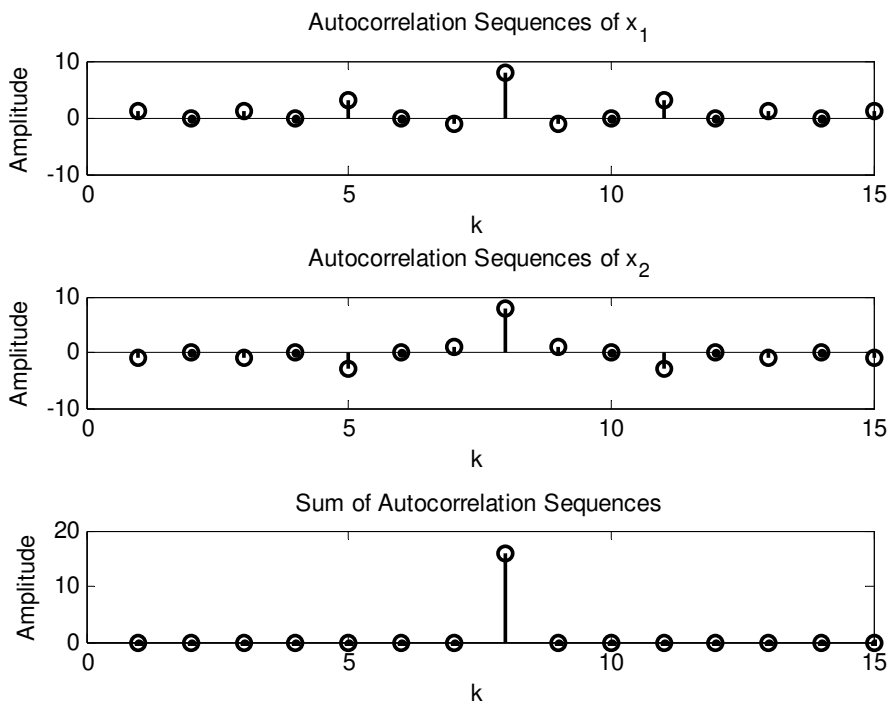


Figure 2-6: Individual autocorrelation sequences and sum of autocorrelation sequences of complementary sequences examples

2.4.2 Existence

Golay introduced complementary sequences of length 2^α. He also gave some examples of complementary sequences for lengths 10 and 26 and a method to construct the length of $N = 2^\alpha 10^\beta 26^\gamma$ complementary pairs. In [12] it is showed

that the length of a complementary sequence must be the sum of the two squares (one square may be 0) and even. Therefore, the admissible lengths up to 50 are

$$1, 2, 4, 8, 10, 16, 20, 26, 32, 34^*, 40, 50$$

In recent years some computer searches have been performed to test the existence of complementary sequences which have lengths given above and the length marked by '*' have been eliminated.

2.4.3 General Properties

Below are the general properties of the complementary sequences.

1. A complementary sequences pair can be expressed as polynomial like in Eq. (2-64).

$$x_i(z) = x_i[0] + x_i[1]z^{-1} + \cdots + x_i[N-2]z^{-(N-2)} + x_i[N-1]z^{-(N-1)} \quad (2\text{-}64)$$

   Using the mathematical property in Eq. (2-65) we can say that a complementary sequences pair satisfies the condition in Eq. (2-66)

$$x_i(z)x_i(z^{-1}) = R_{x_i x_i}[0] + \sum_{k=1}^{N-1} R_{x_i x_i}[k](z^k + z^{-k}) \quad (2\text{-}65)$$

$$x_1(z)x_1(z^{-1}) + x_2(z)x_2(z^{-1}) = 2N \quad (2\text{-}66)$$

   Restricting z in unit circle in complex plain and inserting e^{jw} instead of z into Eq. (2-66), we get expression in Eq. (2-67).

$$\left|x_1\left(e^{jw}\right)\right|^2 + \left|x_2\left(e^{jw}\right)\right|^2 = 2N \quad (2\text{-}67)$$

   The Eq. (2-67) says that a complementary sequences pair has complementary spectra. This means that the sum of the squared magnitude of the Fourier transform of the sequences in a complementary pair is constant [13].

2. If the signs of a complementary pair sequence are reversed (multiplied by -1) , the pair remains complementary [12]. For instance, if we inverted the second

sequences of complementary sequences pair CSP-1, we get another complementary sequences pair, CSP-2.

CSP-1	-1 +1 -1 -1 +1 -1 -1 -1	and	-1 -1 -1 +1 +1 +1 -1 +1
CSP-2	-1 +1 -1 -1 +1 -1 -1 -1	and	+1 +1 +1 -1 -1 -1 +1 -1

3. If any of the sequences in a complementary pair is reverted (inverted in time), they remain complementary [12]. For example, the second sequences of CSP-1 is reverted, another complementary sequences pair, CSP-3, is achieved.

CSP-1	-1 +1 -1 -1 +1 -1 -1 -1	and	-1 -1 -1 +1 +1 +1 -1 +1
CSP-3	-1 +1 -1 -1 +1 -1 -1 -1	and	+1 -1 +1 +1 +1 -1 -1 -1

4. If any of the sequences in a complementary pair is delayed, they remain complementary [12].
5. If sequences in a complementary pair are interchanged as illustrated below they remain complementary [12]. For example, the sequences of CSP-1 are interchanged, we get another complementary sequences pair, CSP-4.

CSP-1	-1 +1 -1 -1 +1 -1 -1 -1	and	-1 -1 -1 +1 +1 +1 -1 +1
CSP-4	-1 -1 -1 +1 +1 +1 -1 +1	and	-1 +1 -1 -1 +1 -1 -1 -1

6. If both sequences in a complementary pair are multiplied by a complex constant, they remain complementary [12]. For example, if sequences of CSP-1 are multiplied by $1 - j$, the achieved sequences pair, CSP-5, is complementary.

CSP-1	-1 +1 -1 -1 +1 -1 -1 -1	and	-1 -1 -1 +1 +1 +1 -1 +1
CSP-5	-1+j, +1-j, -1+j, -1+j, +1-j, -1+j, -1+j, -1+j	and	-1+j, -1+j, -1+j, +1-j, +1-j, +1-j, -1+j, +1-j

7. If both of the sequences in a complementary pair are decimated in time they remain complementary [12].
8. Let (x_1, x_2) be a complementary sequences pair. The new sequences pair formed by operations $x_3 = \{x_1, x_2\}$ and $x_4 = \{x_1, -x_2\}$ is a complementary pair [12],

[13]. For instance, let CSP-6 be a length of 2 complementary sequences pair. Then the new pair CSP-7 formed from CSP-6 by operation defined in previous sentence is a complementary sequences pair.

CSP-6	+1 +1	and	+1 -1
CSP-7	+1 +1 +1 -1	and	+1 +1 -1 +1

9. Let (x_1, x_2) and (x_3, x_4) be complementary sequences pairs of lengths M and N respectively. The new sequences pair, (x_5, x_6), formed by concatenation of (x_1, x_2) and (x_3, x_4) is also complementary [12], [13]. The concatenation operation of complementary pairs is expressed in z-domain as in Eq. (2-68) and Eq. (2-69).

$$x_5(z) = x_1(z^N)x_3(z) + x_2(z^N)x_4(z)z^{MN} \tag{2-68}$$

$$x_6(z) = \tilde{x}_2(z^N)x_3(z) - \tilde{x}_1(z^N)x_4(z)z^{MN} \tag{2-69}$$

As an example, the new pair, CSP-8, formed by concatenation of CSP-6 and CSP-7 is a complementary sequences pair.

CSP-6	+1 +1	and	+1 -1
CSP-7	+1 +1 +1 -1	and	+1 +1 -1 +1
CSP-8	+1 +1 +1 -1 +1 +1 +1 -1 +1 +1 -1 +1 -1 -1 +1 -1	and	-1 -1 -1 +1 +1 +1 +1 -1 -1 -1 +1 -1 -1 -1 +1 -1

10. Let (x_1, x_2) and (x_3, x_4) be complementary sequences pairs of lengths M and N respectively. The new sequences pair, (x_7, x_8), formed by interleaving (x_1, x_2) and (x_3, x_4) is also complementary [12], [13]. The interleaving operation of complementary pairs is expressed in z-domain as in Eq. (2-70) and Eq. (2-71).

$$x_7(z) = x_1(z^{2N})x_3(z^2) + x_2(z^{2N})x_4(z^2)z \tag{2-70}$$

$$x_8(z) = \tilde{x}_2(z^{2N})x_3(z^2) - \tilde{x}_1(z^{2N})x_4(z^2)z \tag{2-71}$$

As an example, the new pair, CSP-9, formed by interleaving CSP-6 and CSP-7 is a complementary sequences pair.

CSP-6	+1 +1	and	+1 -1
CSP-7	+1 +1 +1 -1	and	+1 +1 -1 +1
CSP-8	+1 +1 +1 +1 +1 -1 -1 +1 +1 -1 +1 -1 +1 +1 -1 -1	and	-1 -1 -1 -1 -1 +1 +1 -1 +1 -1 +1 -1 +1 +1 -1 -1

CHAPTER 3

JOINT TRAINING BASED FREQUENCY OFFSET AND CHANNEL ESTIMATION

In this chapter, training based joint frequency offset and channel estimation schemes are analyzed. In Section 3.1, signal models for a SISO communication system over frequency selective channels is introduced for single and multiple frequency offset cases. Section 3.2 describes two ML based estimation method using aperiodic and periodic training sequences. In Section 3.3, an ad hoc estimation method is presented. In Section 3.4 the new ML based estimation method using complementary training sequences is proposed. Finally, in Section 3.5 some extensions to the ML based methods are introduced.

3.1 Signal Model

In dealing with channel estimation problem of Single Input Single Output (SISO) communication systems, most of the channel estimation algorithms assume zero frequency offset between the carrier of transmitted signal and reference signal produced by local oscillator of the receiver. In practical applications, zero frequency offset means that unwanted frequency drift or variations are negligible and the demodulation process is not significantly affected. In practice stable oscillators are used in receiver structures in order to prevent the significant phase variations of the

carrier during the demodulation. Although this provide an improvement in demodulation, it is not so easy to prevent frequency offset.

In wireless communication, there are two different reasons causing the frequency offset.

1. Poor frequency synchronization between the carrier of the transmitted signal and the carrier signal produced by local oscillator of the receiver.
2. Doppler-shift of the transmitted signal

Synchronization problem occurs due to the usage of different reference oscillators in transmitting and receiving side of the communication. This is the main reason of the frequency offset problem. If there is no other reason causing the phase variation, the frequency offset caused by poor synchronization is same for different path of a frequency-selective channel (single frequency offset).

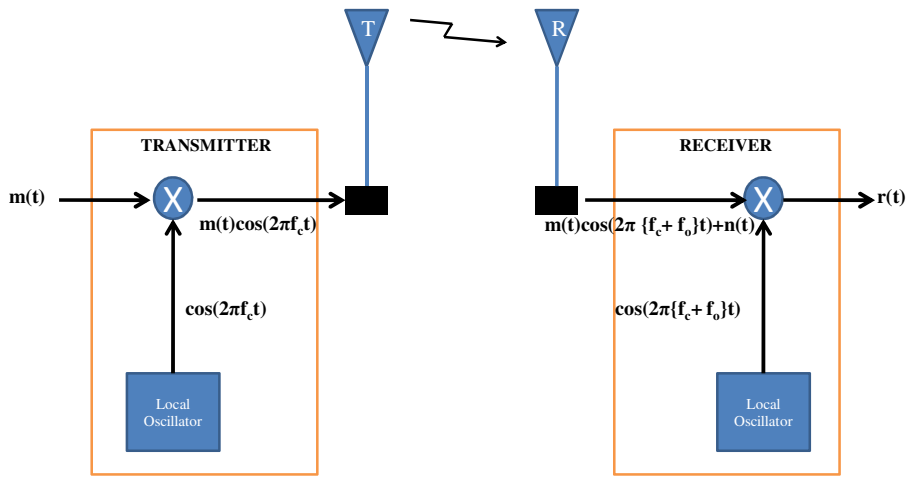


Figure 3-1: Frequency Offset Due to Local Oscillators

By assuming a linear modulation (PSK, QAM, etc.) scheme over a frequency selective channel, the received signal samples at the receiving side are expressed by

$$y[n] = e^{j2\pi nv}s[n] + w[n] \qquad n = 0, 1, \dots, N-1 \tag{3-1}$$

In Eq. (3-1), v is the frequency offset normalized to $1/T_s$ where T_s is the signaling interval and the term $s[n]$ is given by

$$s[n] = \sum_{l=0}^{L-1} h[l]x[n-l] \qquad n = 0, 1, 2, \dots, N-1 \tag{3-2}$$

Here, $h[0], h[1], \dots, h[L-1]$ are the channel gains and $x[-L+1], \dots, x[0], x[1], \dots, x[N-1]$ are the training symbols. The symbols $x[-L+1], \dots, x[-1]$ are precursors in the preamble.

Eq. (3-1) represents the case either the signal is sampled after IF filtering or sampled after matched filtering operation. If signal is considered as matched filtered, then v $(f_o T_s)$ should be small for the validity of Eq. (3-1).

Considering the received signal samples from $n = 0$ to $n = N-1$, Eq. (3-1) can be expressed in matrix form as

$$\boldsymbol{y} = \boldsymbol{FXh} + \boldsymbol{w} \tag{3-3}$$

where $\boldsymbol{F}$ is a N-by-N diagonal matrix containing frequency offsets on received signal samples

$$\boldsymbol{F} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & e^{j2\pi v} & 0 & \cdots & 0 \\ 0 & 0 & e^{j4\pi v} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & e^{j2\pi(N-1)v} \end{bmatrix} \tag{3-4}$$

$\boldsymbol{X}$ is a N-by-L matrix containing entries $x[n-l]$ $(n = 0, 1, \dots, N-1,\ l = 0, 1, \dots, L-1)$ and $\boldsymbol{h}$ is the vector containing the channel gains $h[0], h[1], \dots, h[L-1]$.

$$\boldsymbol{y} = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{j2\pi(N-1)v} \end{bmatrix} \begin{bmatrix} x[0] & \cdots & x[-L+1] \\ \vdots & \ddots & \vdots \\ x[N-1] & \cdots & x[N-L] \end{bmatrix} \begin{bmatrix} h[0] \\ \vdots \\ h[L-1] \end{bmatrix}$$

$$+ \begin{bmatrix} w[0] \\ \vdots \\ w[N-1] \end{bmatrix} \tag{3-5}$$

Doppler-shift is generally seen in mobile communication systems. The frequency offset caused by Doppler-shift occurs due to the relative motion between mobile station and stationary party. In presence of Doppler-shift the frequency offset is the sum of the frequency offsets due to Doppler-shift and poor synchronization. In this case the frequency offsets for different paths of a frequency-selective channel are different (multiple frequency offset). Under such a condition, the received signal samples are expressed as

$$y[n] = \sum_{l=0}^{L-1} h[l]x[n-l]\, e^{j2\pi n v_l} + w[n] \tag{3-6}$$

where v_l is normalized (by $1/T_s$) frequency offset for lth path of the communication channel.

It is not possible to put Eq. (3-6) into a closed matrix form as easily as Eq. (3-2). However, profiting from the signal definition in [10] and defining a matrix $\boldsymbol{A} = [\boldsymbol{a}_0\, \boldsymbol{a}_1 \ldots \boldsymbol{a}_{L-1}]$ we can define a closed form for Eq. (3-6) as

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{h} + \boldsymbol{w} \tag{3-7}$$

where $\boldsymbol{a}_l$ is elementwise product of vectors containing training symbols and samples of frequency offset for lth channel path.

$$\boldsymbol{a}_l = \boldsymbol{x}_l \odot \boldsymbol{f}_{ol} \tag{3-8}$$

$$\boldsymbol{x}_l = [x[n-l] \ldots x[n-l-N+1]]^T \quad l = 0, 1, \ldots, L-1 \tag{3-9}$$

$$\boldsymbol{f}_{ol} = \left[e^{j2\pi n v_l} \ldots e^{j2\pi(n-N+1)v_l}\right]^T \quad l = 0, 1, \ldots, L-1 \tag{3-10}$$

In both Eq. (3-3) and Eq. (3-7), vector $\boldsymbol{w}$ is zero-mean Gaussian noise with covariance matrix

$$\boldsymbol{C}_w = E\{\boldsymbol{w}\boldsymbol{w}^H\} = \sigma_w^2 \boldsymbol{I}_N \tag{3-11}$$

where $\boldsymbol{I}_N$ is N-by-N identity matrix and σ_w^2 is variance of the noise.

For both of Eq. (3-3) and Eq. (3-7) the receive SNR at the output of the receiver is defined as

$$SNR = \sigma_s^2 / \sigma_w^2 \tag{3-12}$$

where σ_s^2 is energy of the sequence $s[n]$ (Eq. (3-2)) and defined with Eq. (3-13)

$$\sigma_s^2 = \frac{1}{N} \sum_{n=0}^{N-1} |s[n]|^2 \tag{3-13}$$

3.2 Maximum Likelihood Frequency Offset and Channel Estimation in Presence of Single Frequency Offset [5]

A basic maximum likelihood frequency offset and channel estimation method has been introduced by [5]. In this section this method is explained.

For a given channel impulse response vector and normalized frequency offset pair $(\boldsymbol{h}, v)$, the received signal samples vector $\boldsymbol{y}$ is Guassian with mean $\boldsymbol{FXh}$ and covariance matrix $\sigma_w^2 \boldsymbol{I}_N$. Under this condition, the likelihood function for pair $(\boldsymbol{h}, v)$ is given as in Eq. (3-14).

$$\Lambda(\boldsymbol{y}; \boldsymbol{h}, v) = \frac{1}{(\pi \sigma_w^2)^N} . \exp\left\{-\frac{1}{\sigma_w^2} [\boldsymbol{y} - \boldsymbol{FXh}]^H [\boldsymbol{y} - \boldsymbol{FXh}]\right\} \tag{3-14}$$

The aim of maximum likelihood approach is to maximize the Eq. (3-14) over the parameters $\boldsymbol{h}$ and v. The location of the maximum gives the joint ML estimate of $\boldsymbol{h}$ and v. [5]

In order to proceed, firstly the frequency offset v accept as fixed and the maximum of likelihood function with respect to $\boldsymbol{h}$ is achieved by least square solution of cost function ℓ in Eq. (3-15).

$$\ell = [\boldsymbol{y} - \boldsymbol{FXh}]^H[\boldsymbol{y} - \boldsymbol{FXh}] \tag{3-15}$$

$$= \boldsymbol{y}^H\boldsymbol{y} - 2\boldsymbol{h}^H\boldsymbol{X}^H\boldsymbol{F}^H\boldsymbol{y} + \boldsymbol{h}^H\boldsymbol{X}^H\boldsymbol{Xh}$$

Taking the derivative of the cost function ℓ with respect to $\boldsymbol{h}^H$ and equating the derivation to zero, the least square solution for $\boldsymbol{h}$ takes the form

$$\boldsymbol{h}_{ls} = (\boldsymbol{X}^H\boldsymbol{X})^{-1}\boldsymbol{X}^H\boldsymbol{F}^H\boldsymbol{y} \tag{3-16}$$

Substituting Eq. (3-16) into Eq. (3-15), the cost function ℓ becomes

$$\ell = \boldsymbol{y}^H\boldsymbol{y} - \boldsymbol{y}^H\boldsymbol{FX}(\boldsymbol{X}^H\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{X}^H\boldsymbol{F}^H\boldsymbol{y} \tag{3-17}$$

Optimum value for frequency offset v is achieved by minimizing the function ℓ. This means maximizing the function $g(v)$ in Eq. (3-18).

$$g(v) = \boldsymbol{y}^H\boldsymbol{FX}(\boldsymbol{X}^H\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{X}^H\boldsymbol{F}^H\boldsymbol{y} \tag{3-18}$$

Maximum of function is found by a search over frequency offset v. Changing the value of v in a suitable range and with a suitable step size the optimum estimated value $\hat{v}$ for frequency offset v is found and expressed like in Eq. (3-19)

$$\hat{v} = \arg\max_{\hat{v}}\{g(\hat{v})\} \tag{3-19}$$

3.2.1 ML Estimation with Aperiodic Training Sequences (MLE#1) [5]

In use of aperiodic training sequences the function $g(\hat{v})$ in Eq. (3-18) may be put in the form

$$g(\hat{v}) = -\rho_a(0) + 2Re\left\{\sum_{m=0}^{N-1} \rho_a(m)\, e^{-j2\pi n\hat{v}}\right\} \tag{3-20}$$

where $\rho(m)$ is a weighted correlation of the training sequence and defined by

$$\rho_a(m) = \sum_{n=m}^{N-1} \boldsymbol{B}(n-m,n)y[n]y^*[n-m] \tag{3-21}$$

In Eq. (3-21) $\boldsymbol{B}$ is the projection matrix defined by Eq. (3-22) and $\boldsymbol{B}(i,j)$ is (i,j) entry of the matrix.

$$\boldsymbol{B} = \boldsymbol{X}(\boldsymbol{X}^H\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{X}^H \tag{3-22}$$

Maximization of $g(\hat{v})$ in Eq. (3-20) requires a two step search. In the first step, maximum of $g(\hat{v})$ is searched over the grid values of $\hat{v}$ and the location of the maximum is found. In the second step, the function $g(\hat{v})$ is searched with smaller steps of $\hat{v}$ and the more accurate location of the maximum is found. Finally, the location found in the second step is accepted as the optimal estimation value for frequency offset v. After estimating the frequency offset, the optimum value is inserted into Eq. (3-16) and the optimum estimation, $\widehat{\boldsymbol{h}}$, for channel impulse response $\boldsymbol{h}$ is found. The closed form for channel estimation is expressed in Eq. (3-23) where $\widehat{\boldsymbol{F}}$ is matrix containing frequency offset estimation.

$$\widehat{\boldsymbol{h}} = (\boldsymbol{X}^H\boldsymbol{X})^{-1}\boldsymbol{X}^H\widehat{\boldsymbol{F}}^H\boldsymbol{y} \tag{3-23}$$

In Eq. (3-22), if matrix $\boldsymbol{X}$ is square ($N = L$) and nonsingular, the projection matrix $\boldsymbol{B}$ becomes a N-by-N identity matrix and Eq. (3-20) becomes independent of frequency offset. This means that the length of aperiodic training sequence N should be larger than the number of the channel taps L.

Examining the Eq. (3-20), it is seen that $g(\hat{v})$ is periodic with unit period. Hence, its maximum occurs at a distance 1 from each other. This says that true estimation value for v is found only in range $|v| \leq 1/2$.

Assuming a high SNR and an observation window containing many symbol intervals the expected value and variation of ML estimation with aperiodic training sequence are approximated by Eq. (3-24) and (3-25) , [5]

$$E\{\hat{v}\} = v \tag{3-24}$$

$$var\{\hat{v}\} = \frac{1}{2N}\frac{\boldsymbol{s}^H\boldsymbol{s}}{\boldsymbol{z}^H(\boldsymbol{I}_N - \boldsymbol{B})\boldsymbol{z}}(SNR)^{-1} \tag{3-25}$$

where $\boldsymbol{s} = \boldsymbol{Xh}$ and $\boldsymbol{z}$ is a vector with components $z[n] = 2\pi n.s[n]$.

Eq. (3-24) indicates that MLE#1 is an unbiased estimator. Because of the asymptotic efficiency property of the ML estimation approach, the Eq. (3-25) constitutes a CRB for frequency offset estimation [5], [17]. (See Appendix C)

From Eq. (3-25), it observed that the estimation accuracy for frequency offset depends on SNR and channel characteristics. Training sequence is also an important factor which affects the variance of the estimate. In order to achieve required estimation performance training sequences which are optimal for joint frequency offset and channel estimation need to be investigated.

In [8], two criterion have been defined for design of an optimal training sequence providing a good frequency offset and channel estimation performance:

Criterion 1: To achieve an efficient channel estimation performance training sequence should provide a diagonal autocorrelation matrix satisfying Eq. (3-26).

$$\boldsymbol{X}^H\boldsymbol{X} = P_x\boldsymbol{I}_N \tag{3-26}$$

where P_x is power of training sequence, which is N in BPSK signaling, and $\boldsymbol{I}_N$ is N-by-N identity matrix.

Criterion 2: For optimal frequency offset estimation training sequence should satisfy Eq. (3-27)

$$\boldsymbol{X} = \arg\max_{\boldsymbol{X}} \; tr\{\boldsymbol{X}^H\boldsymbol{D}^H(\boldsymbol{I}_N - \boldsymbol{B})\boldsymbol{DXR_h}\} \tag{3-27}$$

In Eq. (3-27), $\boldsymbol{D}$ is diagonal matrix defined by Eq. (3-27), $\boldsymbol{B}$ is projection matrix defined by Eq. (3-22) and $\boldsymbol{R_h}$ is autocorrelation matrix of channel impulse response given by Eq. (3-29).

$$\boldsymbol{D} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & N-1 \end{bmatrix} \tag{3-28}$$

$$\boldsymbol{R_h} = \boldsymbol{h}\boldsymbol{h}^H \tag{3-29}$$

With a careful inspection it is seen that Eq. (3-27) is equal to the statement, $\boldsymbol{z}^H(\boldsymbol{I}_N\text{-}\boldsymbol{B})\boldsymbol{z}/4\pi^2$,which is obtained from Eq. (3-25). This equality indicates that the optimal training sequences according to this criteria minimize frequency estimation error.

3.2.2 ML Estimation with Periodic Training Sequences (MLE#2) [5]

Application of periodic training sequences can reduce the complexity of search operation required for ML estimation. A periodic training sequence consists of applicatino of the same training sequence repetitively. For convenience, the length of the training block is chosen equal to the length of the channel impulse response, L. Hence, repeating the symbol block P times a training sequence has a length of multiple of L is constructed ($N = LP$).

Using the periodic training sequence, the training sequence matrix $\boldsymbol{X}$ becomes a repetition of an L-by-L matrix $\boldsymbol{C}$.

$$\boldsymbol{X} = [\boldsymbol{C} \quad \boldsymbol{C} \quad \dots \quad \boldsymbol{C}]^T \tag{3-30}$$

In Eq. (3-30), $\boldsymbol{C}$ is L-by-L circulant matrix with elements

$$\boldsymbol{C}(i,j) = x[(i-j)modL] \qquad i = 0,1,\dots,L-1 \;\; j = 0,1,\dots,L-1 \tag{3-31}$$

Substituting Eq. (3-30) into Eq. (3-22), the projection matrix $\boldsymbol{B}$ takes the form

$$\boldsymbol{B} = \frac{1}{P}\begin{bmatrix} \boldsymbol{I}_L & \boldsymbol{I}_L & \cdots & \boldsymbol{I}_L \\ \boldsymbol{I}_L & \boldsymbol{I}_L & \cdots & \boldsymbol{I}_L \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{I}_L & \boldsymbol{I}_L & \cdots & \boldsymbol{I}_L \end{bmatrix} \tag{3-32}$$

where $\boldsymbol{I}_L$ is L-by-L identity matrix. With new definition the elements of $\boldsymbol{B}$ are expressed as

$$\boldsymbol{B}(n-m,n) = \begin{cases} 1/P & m = iL, \quad i = 0,1,\dots,P-1 \\ 0 & otherwise \end{cases} \tag{3-33}$$

In consequence with the new definition of projection matrix $\boldsymbol{B}$, Eq. (3-20) can be rewritten in a new form

$$g(\hat{v}) = -\rho_p(0) + 2Re\left\{\sum_{m=0}^{P-1} \rho_p(m)\, e^{-j2\pi nL\hat{v}}\right\} \tag{3-34}$$

with

$$\rho_p(m) = \frac{1}{P}\sum_{n=mL}^{N-1} y[n]y^*[n-m] \qquad m = 0,1,\dots,P-1 \tag{3-35}$$

As in aperiodic training sequence case the location of maximum for function $g(\hat{v})$ is found by a two step search. However this time only number of P correlations are processed instead of number of N weighted correlations. This reduces the computational complexity of search process by a factor L.

Examining Eq. (3-34) it is observed that $g(\hat{v})$ is periodic with period $1/L$. This means that the ML estimation gives correct estimation results in the range of $|v| \leq 1/2L$. Hence reducing the computational complexity also reduces the range of unambiguous frequency estimation.

The estimation variance for MLE#2 is achieved by inserting $\boldsymbol{B}$ matrix form in Eq. (3-32) into Eq. (3-25).

$$var\{\hat{v}\} = \frac{3(SNR)^{-1}}{2\pi^2 N(N^2 - L)} \tag{3-36}$$

According to the results of [5], the RHS of Eq. (3-36) still coincides with CRB for periodic sequences. From Eq. (3.36) it is seen that the accuracy of estimation with periodic sequences depends on the length of the channel response and the training sequence as far as SNR.

3.3 An Ad Hoc Frequency Offset Estimator (AHE) in Presence of Single Frequency Offset [5]

An ad hoc frequency offset estimation method (AHE) has been introduced by [5]. This new estimator does not need any grid or fine search like MLE#1 or MLE#2 given in previous sections.

For AHE the same assumptions as with MLE#2 are made and a fixed symbol block of length L is used. The symbol block is repeated P times so that training sequence of length $N = LP$ is achieved.

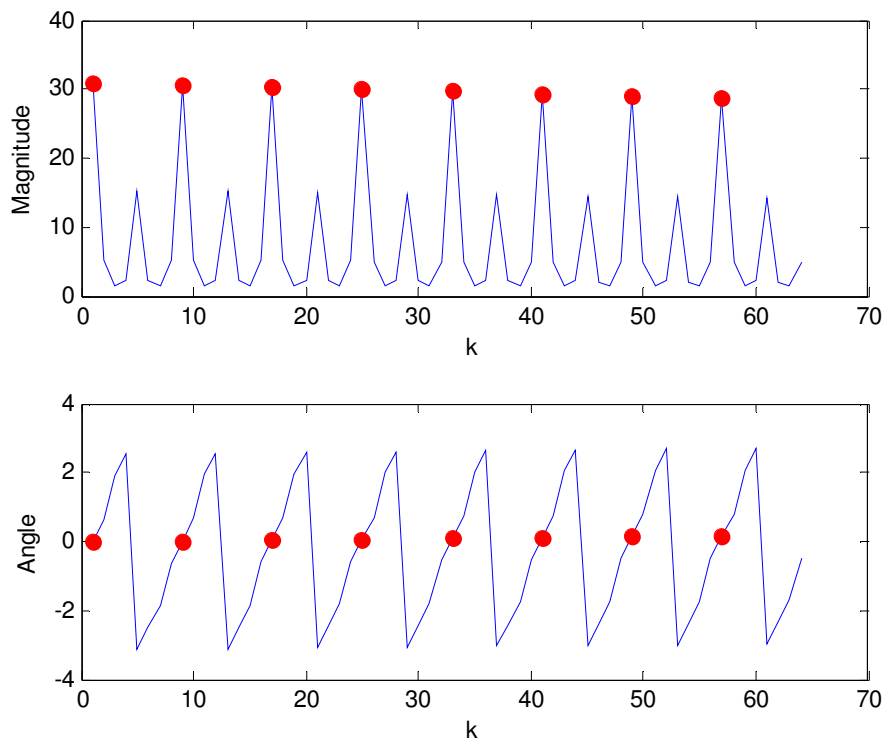


Figure 3-2: Periodic autocorrelation of the received signal samples (magnitude and angle)

The AHE exploits the angle of the periodic points of the autocorrelation of the received signal samples shown in Figure 3-2 in order to find the optimum estimation of the frequency offset. The periodic autocorrelation points shown in the figure can be expressed as in Eq. (3-37).

$$R_{AHE}[m] = \frac{1}{N - mL} \sum_{n=mL}^{N-1} y\,[n] y^*[n - mL] \quad m = 0, 1, \dots, M \tag{3-37}$$

In Eq. (3-37), M is a design parameter which is chosen as not greater than $P - 1$. In order to see the working principle of the AHE, we can rewrite the signal model given by Eq. (3-1) in the form of Eq. (3-38)

$$y[n] = e^{j2\pi n v}[s[n] + \widetilde{w}[n]] \tag{3-38}$$

where $\widetilde{w}[n] = e^{-j2\pi n v} w[n]$. Inserting Eq. (3-38) into Eq. (3-37) the correlation R_{AHE} takes the form in Eq. (3-39)

$$R_{AHE}[m] = \sigma_s^2 e^{j2\pi n L v}[1 + \gamma[m]] \tag{3-39}$$

where

$$\gamma[m] = \frac{1}{\sigma_s^2 (N - mL)} \sum_{n=mL}^{N-1} [s[n]\widetilde{w}^*[k - mL] + s^*[n]\widetilde{w}[n] + \widetilde{w}[n] w^*[n - m]] \tag{3-40}$$

and

$$\sigma_s^2 = \frac{1}{L} \sum_{n=0}^{L-1} |s[n]|^2 \tag{3-41}$$

Profiting from the angle defined by Eq. (3-42) and expressing the $\gamma[m]$ by its imaginary part $\gamma_I[m]$ the approximation in Eq. (3-43) can be written. Analyzing the Eq. (3-42) it is obviously comprehended that the angle $\varphi[m]$ is the angle difference between the mth and $(m - 1)th$ periodic points of the autocorrelation.

$$\varphi[m] = arg\{R_{AHE}[m]R_{AHE}^{*}[m-1]\} \tag{3-42}$$

$$\varphi[m] \cong 2\pi vL + \gamma_I[m] - \gamma_I[m-1] \quad m = 0, 1, \dots, M \tag{3-43}$$

In Eq. (3-43) two important assumptions can be made.

- L satisfies the condition $|2\pi vL| < \pi$
- The terms $\gamma_I[m]$ and $\gamma_I[m-1]$ can be eliminated in high SNR due to their smallness in amplitude compared to unity

Using assumptions above it can be said that the angle $\varphi[m]$ is the sum of a deterministic component proportional to normalized frequency offset v and a zero-mean random component. Hence, the set of angles $\{\varphi[m]\}$ can be used to design an unbiased estimator of normalized frequency offset given by Eq. (3-44).

$$\hat{v} = \frac{1}{2\pi L}\sum_{m=1}^{M} \alpha[m]\varphi[m] \tag{3-44}$$

In Eq. (3-44), $\alpha[m]$ are the components of vector $\boldsymbol{\alpha}$ defined by Eq. (3-45)

$$\boldsymbol{\alpha} = \frac{\boldsymbol{C}_{\varphi}^{-1}\mathbf{1}}{\mathbf{1}^T\boldsymbol{C}_{\varphi}^{-1}\mathbf{1}} \tag{3-45}$$

where $\boldsymbol{C}_{\varphi}$ is the covariance matrix of angle φ and $\mathbf{1}$ is a vector of M ones, [17, p. 138]

According to the results of [5], $\boldsymbol{C}_{\varphi}$ is singular for $M > P/2$ and the angles $\varphi[m]$ for $m > P/2$ are linearly dependent on angles with $1 < m < P/2$. This means that the frequency offset information is carried by angle first $P/2$ components of φ. For that reason it is assumed that $1 < M < P/2$. Exploiting this assumption the components α can be rewritten in the form of Eq. (3-46).

$$\alpha[m] = 3\frac{(P-m)(P-m+1) - M(P-M)}{M(4M^2 - 6PM + 3P^2 - 1)} \quad 1 < m < M \tag{3-46}$$

The variance of ad hoc estimator can be expressed as in Eq. (3-47).

$$var\{\hat{v}\} = \frac{1}{(2\pi L)^2}\frac{1}{\mathbf{1}^T \boldsymbol{C}_{\varphi}^{-1}\mathbf{1}} \tag{3-47}$$

Using the assumption $1 < M < P/2$ the estimation variance can be rewritten in the form of Eq.(3-48).

$$var\{\hat{v}\} = \frac{3(SNR)^{-1}}{4\pi^2 ML^3(4M^2 - 6PM + 3P^2 - 1)} \tag{3-48}$$

From the Eq. (3-48), it can be said that the estimation variance achieves its minimum when $M = P/2$. Inserting $P/2$ instead of M in Eq. (3-48), it is obviously seen that minimum value expression of estimation variance coincides with Eq. (3-36) which is the estimation variance and CRB of MLE#2. Therefore, Eq. (3-48) with $M = P/2$ can be assumed as CRB of ad hoc estimator.

The weights $\alpha[m]$ are independent of the channel impulse response. Hence for frequency offset estimation AHE does not need any channel information. Because the AHE provides a closed form expression for frequency offset, no search operation in estimation is needed. For that reason the computational load is reduced compared with MLE#1 and MLE#2 methods.

As stated before the AHE gives sufficient results as far as the Eq. (3-43) holds. Otherwise, the relation between the normalized frequency offset v and angle $\varphi[m]$ becomes highly nonlinear. Assuming that the Eq. (3-43) holds, the estimation range of AHE is limited to $|v| < 1/2L$.

3.4 Maximum Likelihood Frequency Offset and Channel Estimation Using Complementary Training Sequences in Presence of Single Frequency Offset

In this section, the new frequency offset and channel estimation method using complementary sequences pairs is introduced . This new method is based on the maximum likelihood (ML) frequency offset and channel estimation method described in Section 3.2. The aim of this study is to use perfect autocorrelation

characteristics of complementary sequences in estimation of frequency offset as well as channel estimation.

In order to estimate the frequency offset and the impulse response, a special preamble structure consists of a binary complementary pair is put into the communication channel. The structure of this special preamble is shown in Figure 3-3.

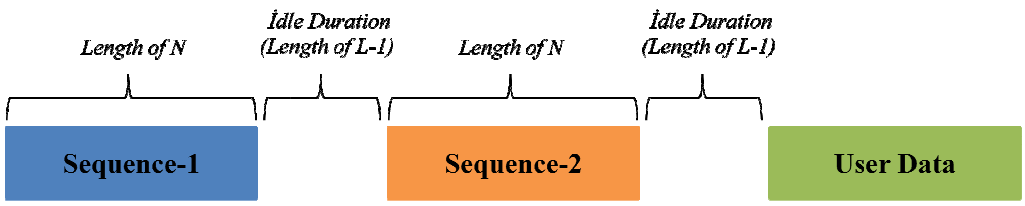


Figure 3-3: Preamble structure for ML frequency offset and channel estimation using complementary training sequences

In the preamble, two sequences in a complementary pair are placed into the channel in successive order. The sequences in the preamble do not contain any precursor or postcursor symbols. Instead of that, between complementary pair and between the preamble and the user data the channel left as idle. The aim of this application is to protect the complementary autocorrelation characteristics of the pair. For easy application and providing a continuous structured frequency offset matrix, the idle durations between the sequences should be chosen as much as $L-1$ symbol duration. The sequences in the preamble with this structure can be resembled to the training sequences with zero-amplitude precursor and postcursor.

With usage of complementary pairs and preamble structure described above, the signal model for single frequency offset case differs from the model described in Section 3.1. After transmission of the preamble, the first $N+L-1$ signal samples in the received side can be expressed in closed form as in Eq. (3-49)

$$\boldsymbol{y}_1 = \boldsymbol{F}_1 \boldsymbol{X}_1 \boldsymbol{h} + \boldsymbol{w}_1 \tag{3-49}$$

where $\boldsymbol{F}_1$ is (N+L-1)-(by-N+L-1) diagonal matrix of frequency offset, $\boldsymbol{X}_1$ is (N+L-1)-by-L matrix containing the elements of the first sequence and $\boldsymbol{h}$ is L-by-1 channel impulse response vector. $\boldsymbol{w}_1$ is (N+L-1)-by-1 AWGN vector with zero mean and variance of $\sigma_w^2/2$. Structures of $\boldsymbol{y}_1$, $\boldsymbol{F}_1$ and $\boldsymbol{X}_1$ are expressed in Eq. (3-50), Eq. (3-51) and Eq. (3-52) respectively. In the matrix $\boldsymbol{X}_1$, zeros come from the idle states of the channel before the transmission of the first sequence and between the transmissions of the first and the second sequences.

$$\boldsymbol{y}_1 = [y_1[0]\, y_1[1] \dots\, y_1[N-1]\, y_1[n] \dots\, y_1[N+L-1]]^T \tag{3-50}$$

$$\boldsymbol{F}_1 = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & e^{j2\pi v} & 0 & \cdots & 0 \\ 0 & 0 & e^{j4\pi v} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & e^{j2\pi(N+L-2)v} \end{bmatrix} \tag{3-51}$$

$$\boldsymbol{X}_1 = \begin{bmatrix} x_1[0] & 0 & 0 & \cdots & \cdots & 0 \\ x_1[1] & x_1[0] & 0 & \cdots & \cdots & 0 \\ x_1[2] & x_1[1] & x_1[0] & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & x_1[N-1] & x_1[N-2] \\ 0 & 0 & 0 & \cdots & 0 & x_1[N-1] \end{bmatrix} \tag{3-52}$$

After receiving the samples belong to the second sequence, the received signal samples can be expressed in closed form as in Eq. (3-53).

$$\boldsymbol{y}_2 = \boldsymbol{F}_2\boldsymbol{X}_2\boldsymbol{h} + \boldsymbol{w}_2 \tag{3-53}$$

In Eq. (3-51), $\boldsymbol{F}_2$ is (N+L-1)-by-(N+L-1) frequency offset matrix, $\boldsymbol{X}_2$ is (N+L-1)-by-L matrix containing the elements of the second complementary training sequence and $\boldsymbol{w}_2$ is (N+L-1)-by-1 AWGN vector in a similar manner to Eq. (3-49). The channel impulse response vector, $\boldsymbol{h}$, is the same vector as in Eq. (3-49). Structures of $\boldsymbol{y}_2$, $\boldsymbol{F}_2$ and $\boldsymbol{X}_2$ are expressed in Eq. (3-54), Eq. (3-55) and Eq. (3-56) respectively.

$$\boldsymbol{y}_2 = [y_2[0]\, y_2[1] \dots\, y_2[N-1]\, y_2[n] \dots\, y_2[N+L-1]]^T \tag{3-54}$$

$$\boldsymbol{F}_2 = \begin{bmatrix} e^{j2\pi(N+L-1)v} & 0 & 0 & \cdots & 0 \\ 0 & e^{j2\pi(N+L)v} & 0 & \cdots & 0 \\ 0 & 0 & e^{j2\pi(N+L+1)v} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & e^{j2\pi(2N+2L-3)v} \end{bmatrix} \tag{3-55}$$

$$\boldsymbol{X}_2 = \begin{bmatrix} x_2[0] & 0 & 0 & \cdots & \cdots & 0 \\ x_2[1] & x_2[0] & 0 & \cdots & \cdots & 0 \\ x_2[2] & x_2[1] & x_2[0] & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & x_2[N-1] & x_2[N-2] \\ 0 & 0 & 0 & \cdots & 0 & x_2[N-1] \end{bmatrix} \tag{3-56}$$

With a careful inspection, it is seen that the frequency offset matrix, $\boldsymbol{F}_2$, in the second observation is multiplication of the frequency offset matrix, $\boldsymbol{F}_1$, in the first observation and a complex scalar. This complex scalar comes from the phase variations during the idle state of the channel between transmissions of complementary pair.

$$\boldsymbol{F}_2 = e^{j2\pi(N+L-1)v} \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & e^{j2\pi v} & 0 & \cdots & 0 \\ 0 & 0 & e^{j4\pi v} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & e^{j2\pi(N+L-2)v} \end{bmatrix} \tag{3-57}$$

Using Eq. (3-57), the relation between matrices $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ is obtained as in Eq. (3-58).

$$\boldsymbol{F}_2 = e^{j2\pi(N+L-1)v} \boldsymbol{F}_1 \tag{3-58}$$

Collecting the signal samples $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ in the same vector, $\boldsymbol{y}$, the received signal samples are put into the closed form as in Eq. (3-59). Structures of the matrices $\boldsymbol{F}$, $\boldsymbol{X}$ and $\boldsymbol{w}$ in Eq. (3-60), Eq. (3-61) and Eq. (3-62) respectively.

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{bmatrix} = \boldsymbol{FXh} + \boldsymbol{w} \tag{3-59}$$

$$\boldsymbol{F} = \begin{bmatrix} \boldsymbol{F}_1 & 0 \\ 0 & \boldsymbol{F}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{F}_1 & 0 \\ 0 & e^{j2\pi(N+L-1)v} \boldsymbol{F}_1 \end{bmatrix} \tag{3-60}$$

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{bmatrix} \tag{3-61}$$

$$\boldsymbol{w} = \begin{bmatrix} \boldsymbol{w}_1 \\ \boldsymbol{w}_2 \end{bmatrix} \tag{3-62}$$

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{F}_1 & 0 \\ 0 & e^{j2\pi(N+L-1)v}\boldsymbol{F}_1 \end{bmatrix} \begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{bmatrix} \boldsymbol{h} + \begin{bmatrix} \boldsymbol{w}_1 \\ \boldsymbol{w}_2 \end{bmatrix} \tag{3-63}$$

In Eq. (3-63), $\boldsymbol{h}$ and v are unknown parameters to be estimated. Based on the statistical characteristics of parameters in this equation the expectation and variance of received signal samples vector $\boldsymbol{y}$ is given in Eq. (3-64) and Eq. (3-65) respectively

$$E\{\boldsymbol{y}\} = \boldsymbol{FXh} \tag{3-64}$$

$$Cov\{\boldsymbol{y}\} = \sigma_w^2 I \tag{3-65}$$

Using the statistical characteristics, the likelihood function for parameters pair $(\boldsymbol{h}, v)$ is obtained as in Eq. (3-66).

$$\Lambda(\boldsymbol{y}; \boldsymbol{h}, v) = \frac{1}{(\pi\sigma_w^2)^N} . \exp\{-\frac{1}{\sigma_w^2}[\boldsymbol{y} - \boldsymbol{FXh}]^H[\boldsymbol{y} - \boldsymbol{FXh}]\} \tag{3-66}$$

Here the aim is to find the optimal estimation values for the parameters pair $(\boldsymbol{h}, v)$ by maximizing the likelihood function over these parameters. Maximizing likelihood function equals to minimizing function ℓ in Eq. (3-67)

$$\ell = [\boldsymbol{y} - \boldsymbol{FXh}]^H[\boldsymbol{y} - \boldsymbol{FXh}] \tag{3-67}$$

Multiplying the two terms on the right side, which are Hermitian of each other, Eq. (3-67) can be rewritten like in Eq. (3-68)

$$\ell = \boldsymbol{y}^H\boldsymbol{y} - 2\boldsymbol{h}^H\boldsymbol{X}^H\boldsymbol{F}^H\boldsymbol{y} + \boldsymbol{h}^H\boldsymbol{X}^H\boldsymbol{F}^H\boldsymbol{FXh} \tag{3-68}$$

Multiplication of frequency offset matrix with its Hermitian is an identity matrix.

$$\boldsymbol{F}^H\boldsymbol{F} = \begin{bmatrix} \boldsymbol{F}_1^H & 0 \\ 0 & e^{-j2\pi(N+L-1)v}\boldsymbol{F}_1^H \end{bmatrix} \begin{bmatrix} \boldsymbol{F}_1 & 0 \\ 0 & e^{j2\pi(N+L-1)v}\boldsymbol{F}_1 \end{bmatrix}$$

$$= \begin{bmatrix} \boldsymbol{F}_1^H\boldsymbol{F}_1 & 0 \\ 0 & e^{-j2\pi(N+L-1)v}\boldsymbol{F}_1^H\boldsymbol{F}_1 e^{j2\pi(N+L-1)v} \end{bmatrix}$$

$$= \begin{bmatrix} \boldsymbol{I} & 0 \\ 0 & \boldsymbol{I} \end{bmatrix} \tag{3-69}$$

Inserting Eq. (3-69) into Eq. (3-68) the function ℓ takes the form in Eq. (3-70)

$$\ell = \boldsymbol{y}^H\boldsymbol{y} - 2\boldsymbol{h}^H\boldsymbol{X}^H\boldsymbol{F}^H\boldsymbol{y} + \boldsymbol{h}^H\boldsymbol{X}^H\boldsymbol{X}\boldsymbol{h} \tag{3-70}$$

Like ML method described in Section 3.2, the estimation process begins by accepting the frequency offset, v, as fixed and minimizing the function ℓ over the channel impulse response, $\boldsymbol{h}$. The optimum value for vector $\boldsymbol{h}$ is found by least square solution of Eq. (3-70). Least square solution is achieved by taking the derivative of function ℓ and equating it to zero. Performing this mathematical operation, the least square solution of $\boldsymbol{h}$ found as in Eq. (3-71).

$$\boldsymbol{h}_{ls} = (\boldsymbol{X}^H\boldsymbol{X})^{-1}\boldsymbol{X}^H\boldsymbol{F}^H\boldsymbol{y} \tag{3-71}$$

In Eq. (3-71), $(\boldsymbol{X}^H\boldsymbol{X})$ is L-by-L autocorrelation matrix which is sum of the autocorrelation matrices of complementary training sequences. Because of the complementary autocorrelation characteristics of complementary sequences, the autocorrelation matrix is the multiplication of the identity matrix with a scalar constant which is the sum of the energies of the complementary pair.

$$\boldsymbol{R}_{xx} = [\boldsymbol{X}_1^H\ \boldsymbol{X}_2^H] \begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{bmatrix}$$

$$= \boldsymbol{X}_1^H\boldsymbol{X}_1 + \boldsymbol{X}_2^H\boldsymbol{X}_2$$

$$= 2N\boldsymbol{I} \tag{3-72}$$

Inserting Eq. (3-72) into Eq. (3-71) the least square solution of channel impulse response takes a simpler form as in Eq. (3-73).

$$\boldsymbol{h}_{ls} = \frac{1}{2N}\boldsymbol{X}^H\boldsymbol{F}^H\boldsymbol{y} \tag{3-73}$$

Inserting $\boldsymbol{h}_{ls}$ into Eq. (3-70), the cost function, ℓ, to be minimized for frequency offset estimation is achieved.

$$\begin{aligned}\ell &= \boldsymbol{y}^H\boldsymbol{y} - 2\boldsymbol{h}^H\boldsymbol{X}^H\boldsymbol{F}^H\boldsymbol{y} + \boldsymbol{h}^H\boldsymbol{X}^H\boldsymbol{X}\boldsymbol{h} \\ &= \boldsymbol{y}^H\boldsymbol{y} - \boldsymbol{2}\frac{1}{2N}\boldsymbol{y}^H\boldsymbol{F}\boldsymbol{X}\boldsymbol{X}^H\boldsymbol{F}^H\boldsymbol{y} + \frac{1}{2N}\boldsymbol{y}^H\boldsymbol{F}\boldsymbol{X}(2N\boldsymbol{I})\boldsymbol{X}^H\boldsymbol{F}^H\boldsymbol{y}\frac{1}{2N} \\ &= \boldsymbol{y}^H\boldsymbol{y} - \frac{1}{2N}\boldsymbol{y}^H\boldsymbol{F}\boldsymbol{X}\boldsymbol{X}^H\boldsymbol{F}^H\boldsymbol{y}\end{aligned} \tag{3-74}$$

Because the multiplication $\boldsymbol{y}^H\boldsymbol{y}$ is a positive scalar, minimizing the function ℓ is equivalent to maximizing the function $g(v)$ given in Eq. (3-75) over the frequency offset, v.

$$g(v) = \boldsymbol{y}^H\boldsymbol{F}\boldsymbol{X}\boldsymbol{X}^H\boldsymbol{F}^H\boldsymbol{y} \tag{3-75}$$

With a careful inspection, it is obviously perceived that the function $g(v)$ is magnitude square of vector $\boldsymbol{X}^H\boldsymbol{F}^H\boldsymbol{y}$.

$$g(v) = |\boldsymbol{X}^H\boldsymbol{F}^H\boldsymbol{y}|^2 \tag{3-76}$$

The optimal value for the frequency offset, v, is the value maximizing the function $g(v)$. In summary, the estimator for the frequency offset is written as in Eq. (3-77) .

$$\hat{v} = \arg\max_{\hat{v}}\{g(\hat{v})\} \tag{3-77}$$

Inspecting the matrix structure in Eq. (3-60), it can be seen that the frequency offset matrix $\boldsymbol{F}$ has a continuous structure as in Eq. (3-78)

$$\boldsymbol{F} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & e^{j2\pi v} & 0 & \cdots & 0 \\ 0 & 0 & e^{j4\pi v} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & e^{j2\pi(2N+2L-3)v} \end{bmatrix} \tag{3-78}$$

Profiting from Eq. (3-78), the function $g(\hat{v})$ may be put in the form in Eq. (3-79).

$$g(\hat{v}) = -\rho_c(0) + 2Re\left\{\sum_{m=0}^{N-1} \rho_c(m)\, e^{-j2\pi n\hat{v}}\right\} \tag{3-79}$$

where $\rho_c(m)$ is a weighted correlation defined by Eq. (3-80).

$$\rho_c(m) = \sum_{n=m}^{N-1} \boldsymbol{B}(n-m,n)y[n]y^*\,[n-m] \tag{3-80}$$

Here, the matrix $\boldsymbol{B}$ is projection matrix obtained by multiplication of training sequence matrix, $\boldsymbol{X}$, with its Hermitian.

$$\boldsymbol{B} = \boldsymbol{X}\boldsymbol{X}^H \tag{3-81}$$

After finding optimum estimation for frequency offset the optimal value for channel impulse response can be found by inserting the frequency offset matrix created with optimal frequency offset estimation into Eq. (3-73).

$$\hat{\boldsymbol{h}} = \frac{1}{2N}\boldsymbol{X}^H\hat{\boldsymbol{F}}^H\boldsymbol{y} \tag{3-82}$$

Exploiting the results of maximum likelihood estimation with complementary training sequences some observations can be noted as follows.

1. The maximum likelihood (ML) approach with complementary training sequences is a two step process. In the first step frequency offset value is estimated by accepted the channel impulse response vector as fixed. In the second step channel impulse response is estimated by using the estimated value of frequency offset.
2. When the frequency offset is zero, the channel impulse response estimation in Eq. (3-82) reduces to form in Eq. (3-83).

$$\hat{\boldsymbol{h}} = \frac{1}{2N}\boldsymbol{X}^H\boldsymbol{y} \tag{3-83}$$

This result coincides with the results of LSSE channel estimation method described in Section 2.2.2. Because of perfectly diagonal autocorrelation properties of complementary sequences, the estimation in Eq. (3-83) reduces the mean square error in channel estimation.

3. Usage of complementary sequences in frequency offset estimation reduces the complexity of the estimation process. The goal becomes maximizing the magnitude square of a vector over frequency offset, v. This is not possible with some other ML estimators such as MLE #1 due to the matrix inverse operation as described in Section 3.2.1.
4. Maximization of $g(\hat{v})$ in Eq. (3-75) requires a two step search. In the first step, maximum of $g(\hat{v})$ is searched over a grid of $\hat{v}$-values and the location of the maximum is found. In the second step, the function $g(\hat{v})$ is searched by smaller steps of $\hat{v}$ and the location of maximum nearest to location found in the first step is exposed. The location found in the second step accept as optimal estimation value for the frequency offset, v.
5. The function $g(\hat{v})$ can be easily computed by FFT (Fast Fourier Transform) method by profiting the function form in Eq. (3-79). This reduces the computational complexity in grid search step of the frequency offset estimation.
6. From Eq. (3-79), it is observed that the cost function, $g(v)$ is periodic with unit period and its maxima occur at a distance of 1 from each other. Hence, the estimation range of ML frequency offset estimation with complementary sequences is $[-0.5,0.5]$ as can be seen in Figure 3-3.

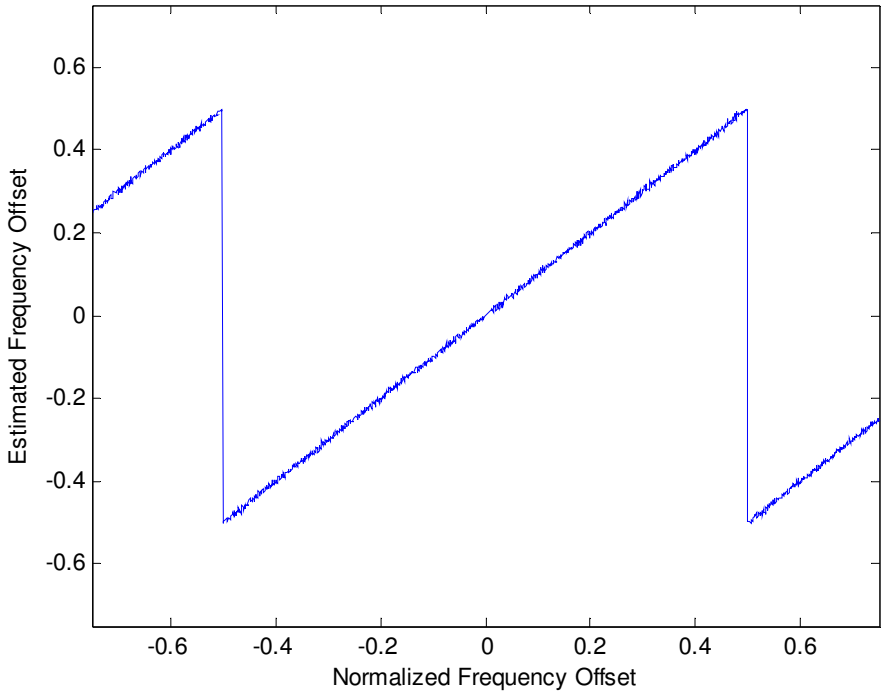


Figure 3-4: Estimated Frequency Offset vs Normalized Frequency Offset

7. Profiting from the results of [5] and [8] the expectation and variance of ML frequency offset estimation with complementary sequences are obtained as in Eq. (3-84) and Eq. (3-85) respectively

$$E\{\hat{v}\} = v \tag{3-84}$$

$$var\{\hat{v}\} = \frac{1}{16\pi^2(N+L-1)} \frac{\boldsymbol{s}^H \boldsymbol{s}}{\boldsymbol{s}^H \boldsymbol{D}^H \left(\boldsymbol{I}_{2(N+L-1)} - \frac{1}{2N}\boldsymbol{X}\boldsymbol{X}^H\right)\boldsymbol{D}\boldsymbol{s}} (SNR)^{-1} \tag{3-85}$$

In Eq. (3-85), $\boldsymbol{s} = \boldsymbol{X}\boldsymbol{h}$ and $\boldsymbol{D}$ is diagonal matrix given in Eq. (3-86)

$$\boldsymbol{D} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 2(N+L-1) \end{bmatrix} \tag{3-86}$$

We also note the following the application of complementary sequences in the presence of multiple frequency offset does not present useful simplifications. The reason for that is the disturbance of perfect autocorrelation structure of complementary sequences by mathematical model of signal structure in presence of multiple frequency offset.

3.5 Extensions

In this section we present some extensions on the work present for joint channel and frequency offset estimation. The extensions are described to illustrate the differences and similarities of the studied problem for different systems.

3.5.1 Maximum Likelihood (ML) Frequency Offset and Channel Estimation in MIMO Systems Using Complementary Training Sequences

Recently a channel estimation schemes for MIMO communication systems using complementary sequences has been reported [7]. Complementary training sequences can also be used for joint frequency offset and channel estimation of MIMO systems. However, the complementary sequences can be applied to MIMO systems under some constraints.

- The communication system should only have two transmit antennas
- For all of the different paths of the communication channel, only one frequency offset should exist.

Usage of the complementary sequences in ML frequency offset and channel estimation of MIMO systems requires a special training method. For training purposes two different preambles created from the same complementary sequences pair are transmitted from each of the transmit antennas. In the first transmission, one of the sequences in the complementary pair, $\boldsymbol{x}_1$, is sent from the first antenna while the other, $\boldsymbol{x}_2$, is transmitted from the second antenna. In the second transmission, the reverted (inverted in time) and minus signed (inverted in

amplitude) version of sequence $\boldsymbol{x}_2$ is sent from the first antenna while the reverted version of the sequence $\boldsymbol{x}_1$ is transmitted from the second antenna. In order to protect the complementary autocorrelation structure of the complementary pair, idle durations as much as $L-1$ symbol duration are left between the transmission the first and the second training sequences and between the second training sequences and the user data. The preambling scheme is illustrated in Figure 3-4.

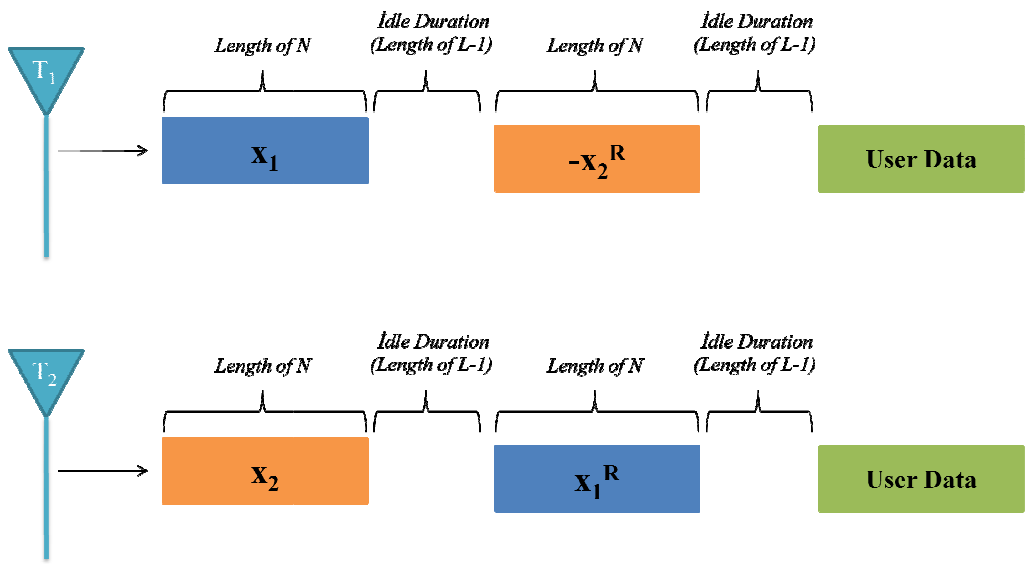


Figure 3-5: Preamble signaling structure for ML frequency offset and channel estimation using complementary training sequences (MIMO communication systems)

After the transmission of the first sequences, the received signal samples in the qth receive antenna can be modeled as in Eq. (3-87)

$$\boldsymbol{y}_q{}^1 = \boldsymbol{F}_1(v_q)\boldsymbol{X}_1\boldsymbol{h}_{1,q} + \boldsymbol{F}_1(v_q)\boldsymbol{X}_2\boldsymbol{h}_{2,q} + \boldsymbol{w}_q{}^1 \tag{3-87}$$

where $\boldsymbol{X}_1$ is (N+L-1)-by-L matrix containing the symbols in sequence $\boldsymbol{x}_1$ and $\boldsymbol{X}_2$ is (N+L-1)-by-L matrix containing symbols of the sequence $\boldsymbol{x}_2$. $\boldsymbol{h}_{1,q}$ is L-by-1 channel impulse response vector between the first transmit and the qth receive

antennas while $\boldsymbol{h}_{2,q}$ is L-by-1 vector containing channel gains between the second transmit and the qth receive antennas. $\boldsymbol{F}_1(v_q)$ is (N+L-1)-by-(N+L-1) frequency offset matrix in the form of Eq. (3-88). $\boldsymbol{w}_q$ zero-mean AWGN with variance of $\sigma_w^2/2$.

$$\boldsymbol{F}(v_q) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & e^{j2\pi v_q} & 0 & \cdots & 0 \\ 0 & 0 & e^{j4\pi v_q} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & e^{j2\pi(N+L-2)v_q} \end{bmatrix} \tag{3-88}$$

After the transmission of the second sequences the received signal samples in the qth receive antenna is given as

$$\boldsymbol{y}_q{}^2 = -\boldsymbol{F}_2(v_q)\boldsymbol{X}_2^R\boldsymbol{h}_{1,q} + \boldsymbol{F}_2(v_q)\boldsymbol{X}_1^R\boldsymbol{h}_{2,q} + \boldsymbol{w}_q{}^2 \tag{3-89}$$

where $\boldsymbol{F}_2(v_q)$ is (N+L-1)-by-(N+L-1) frequency offset matrix in the form of Eq. (3-90).

$$\boldsymbol{F}_2(v_q) = e^{j2\pi(N+L-1)v_q} * \boldsymbol{F}_1(v_q) \tag{3-90}$$

After the training, total received signal samples at the receive antenna may be put in the form of Eq. (3-91) and (3-92).

$$\boldsymbol{y}_q = \begin{bmatrix} \boldsymbol{y}_q{}^1 \\ \boldsymbol{y}_q{}^2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{F}_1(v_q) & 0 \\ 0 & e^{j2\pi T_i v_q} * \boldsymbol{F}_1(v_q) \end{bmatrix} \begin{bmatrix} \boldsymbol{X}_1 & \boldsymbol{X}_2 \\ -\boldsymbol{X}_2^R & \boldsymbol{X}_1^R \end{bmatrix} \begin{bmatrix} \boldsymbol{h}_{1,q} \\ \boldsymbol{h}_{2,q} \end{bmatrix} + \begin{bmatrix} \boldsymbol{w}_q{}^1 \\ \boldsymbol{w}_q{}^2 \end{bmatrix} \tag{3-91}$$

$$\boldsymbol{y}_q = \boldsymbol{F}(v_q)\boldsymbol{X}_q\boldsymbol{h}_q + \boldsymbol{w}_q \tag{3-92}$$

Using Eq. (3-92), likelihood function for received signal at the qth receive antenna is obtained as in Eq. (3-93).

$$\Lambda(\boldsymbol{y}_q; \boldsymbol{h}_q, v_q) = \frac{1}{(\pi\sigma_w^2)^N} . \exp\left\{-\frac{1}{2\sigma_w^2}\left[\boldsymbol{y}_q - \boldsymbol{F}(v_q)\boldsymbol{X}_q\boldsymbol{h}_q\right]^H\left[\boldsymbol{y}_q - \boldsymbol{F}(v_q)\boldsymbol{X}_q\boldsymbol{h}_q\right]\right\} \tag{3-93}$$

Our aim is to maximize Eq. (3-93) over parameters v_q and $\boldsymbol{h}_q$. Maximizing Eq. (3-93) is equivalent to minimize the function in Eq. (3-94).

$$\ell_q = \left[\boldsymbol{y}_q - \boldsymbol{F}(v_q)\boldsymbol{X}_q\boldsymbol{h}_q\right]^H\left[\boldsymbol{y}_q - \boldsymbol{F}(v_q)\boldsymbol{X}_q\boldsymbol{h}_q\right] \tag{3-94}$$

Since multiplication of frequency offset matrix $\boldsymbol{F_1}(v_q)$ with its Hermitian is identity, the function ℓ_q takes a simpler form as in Eq. (3-95)

$$\ell_q = \boldsymbol{y}_q^H\boldsymbol{y}_q - 2\boldsymbol{h}_q^H\boldsymbol{X}_q^H\boldsymbol{F}^H\left(v_q\right)\boldsymbol{y}_q + \boldsymbol{h}_q^H\boldsymbol{X}_q^H\boldsymbol{X}_q\boldsymbol{h}_q \tag{3-95}$$

As in ML estimators in previous section, estimation process begins by accepting the frequency offset, v_q, as fixed and finding least squares estimation scheme of the channel impulse response by taking the derivative of function ℓ_q and equating the derivative to zero.

$$\boldsymbol{h}_{q,ls} = (\boldsymbol{X}_q^H\boldsymbol{X}_q)^{-1}\boldsymbol{X}_q^H\boldsymbol{F}^H\left(v_q\right)\boldsymbol{y}_q \tag{3-96}$$

In Eq. (3-96), multiplication $\boldsymbol{X}_q^H\boldsymbol{X}_q$ has the form of Eq. (3-97).

$$\boldsymbol{X}_q^H\boldsymbol{X}_q = \begin{bmatrix} \boldsymbol{X}_1^H\boldsymbol{X}_1 + \boldsymbol{X}_2^H\boldsymbol{X}_2 & \boldsymbol{X}_1^H(-\boldsymbol{X}_2^R) + \boldsymbol{X}_2^H\boldsymbol{X}_1^R \\ (-\boldsymbol{X}_2^H)\boldsymbol{X}_1 + \boldsymbol{X}_1^R\boldsymbol{X}_2 & (-\boldsymbol{X}_2^R)^H(-\boldsymbol{X}_2^R) + (\boldsymbol{X}_1^R)^H\boldsymbol{X}_1^R \end{bmatrix} \tag{3-97}$$

From the auto-correlation ($\boldsymbol{r}_{u^R} = \boldsymbol{r}_u$) and cross-correlation ($\boldsymbol{r}_{z^R u^R} = r_{uz}$) properties of reverted sequences, multiplication $\boldsymbol{X}_q^H\boldsymbol{X}_q$ reduces to a 2L-by-2L identity matrix multiplied by total energy of complementary pair ($2N$) [7].

$$\boldsymbol{X}_q^H\boldsymbol{X}_q = 2N\boldsymbol{I}_{2L} \tag{3-98}$$

Using Eq. (3-98) $\boldsymbol{h}_q$ can be rewritten as in Eq. (3-99).

$$\boldsymbol{h}_q = \frac{1}{2N}\boldsymbol{X}_q^H\boldsymbol{F}^H\left(v_q\right)\boldsymbol{y}_q \tag{3-99}$$

Inserting Eq. (3-99) into Eq. (3-95) function ℓ_q takes the form of Eq. (3-100).

$$\begin{aligned} \ell_q &= \boldsymbol{y}_q^H\boldsymbol{y}_q - \frac{1}{2N}\boldsymbol{y}_q^H\boldsymbol{F}(v_q)\boldsymbol{X}_q\boldsymbol{X}_q^H\boldsymbol{F}^H\left(v_q\right)\boldsymbol{y}_q \\ &= \boldsymbol{y}_q^H\boldsymbol{y}_q - \frac{1}{2N}\left|\boldsymbol{X}_q^H\boldsymbol{F}^H\left(v_q\right)\boldsymbol{y}_q\right|^2 \end{aligned} \tag{3-100}$$

It is obviously seen that minimizing function ℓ_q is equivalent to maximize the function $g(v_q)$ defined by Eq. (3-101).

$$g(v_q) = \left|\boldsymbol{X}_q^H \boldsymbol{F}^H(v_q)\boldsymbol{y}_q\right|^2 \tag{3-101}$$

Function $g(v_q)$ in Eq. (3-101) assumes $\boldsymbol{h}_q$ is fixed and only variable is the frequency offset v_q. Optimum estimation value for the frequency offset v_q is found by searching the maximum value of function $g(v_q)$ over parameter v_q. By this way frequency offset estimator is obtained as in Eq. (3-102).

$$\begin{aligned}\hat{v}_q &= \arg\max_{\hat{v}_q}\{g(\hat{v}_q)\} \\ &= \arg\max_{\hat{v}_q}\{\left|\boldsymbol{X}_q^H \boldsymbol{F}^H(\hat{v}_q)\boldsymbol{y}_q\right|^2\}\end{aligned} \tag{3-102}$$

Estimation range of frequency offset estimator in Eq. (3-98) is $[-0.5,0.5]$. After finding optimum estimation for the frequency offset, optimum estimation of the channel impulse response is found by inserting estimated frequency offset into Eq. (3-99).

$$\hat{\boldsymbol{h}}_q = \frac{1}{2N}\boldsymbol{X}_q^H \boldsymbol{F}^H(\hat{v}_q)\boldsymbol{y}_q \tag{3-103}$$

As it is stated at the beginning of this section, usage of complementary sequences in joint frequency offset and channel estimation of MIMO systems is a special application which is only applicable to two transmit antenna systems. Application of complementary sequences in multiple transmit antenna systems has not fully examined and needs further investigation

3.5.2 Maximum Likelihood (ML) Frequency Offset and Channel Estimation in Presence of Multiple Frequency Offset

Based on the signal model defined by Eq. (3-7), for the case of multiple frequency offset, a maximum likelihood frequency offset and channel estimation approach

have been introduced by [10]. Using closed form signal model given in Eq. (3-7) and statistical characteristics of the signal and noise the likelihood function is obtained as

$$\Lambda(\boldsymbol{y};\boldsymbol{h},v) = \frac{1}{(\pi\sigma_w^2)^N}.\exp\left\{-\frac{1}{\sigma_w^2}[\boldsymbol{y}-\boldsymbol{A}\boldsymbol{h}]^H[\boldsymbol{y}-\boldsymbol{A}\boldsymbol{h}]\right\} \tag{3-104}$$

where $\boldsymbol{A}$ is the matrix whose elements are defined by Eq. (3-8), $\boldsymbol{h}$ is the channel impulse response and σ_w^2 is the variance of the noise in the channel.

Like ML based methods defined in previous sections estimation begins by finding the optimum channel impulse response maximizing the likelihood function. Taking the derivative of the function given in Eq. (3-105) over $\boldsymbol{h}^H$ and equating the derivative to zero the least square estimation scheme for the channel impulse response is achieved as in Eq. (3-106).

$$\ell = [\boldsymbol{y}-\boldsymbol{A}\boldsymbol{h}]^H[\boldsymbol{y}-\boldsymbol{A}\boldsymbol{h}] \tag{3-105}$$

$$\boldsymbol{h}_{ls} = (\boldsymbol{A}^H\boldsymbol{A})^{-\mathbf{1}}\boldsymbol{A}^H\boldsymbol{y} \tag{3-106}$$

Inserting Eq. (3-106) into Eq. (3-105) the function to maximize is obtained as in Eq. (3-107)

$$g(v) = \boldsymbol{y}^H\boldsymbol{A}(\boldsymbol{A}^H\boldsymbol{A})^{-\mathbf{1}}\boldsymbol{A}^H\boldsymbol{y} \tag{3-107}$$

In order to correctly estimate all of the frequency offsets in the channel the writers of [10] advise to use an impulsive training sequences which form a perfectly diagonal autocorrelation matrix. Under this condition, $\boldsymbol{A}^H\boldsymbol{A}$ is dominated by large diagonal terms and function $g(v)$ is reduced to the simpler form in Eq. (3-108)

$$\begin{aligned} g(v) &= \boldsymbol{y}^H\boldsymbol{A}\boldsymbol{A}^H\boldsymbol{y} \\ &= \sum_{l=0}^{L-1}\left|\sum_{n=0}^{N-1} y^*[n]x[n-l]e^{2\pi v_l n}\right|^2 \end{aligned} \tag{3-108}$$

Using the function form in Eq. (3-100), the frequency offset estimator is obtained as

$$\hat{v}_l = \arg\max_{\hat{v}_l}\{\left|\sum_{n=0}^{N-1} y^*[n]x[n-l]e^{2\pi v_l n}\right|^2\} \tag{3-109}$$

where v_l is the frequency offset in lth channel path.

Each individual frequency offset is found by making the search defined by Eq. (3-109) over all possible frequencies. This requires number of L search for L different frequency offsets. After finding the frequency offsets, optimum value for the channel impulse response is calculated by Eq. (3-82).

The presence of different frequency offsets in different paths of a communication channel is most challenging case in frequency offset estimation literature. Although an estimation approach for this case has been defined by [10], this approach is not easy to apply in practice. Optimal training sequences for multiple frequency offsets are not explicitly known and need to be investigated.

CHAPTER 4

NUMERICAL RESULTS

In this chapter, results of the simulations which have been performed in this thesis study are introduced. In Section 4.1, the channel model used in simulations is presented. In Section 4.2, frequency offset and channel estimation performances of MLE#1, MLE#2, AHE and MLEWCS methods are investigated. The performances of the methods are compared in Section 4.3. Finally in Section 4.4, estimation performance of MLEWCS method in MIMO systems is investigated.

4.1 Channel Model

For simulations, the following communication channel model with six paths has been used [5]. The channel taps have the form of Eq. (4-1).

$$h(k) = \sum_{i=0}^{5} A_i \, g_T(kT_s - \tau_i - t_0) \tag{4-1}$$

In Eq. (4-1), $\{A_i\}$ are channel attenuations and chosen as independent Gaussian random variables with zero mean and variances $\{0.50, 1.00, 0.63, 0.25, 0.16, 0.10\}$. T_s is symbol duration and τ_i is delays of paths. The normalized path delays τ_i/T_s are chosen as $\{0, 0.054, 0.135, 0.432, 0.621, 1.351\}$. The timing phase, t_0, is chosen equal to $3T_s$. $g_T(t)$ is raised-cosine roloff filter assumed as used in transmitter to shape symbol pulses, with a rolloff factor of 0.5.

In the simulations, we have generated three taps ($L = 3$) and eight taps ($L = 8$) channel impulse response sequences by sampling the continuous channel model as shown in Eq. (4-1). That is for the three taps channel, $h[0]$, $h[1]$ and $h[2]$ have been generated through the Eq. (4-1).

4.2 Performances of Joint Frequency Offset and Channel Estimation Methods in SISO Communication Systems

Computer simulations have been run to investigate the performance of joint frequency offset and channel estimation with different types of training sequences. For all simulations BPSK signal formats have been used. Mapping $0 \rightarrow -1, 1 \rightarrow 1$ has been applied for simulations of transmissions.

In each simulation run of ML based estimation schemes, frequency offset has been estimated with a two step search. In the first step the frequency offset value that makes the cost function maximum has been found. In the second step a fine search has been performed near the maximum of cost function to find optimum frequency offset estimation. A typical grid and fine search process is illustrated in Figure 4-1. In the grid search, cost function, $g(v)$, is searched over the range of $[-0.5,0.5]$. with a bigger value of the frequency offset, while in the fine search, the cost function is searched over the range of $[-0.075,0.025]$ with a smaller step.

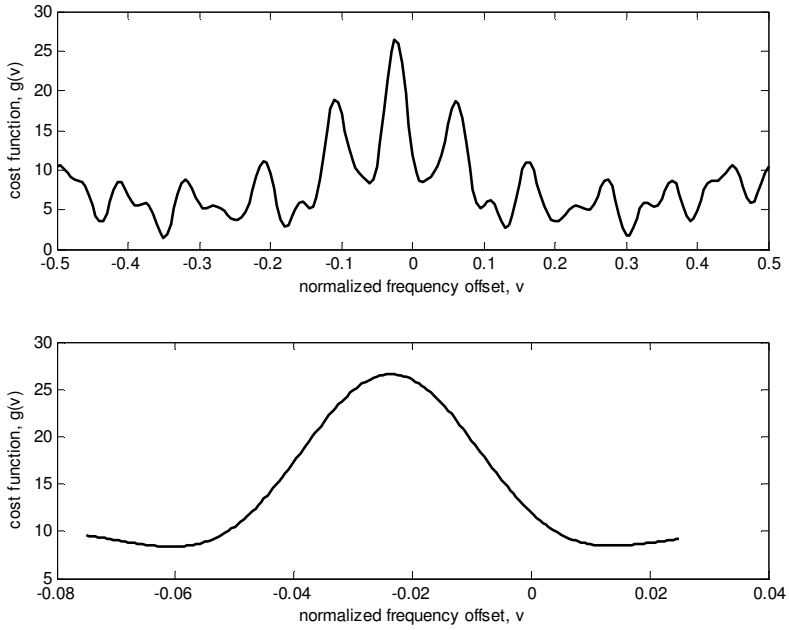


Figure 4-1: Grid and fine search for frequency offset

For each simulation 1000 Monte-Carlo realization have been realized and average values of realization results are used as simulation results. In all of the simulations the performances of the estimation schemes have been observed with the change of received SNR values calculated from Eq. (3-13). SNR value has been scanned from 0 to 30 dB.

4.2.1 Performance of MLE#1

Computer simulations have been run to see the performance of MLE#1 with different types of training sequences. The training sequences which have been used in simulations are given in Table 4-1. In the table, TS#1 and TS#2 are the training sequences which are designed for optimum LSSE channel estimation and they include non-zero and zero amplitude precursors respectively [1] (See Appendix B). TS#3 and TS#4 are the sequences chosen from [2] because of their nearly impulsive autocorrelation characteristics. Finally, TS#5 is a training sequence designed for joint ML based frequency offset and channel estimation [8].

Table 4-1: Training sequences used with MLE#1

Training Sequence	In Binary Format	In Hexadecimal Format
TS#1	0000100100	024
TS#2	0000001001	009
TS#3	0010111110011101	2F9D
TS#4	00110111010110111100	375BC
TS#5	10110111	B7

4.2.1.1 Simulation-1

Simulation-1 has been performed to observe the performance of TS#1 and TS#2 with MLE#1. In each run, the first two bits of the sequences have been used as precursors. As communication medium, a three taps channel ($L = 3$) modeled in accordance with model given in Section 4.1 has been taken. Frequency offset values to be estimated have been chosen from the range of [-0.48, 0.48] randomly. The CRBs for frequency offset estimations have been calculated from RHS of Eq. (3-25).

Mean square error (MSE) curve for frequency offset estimation with TS#1 is illustrated in Figure 4-2. For comparison purposes a minimum CRB curve obtained in 1000 simulation runs is also illustrated in the same figure. Analyzing the figure, it is seen that MSE curve follows the CRB for intermediate and high SNR levels. In high SNR, MSE reaches a level of 10^{-6}. In low SNR level, MSE is distant from the CRB. The reason for this behavior is the effects of noise which makes the maximum of cost function, $g(v)$, faraway from the true frequency offset value. An illustration summarizing this situation is given in Figure 4-4. In the figure, the cost

function $g(v)$ is plotted for $0, 5, 10$ and $20\ dB$ values of the receive SNR. The peak of cost function becomes more clear with increasing the SNR value while in $0\ dB$ SNR cost function has several peak points having close values.

Normalized mean square error (MSE) curve for channel estimation with TS#1 is illustrated in Figure 4-3. As a reference, a normalized MSE curve in the case of ideal frequency offset recovery is also shown in the same figure. Because of the perfectly diagonal autocorrelation structure of TS#1 for $L = 3$, the ideal MSE curve constitutes a minimum MSE bound for length of 8 training sequence in the conditions of Simulation-1. From the same figure it is observed that the normalized MSE curve follows the ideal curve in intermediate and high SNR levels. The larger difference in low SNR level comes from the large error in frequency offset estimation in this region.

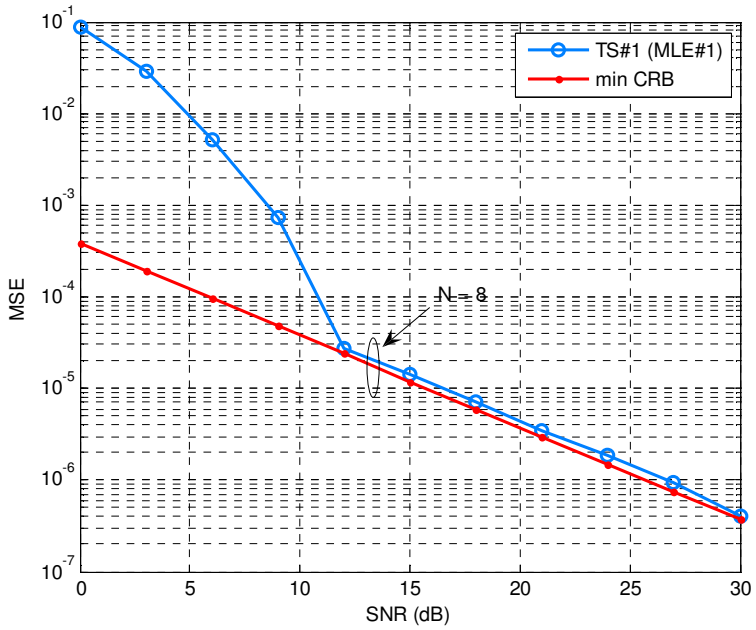


Figure 4-2: MSE for frequency offset estimation with TS#1 (MLE#1) (three taps channel)

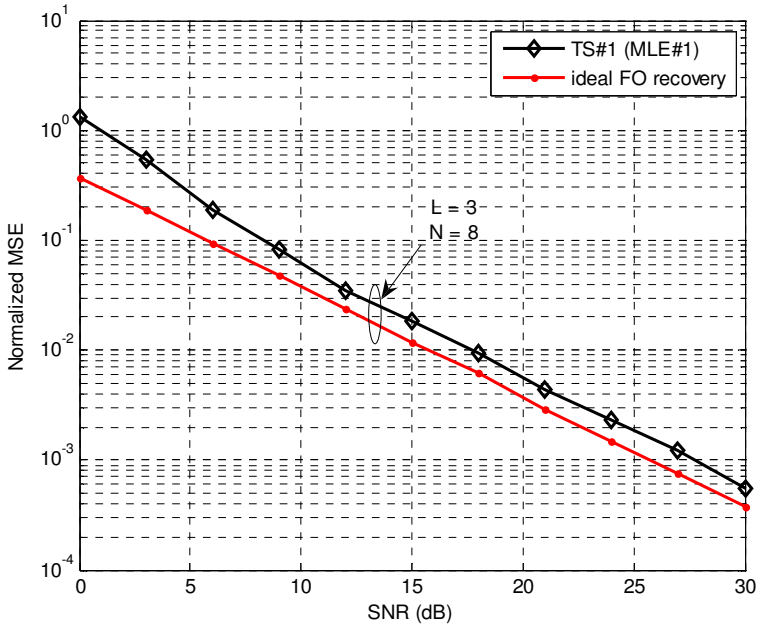


Figure 4-3: Normalized MSE for channel estimation with TS#1 (MLE#1) (three taps channel)

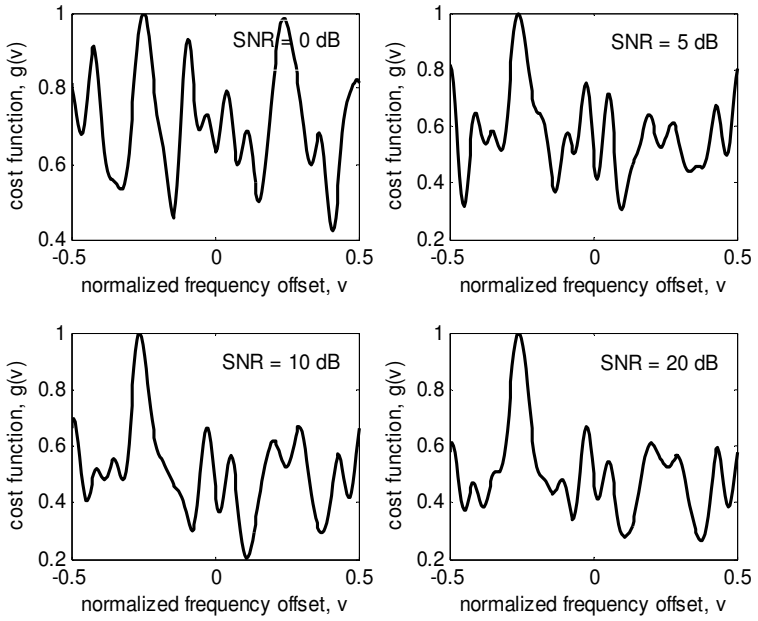


Figure 4-4: Cost function for 0, 5, 10 and 20 dB receive SNR values

MSE curve for frequency offset estimation with TS#2 is plotted in Figure 4-5 with minimum CRB. As can be seen from the figure, the MSE curve follows the CRB for intermediate and high SNR regions and reaches a 10^{-5} MSE level for high SNRs. In low SNR, the MSE curve is far away from the CRB because of the effect of the noise.

Normalized MSE curve for channel estimation with TS#2 is shown in Figure 4-6 in the existence of normalized MSE curve in the case of ideal frequency offset recovery. From the figure, it is seen that normalized MSE curve flows the ideal curve for all SNR levels with a difference of nearly 7 dB. The reason such a big difference is larger estimation error seen in frequency offset estimation with TS#2 compared with the usage of TS#1.

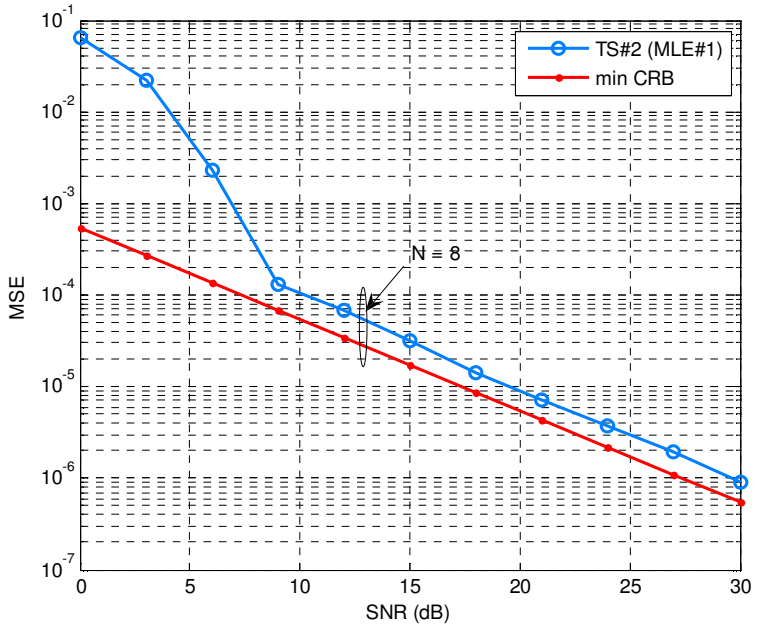


Figure 4-5: MSE for frequency offset estimation with TS#2 (MLE#1) (three taps channel)

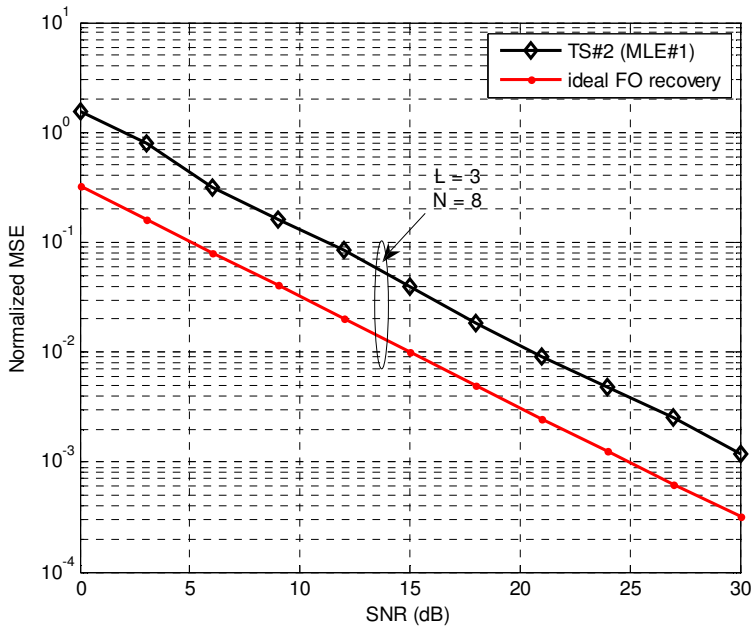


Figure 4-6: Normalized MSE for channel estimation with TS#2 (MLE#1) (three taps channel)

With a comparison of MSE and normalized MSE curves, it is obviously comprehended that TS#1 is better than TS#2 in both of the frequency offset and channel estimation. The reason for that goodness is the transmission of more signal energy with TS#1 due to the non-zero amplitude precursor. In addition, another reason is the perfectly diagonal autocorrelation matrix structure of TS#1.

4.2.1.2 Simulation-2

Simulation-2 has been run in order to see the performance of TS#3 and TS#4 with MLE#1 and effect of the length of the training sequences in estimation. As communication medium, an eight gains channel model ($L = 8$) designed according to the model given in Section 4.1 has been used. In each run, frequency offset values have been chosen from the range of [-0.48, 0.48] randomly.

In Figure 4-7, MSE curves for frequency offset estimation with TS#3 and TS#4 are illustrated. Min CRBs obtained from RHS of Eq. (3-25) in 1000 simulation runs are also plotted in the same graph. From the inspection of the figure, it is seen that both of the MSE curves follow their CRBs for intermediate and high SNR levels. As it is expected, TS#4 exhibits a better frequency offset estimation performance than TS#3. The reason for that is the transmission of more signal energy with TS#4 and using more information provided by excess bits of TS#4.

Normalized MSE curves for channel estimation with TS#3 and TS#4 are shown in Figure 4-8. As it is seen from the figure, normalized MSE curves follow their ideal curves for all SNR levels. As in the frequency offset estimation, longer length training sequence (TS#4) shows a better performance in the channel estimation and create a normalized MSE nearly coinciding ideal MSE curve of TS#3. Although TS#3 and TS#4 have not been designed for LSSE channel estimation purposes, they have a good performance in LSSE estimation. The reason of that is nearly impulsive autocorrelation characteristics of the sequences.

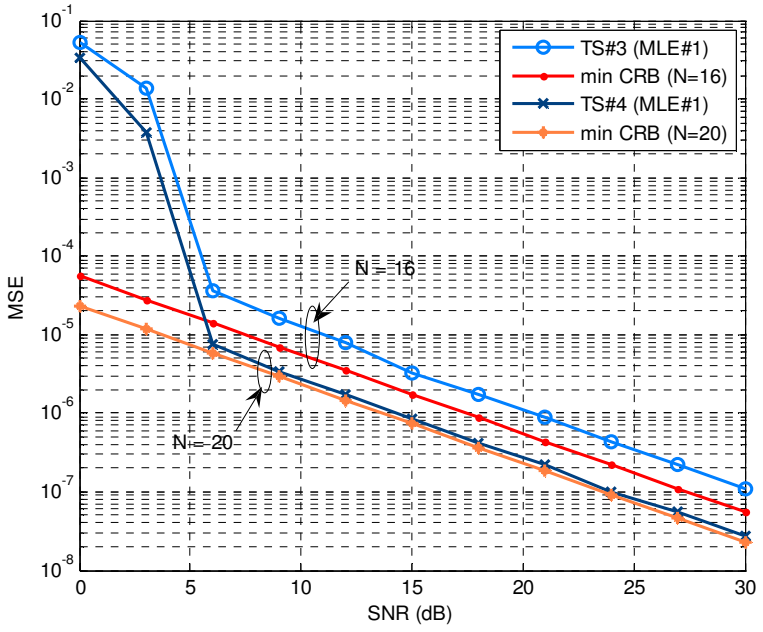


Figure 4-7: MSE for frequency offset estimation with TS#3 and TS#4 (MLE#1) (eight taps channel)

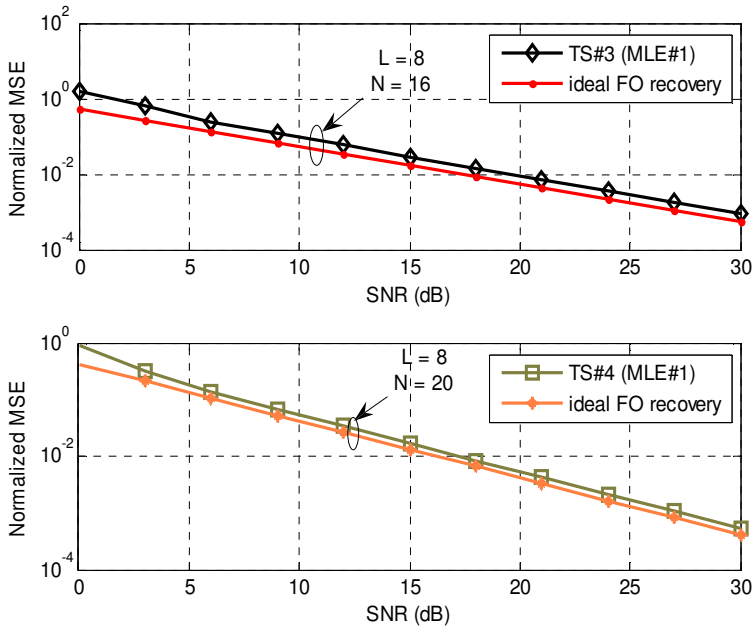


Figure 4-8: Normalized MSE for channel estimation with TS#3 and TS#4 (MLE#1) (eight taps channel)

4.2.1.3 Simulation-3

In order to analyze the performance of TS#5 with MLE#1, Simulation-5 has been carried out. A three taps ($L = 3$) channel model has been used as communication medium. The frequency offset values to be estimated have been chosen from the range of [-0.48, 0.48] randomly.

MSE curve for frequency offset estimation with TS#5 is shown in Figure 4-9. For comparison purposes a min CRB provided as a result of 1000 simulation run is also plotted in the same figure. CRB has been calculated from RHS of Eq. (3-25). As it is observed from the figure, MSE curve follows the CRB for intermediate and high SNR levels. With a careful inspection of the graph it can be seen that MSE curve tends to make a floor in high SNR region. The reason of this floor is the micro-bias appearing in the fine search step of the estimation process.

Normalized MSE curve for channel estimation with TS#5 is displayed in Figure 4-10 in the existence of normalized MSE curve of ideal frequency offset recovery case. As in the case of TS#1, ideal MSE curve is a minimum MSE bound for length of 8 training sequences due to the perfectly diagonal autocorrelation matrix structure of TS#5 (for $L = 3$). As it can be seen from the figure, the normalized MSE curve follows the ideal MSE for intermediate and high SNR levels. In low SNR level, normalized MSE is distant from the ideal curve due to the large error seen in frequency offset estimation.

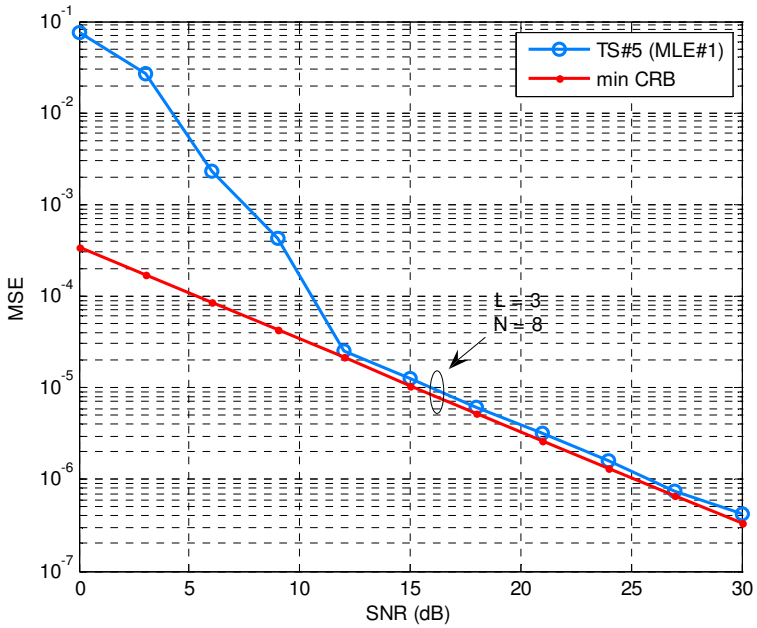


Figure 4-9: MSE for frequency offset estimation with TS#5 (MLE#1) (three taps channel)

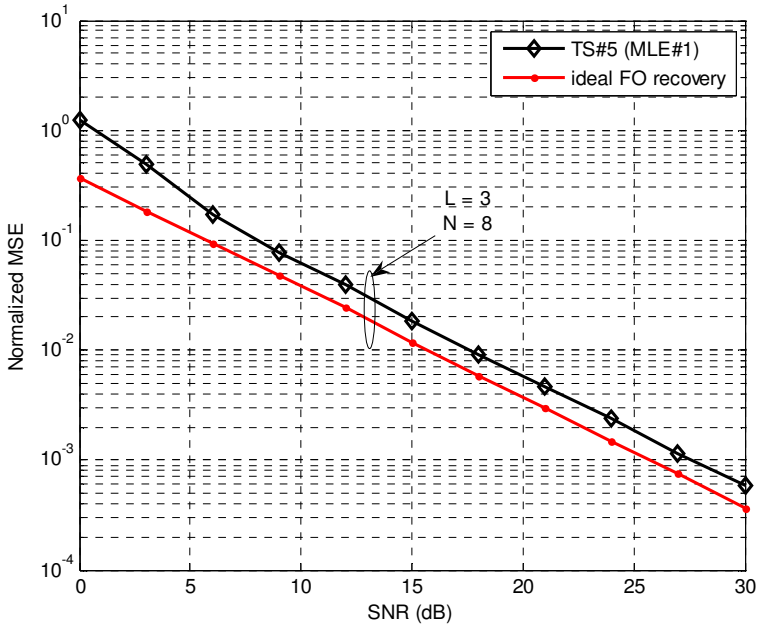


Figure 4-10: Normalized MSE for channel estimation with TS#5 (MLE#1) (three taps channel)

4.2.2 Performance of MLE#2

A computer simulation has been realized to analyze the performance of MLE#2. For simulation a periodic training sequence constructed by repeating the symbol block $C2$ four times has been used. The training sequence is given in Table 4-2.

Table 4-2: Training sequence used with MLE#2

Training Sequence	In Binary Format	In Hexadecimal Format
TS#6	11000010110000101100001011000010	C2C2C2C2

4.2.2.1 Simulation-4

Simulation-4 has been carried out to inspect the performance of TS#6 with MLE#2. An eight taps channel model designed in accordance with the model given in Section 4.1 has been applied. In order to provide the matrix structure in Eq. (3-30) last seven bits of the fixed symbol block $C2$ have been used as precursors before the transmission of TS#6. In each simulation run, the frequency offset values have been chosen from the range of [-0.05, 0.05] due to the narrow estimation range of MLE#2. CRB has been calculated from RHS of Eq. (3-36).

MSE curve for frequency offset estimation with TS#6 is shown in Figure 4-11. CRB is also illustrated in the same figure. From the inspection of the figure, it can be said that the MSE curve coincides with the CRB nearly for all SNR levels and reaches a MSE level of 10^{-8} in high SNR with a smaller number of search compared to MLE#1. In addition, in $0\ dB$ SNR, MSE exhibits a level of 10^{-5} which is a smaller value compared with the MSE level of MLE#1 in this SNR level. As a consequence MLE#2 exhibits a good frequency offset estimation performance in addition to its lower computational complexity.

Normalized MSE curve for channel estimation with TS#6 is plotted in Figure 4-12. As a reference, a normalized MSE in the case of ideal frequency offset recovery is also shown in the same figure. From the investigation of the figure, it is seen that the normalized MSE curve follows the ideal MSE for all SNR regions.

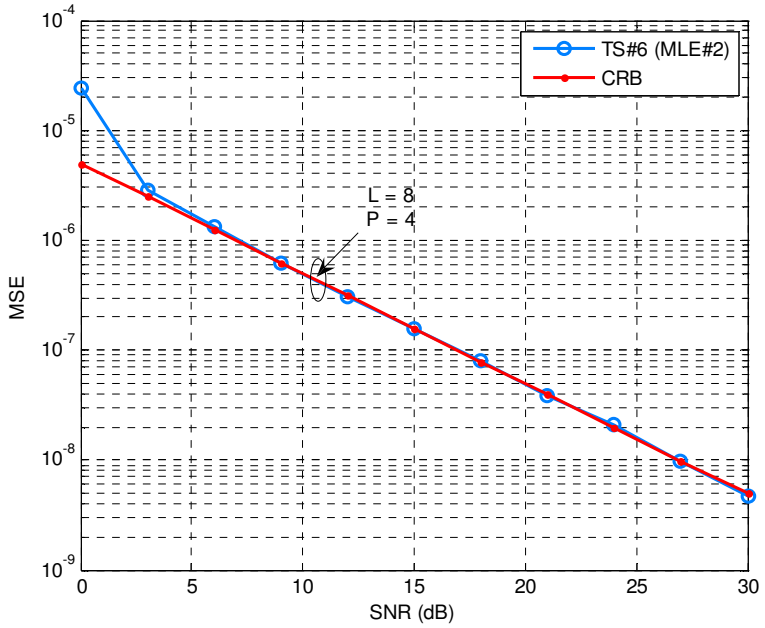


Figure 4-11: MSE for frequency offset estimation with TS#6 (MLE#2) (eight taps channel)

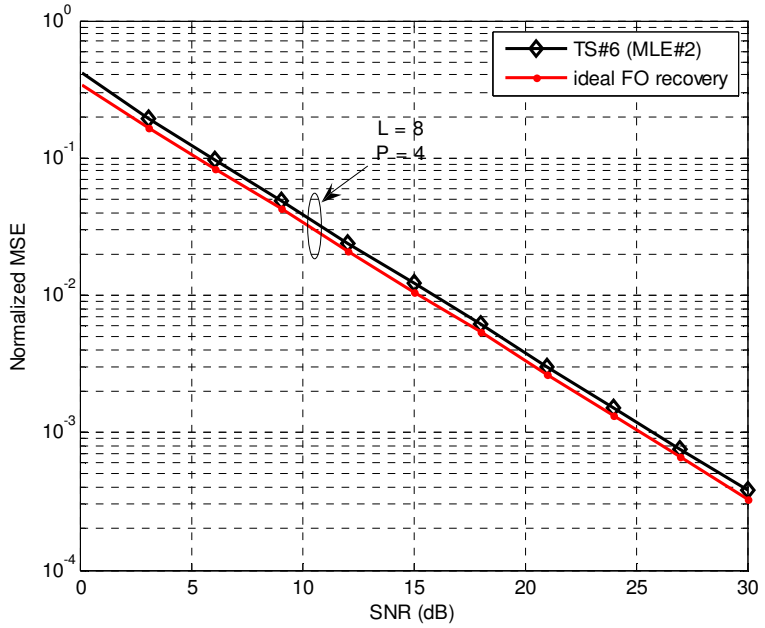


Figure 4-12: Normalized MSE for channel estimation with TS#6 (MLE#2) (eight taps channel)

4.2.3 Performance of AHE

In order to see the performance of AHE with the change of SNR, Simulation-5 has been performed.

4.2.3.1 Simulation-5

For Simulation-5, TS#6 used in Simulation-4 has been taken as periodic training sequence ($P = 4$). An eight taps communication channel has been designed for using as communication medium. In each simulation run the frequency offset values have been selected from the range of [-0.05, 0.05] randomly. A value of 2 has been chosen as design parameter M. The CRB has been calculated from the RHS of Eq. (3-48).

MSE versus SNR for AHE is shown in Figure 4-13. For comparison purposes the CRB is also illustrated in the same figure. Investigating the figure, it is obviously observed that MSE curve trails the CRB and AHE reaches a very accurate estimation in high SNR with a very small computational complexity compared with MLE#1 and MLE#2. In low SNR, AHE also shows a small MSE level compared to MLE#1.

In Figure 4-14, normalized MSE curve for channel estimation is plotted. The channel estimation has been fulfilled substituting the estimated value of frequency offset provided by AHE into Eq. (3-23). Examining the figure, it can be said that normalized MSE curve follows the ideal MSE for all SNR levels.

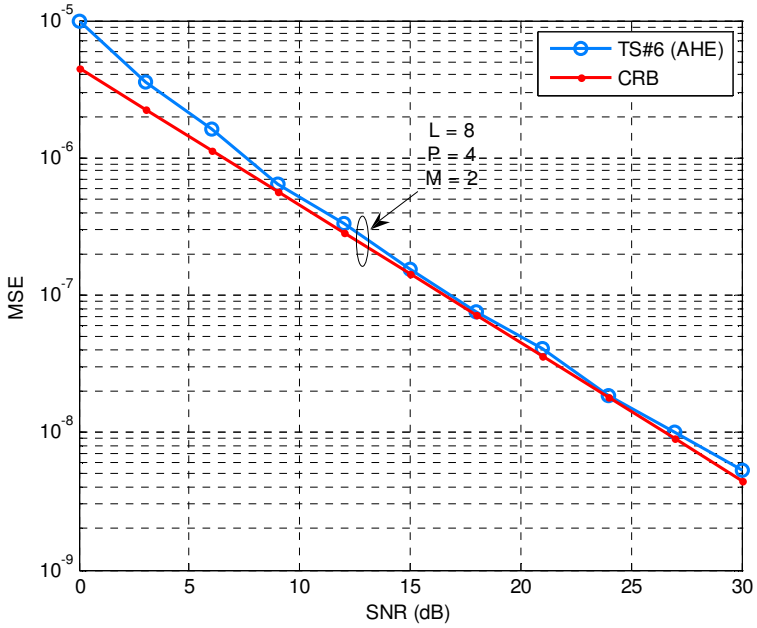


Figure 4-13: MSE for frequency offset estimation with TS#6 (AHE) (eight taps channel)

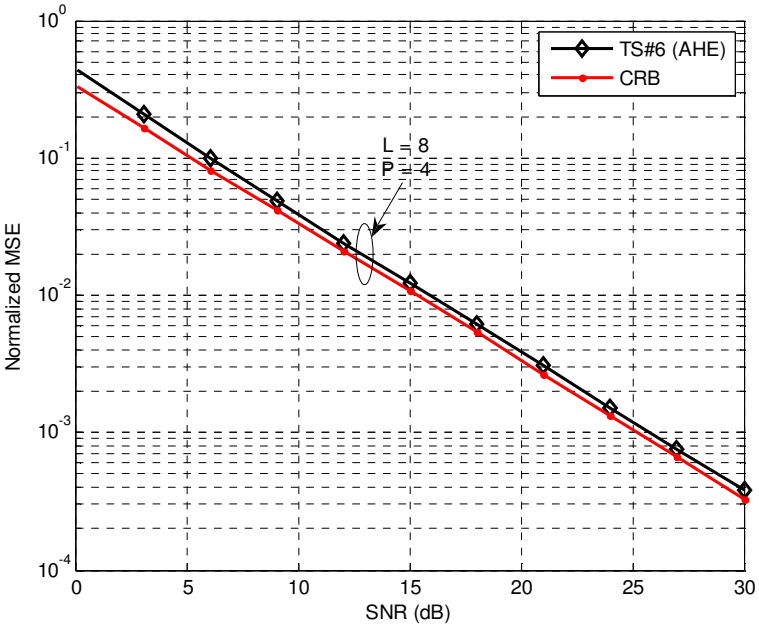


Figure 4-14: Normalized MSE for channel estimation with TS#6 (AHE) (eight taps channel)

4.2.4 Performance of MLEWCS

For analyzing the performance of ML based frequency offset and channel estimation method with complementary training sequences (MLEWCS) a simulation has been carried out. As training sequences complementary pairs (TS#7 and TS#8) shown in Table 4-3 have been used. TS#7 in the table has been chosen from the length of 8 complementary pair examples introduced by Golay [12]. TS#8 has been constructed from TS#7 using the property#8 in Section 2.4.3. Because of the usage of the complementary sequences as pairs, the preambles constructed by the complementary pairs constitute length of $2N$ training sequences. Therefore TS#7 in Table 4-3 forms a length of 16 training sequence while TS#8 constitutes a length of 32 training sequence.

Table 4-3: Complementary pairs used with ML estimator with complementary training sequences

Training Sequence	Sequence-1	Sequence-2
TS#7	00011101	00010010
TS#8	0001110100010010	0001110111101101

4.2.4.1 Simulation-6

Simulation-6 has been run to observe the performance of complementary sequences in frequency offset and channel estimation and to see effect of the length of the sequences on estimation performance. As in some of the previous simulations a communication channel model with eight taps designed according to the model in Section 4.1 has been applied. In simulation runs, the frequency offset values to be estimated have been taken randomly from the range of [-0.48, 0.48]. The CRBs has been calculated from the RHS of Eq. (3-85).

MSE curves for frequency offset estimation with complementary training sequences (TS#7 and TS#8) are illustrated in Figure 4-15. In the same figure minimum CRBs provided as result of 1000 simulation runs are also plotted. From the investigation of the figure, it is seen that the MSE curve for TS#7 trails its CRB for intermediate and high SNR levels while MSE of TS#8 coincides the CRB nearly in all levels beginning from $2.5\ dB$ SNR. The large MSE levels in low SNR region caused by the effect of the noise makes the maximum of cost function faraway from true frequency offset value. As it is expected the estimation performance gets better with increment of the length of the training sequences.

Normalized MSE curves for channel estimation with TS#7 and TS#8 are shown in Figure 4-16 with their ideal normalized MSE curves. Because of the perfectly diagonal complementary autocorrelation characteristics of the complementary sequences, the ideal normalized MSE curves in the figure form minimum MSE bounds for the LSSE channel estimation with length of 16 and length of 32 training sequences. From analyze of the figure it can observed that the normalized MSE curves of training sequences follow their ideal curves. The larger difference between the normalized MSEs and ideal MSEs in low SNR levels come from the large error in frequency offset estimation in this SNR level.

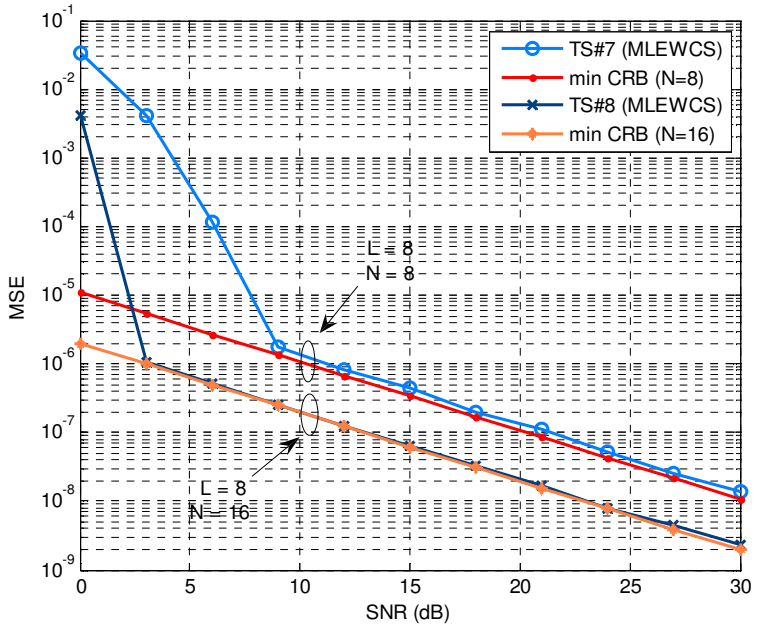


Figure 4-15: MSE for frequency offset estimation with TS#7 and TS#8 (MLEWCS) (eight taps channel)

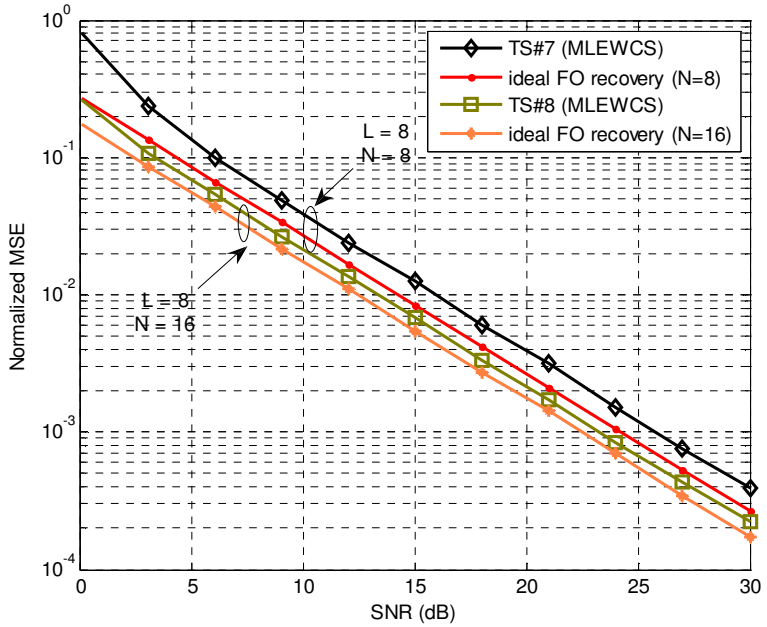


Figure 4-16: Normalized MSE for channel estimation with TS#7 and TS#8 (MLEWCS) (eight taps channel)

4.3 Comparison of the Performances of Joint Frequency Offset and Channel Estimation Methods

4.3.1 Comparison of the Performances of MLE#1 and MLEWCS

A number of computer simulations have been performed in the aim of performance comparison of MLE#1 and ML estimator with complementary training sequences (MLEWCS). In the simulations, the training sequences TS#1, TS#2, TS#3 and TS#5 given in Table 4-1 have been used with MLE#1 while the TS#7 in Table 4-3 and complementary pair (TS#9) in Table 4-4 have been chosen for MLEWCS.

Table 4-4: Complementary pairs used with MLEWCS

Training Sequence	Sequence-1	Sequence-2
TS#9	0001	1101

4.3.1.1 Simulation-7

Simulation-7 has been carried out to compare the performance of MLE#1 and MLEWCS by means of TS#3 and TS#7. In the simulation, the MLE#1 method has been realized with TS#3 while the MLEWCS has been carried out with TS#7. As a communication medium an eight taps channel impulse response sequence has been used. The frequency offset values to be estimated has been determined from the range of [-0.48, 0.48].

An MSE graph comparing the frequency offset estimation performances of MLE#1 and MLEWCS is illustrated in Figure 4-17. From the inspection of the figure, it is obviously seen that the MLEWCS exhibits a better performance than MLE#1 with

spending of nearly the same transmitted signal power. We can explain such a good performance with the following remarks:

1. The perfectly diagonal complementary autocorrelation characteristics of TS#7 reduces the cost function $g(v)$ to the squared magnitude of a vector by eliminating the inverse of the autocorrelation matrix $\boldsymbol{X}^H\boldsymbol{X}$. This elimination puts the $g(v)$ to a function form with sharp peaks making easy to determine the maximum point. This situation is illustrated in Figure 4-19. In the figure, the cost functions obtained with MLE#1 and MLEWCS for the same frequency offset value are given.
2. The usage of more information extracted from the extra received signal samples belong to the idle state of the channel in the application of MLEWCS provides a smaller estimation error.

The channel estimation performances of MLE#1 and MLEWCS are compared in Figure 4-18. From analyze of the figure it can be said that MLEWCS also shows a better performance in channel estimation. This is the result of perfectly diagonal autocorrelation matrix structure of TS#7.

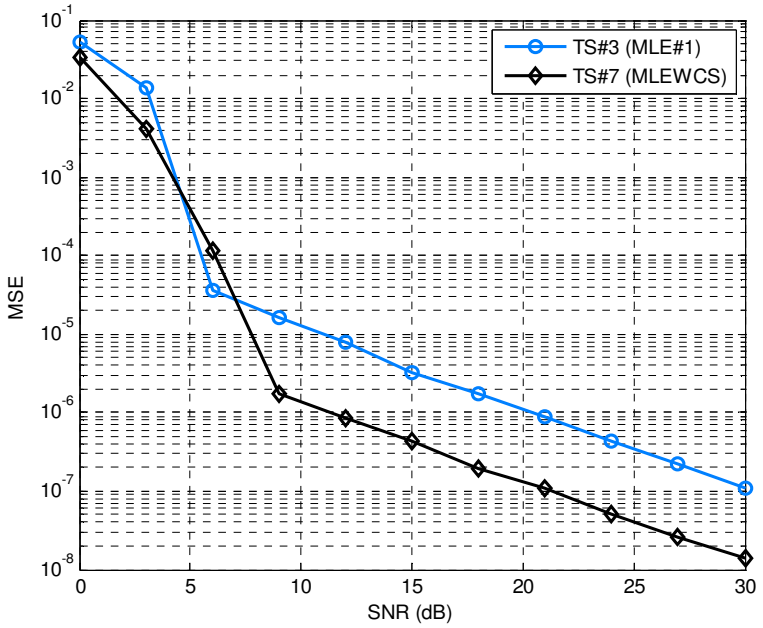


Figure 4-17: MSE for frequency offset estimation with TS#3 (MLE#1) and TS#7 (MLEWCS) (eight taps channel)

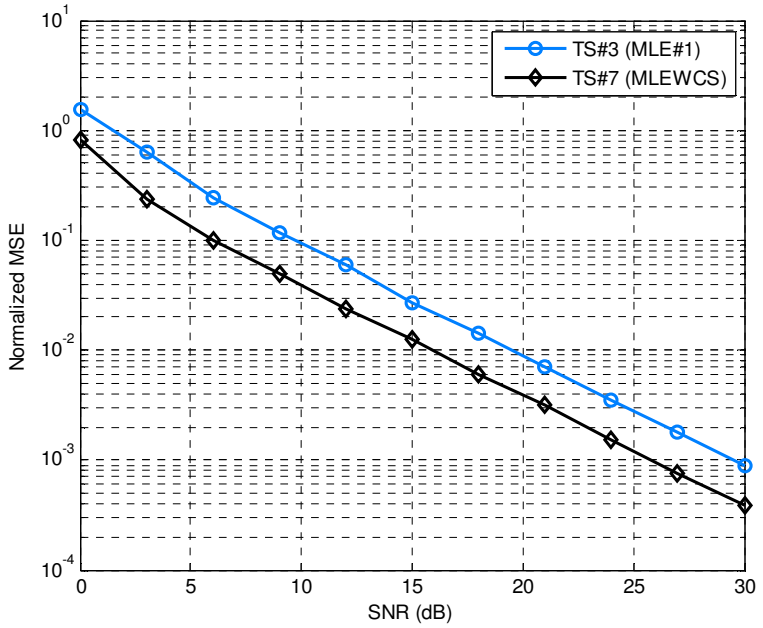
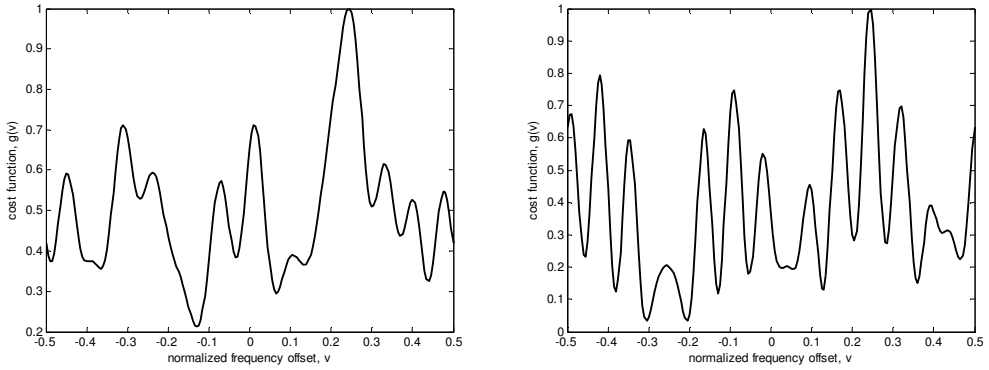


Figure 4-18: MSE for channel estimation with TS#3 (MLE#1) and TS#7 (MLEWCS) (eight taps channel)



(a) (b)

Figure 4-19: Typical cost functions: (a) with MLE#1, (b) with MLEWCS

Although we have tried to use same length of training sequences in the simulation, we still need to compare the performances of MLE#1 and MLEWCS when the transmitted signal energy changes. In Figure 4-20 and Figure 4-21, two plots comparing the frequency offset and channel estimation performances of MLE#1 and MLEWCS, with the change of transmitted signal energy, are given. In the plots, the x-axis show the ratio of transmitted signal energy to the average noise power in the channel. The values of this ratio has been chosen as the receive SNR remains in the range of 0 to $30\,dB$. From the examining of the figures, it can be said that MLE#1 with TS#3 is superior to MLEWCS with TS#7 in the case of the transmission of the same amount of training sequence energy.

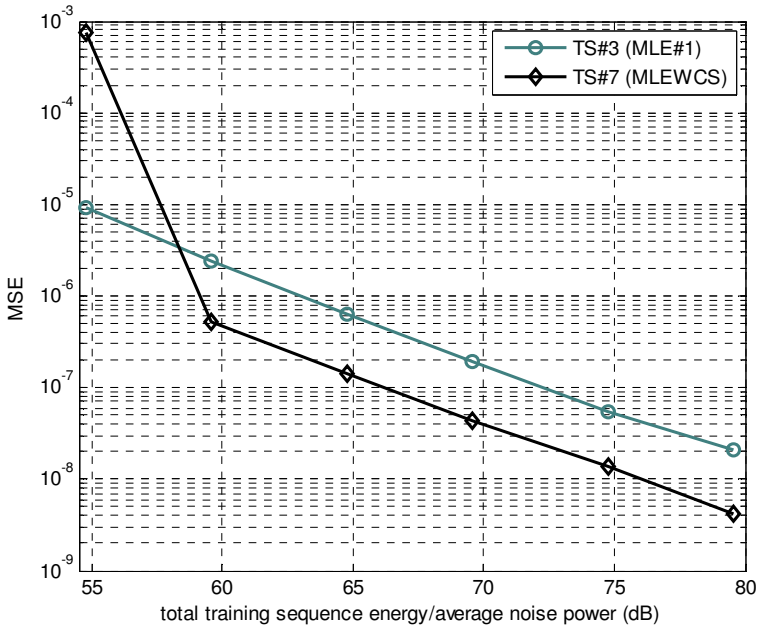


Figure 4-20: Comparison of the frequency offset estimation performances of TS#3 (MLE#1) and TS#7 (MLEWCS) in the change of total training sequence energy (eight taps channel)

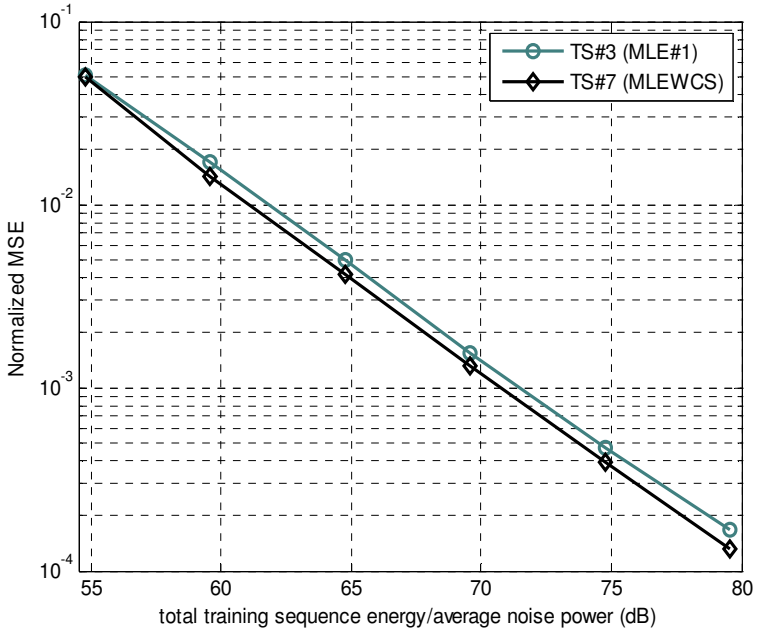


Figure 4-21: Comparison of the channel estimation performances of TS#3 (MLE#1) and TS#7 (MLEWCS) in the change of total training sequence energy (eight taps channel)

The difference of the values of receive SNR and x-axis of Figure 4-20 and Figure 4-21 (total training sequences energy/average noise power) comes from the values of the channel impulse response sequences applied in the simulations. The histogram of the autocorrelation of channel impulse response is shown in Figure 4-22.

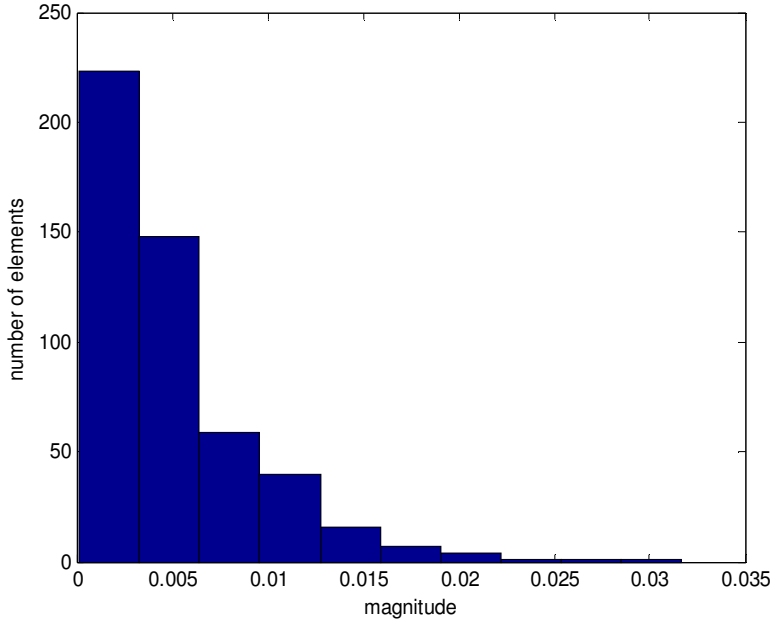


Figure 4-22: Histogram of the autocorrelation of the channel impulse response applied in the simulations

4.3.1.2 Simulation-8

Simulation-8 has been performed with TS#1, TS#2 in Table 4-1 and TS#9 in Table 4-4. The first two bits of TS#1 and TS#2 have been transmitted as non-zero and zero amplitude precursors respectively. As communication medium a three taps channel impulse response sequence has been assumed. The true frequency offset values has been chosen from the range of [-0.48, 0.48].

The frequency offset performances of MLE#1 with TS#1 and TS#2 and MLEWCS with TS#9 is shown in Figure 4-23. As it can be seen from the figure MLEWCS outperforms the MLE#1 for intermediate and high SNR levels due to the same remarks as those made for Simulation-7.

The normalized MSE curves for channel estimation with MLE#1 and MLEWCS are plotted in Figure 4-24. From the figure it can be said that MLEWCS has a better performance than MLE#1 for all SNR levels. Although both of the TS#1 and TS#9

has a perfectly diagonal autocorrelation characteristics, the better performance of MLEWCS can be explained again with the usage of more channel information hidden in the extra received signal samples belong to the idle state of the channel.

Two plots comparing the frequency offset and channel estimation performances of MLE#1 (with TS#1 and TS#2) and MLEWCS (with TS#9) in the change of total training sequence energy are given in Figure 4-25 and Figure 4-26 respectively. From the figure, it is seen that MLEWCS with TS#9 exhibits a better estimation performance than MLE#1 with TS#1 and TS#2 for spending the same amount training sequence energy.

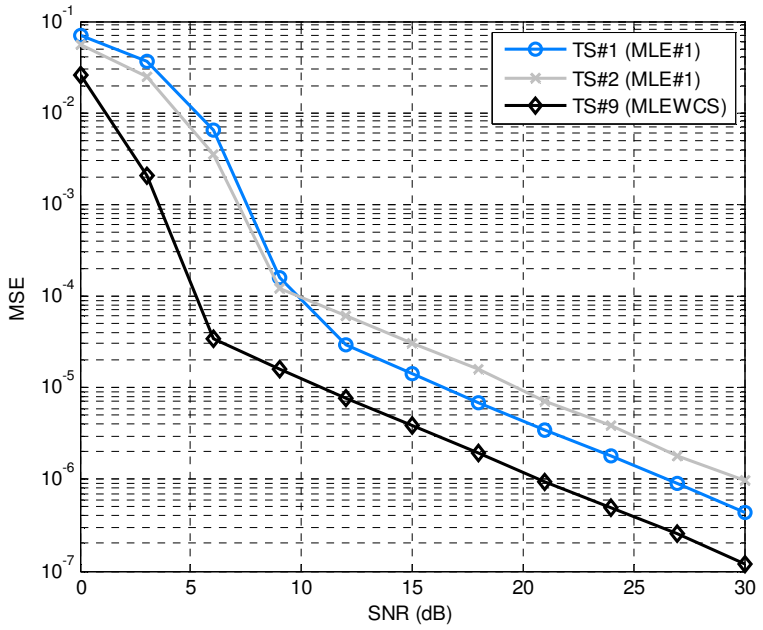


Figure 4-23: MSE for frequency offset estimation with TS#1, TS#2 (MLE#1) and TS#9(MLEWCS) (three taps channel)

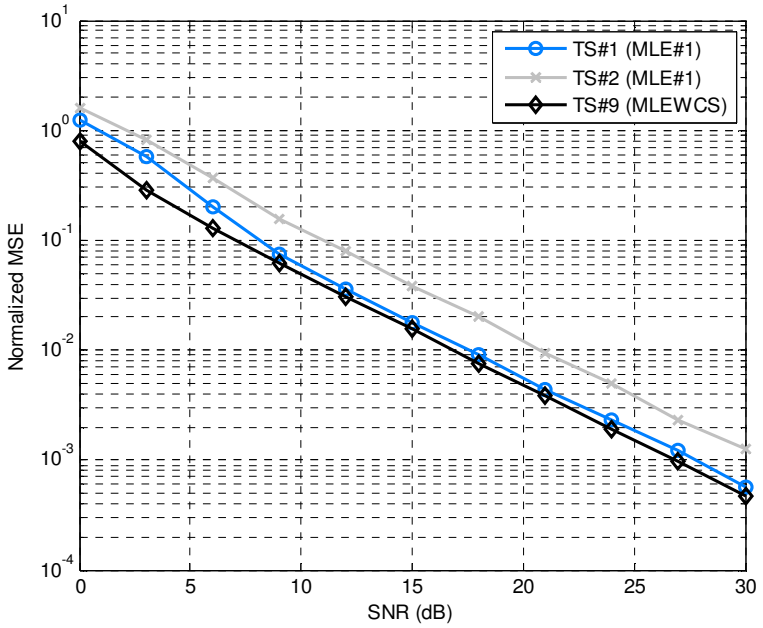


Figure 4-24: MSE for channel estimation with TS#1, TS#2 (MLE#1) and TS#9 (MLEWCS) (three taps channel)

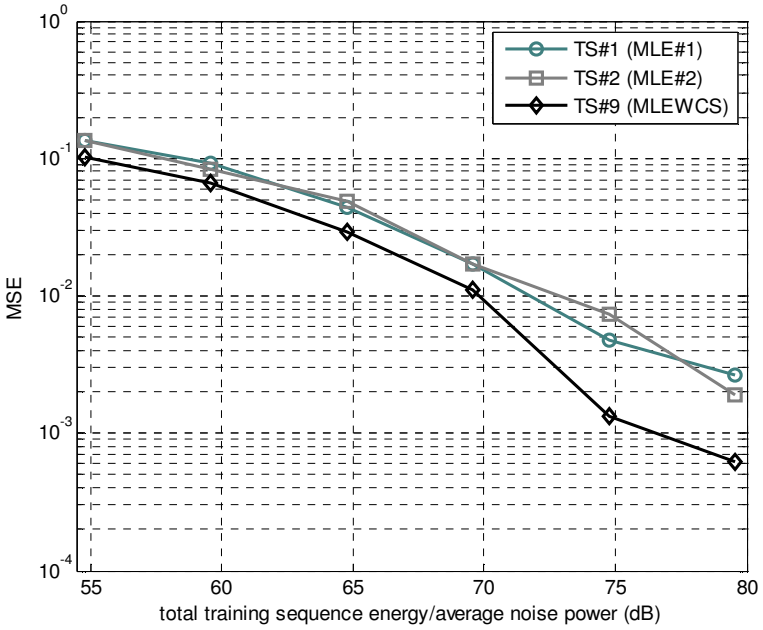


Figure 4-25: Comparison of the frequency offset estimation performances of TS#1, TS#2 (MLE#1) and TS#9 (MLEWCS) in the change of total training sequence enrgy (three taps channel)

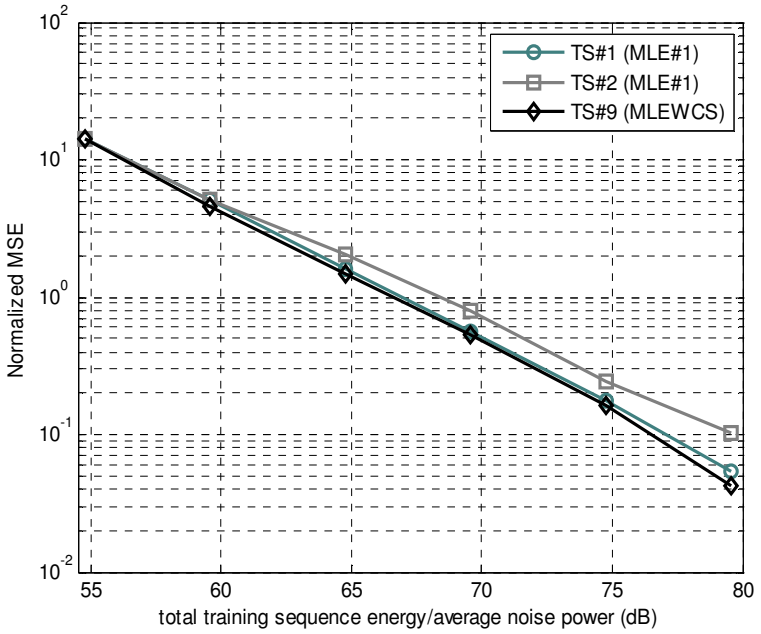


Figure 4-26: Comparison of the channel estimation performances of TS#1, TS#2 (MLE#1) and TS#9 (MLEWCS) in the change of total training sequence energy (three taps channel)

4.3.1.3 Simulation-9

In Simulation-9, TS#5 in Table 4-4 and TS#9 in Table 4-5 has been used as training purposes. As communication medium a three taps channel model designed in accordance with the model given in Section 4.1 has been applied. The frequency offset values to be estimated have been selected from the range of [0.48, -0.48].

The MSE curves for frequency offset estimation with TS#5 (MLE#1) and TS#9 (MLEWCS) are given in Figure 4-27.Analyzing the figure, it can be seen that MLEWCS shows a better estimation performance than MLE#1 with the same reason as in Simulation-7 and Simulation-9.

The normalized MSE curves for channel estimation with TS#5 and TS#9 are shown in Figure 4-28. As it can be observed from the figure, MLEWCS outperforms the MLE#1 in the channel estimation although both of the TS#5 and TS#9 have

perfectly diagonal autocorrelation matrix structures. The reason for such a better performance is the usage of channel information coming with the extra received signal samples due to the idle state of the channel.

In Figure 4-29 and Figure 4-30, two plots comparing the frequency offset and channel estimation performances of MLE#1 (with TS#5) and MLEWCS (with TS#9) in the change of total training sequence energy are given. Inspecting the figures, it is seen that MLE#1 with TS#5 outperforms the MLEWCS with TS#9 in both of the frequency offset and channel estimation by spending the same amount of training sequence energy.

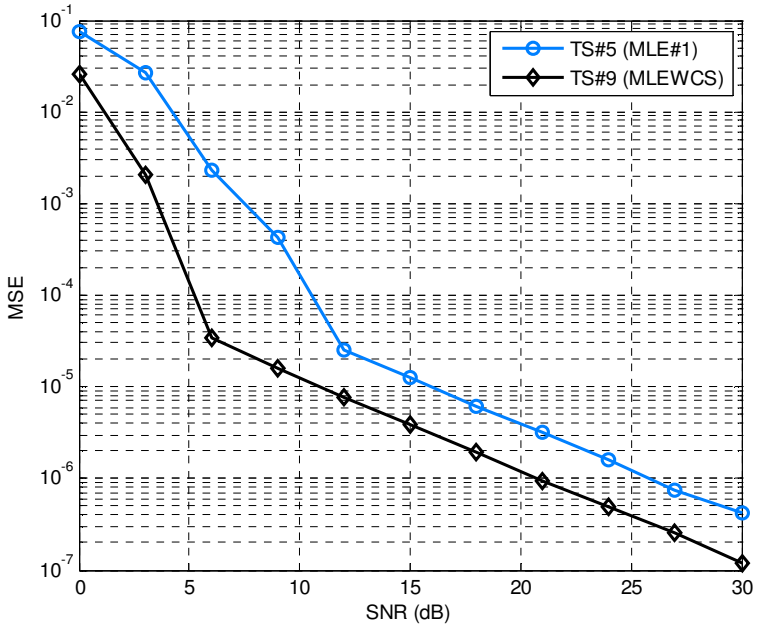


Figure 4-27: MSE for frequency offset estimation with TS#5 (MLE#1) and TS#9 (MLEWCS) (three taps channel)

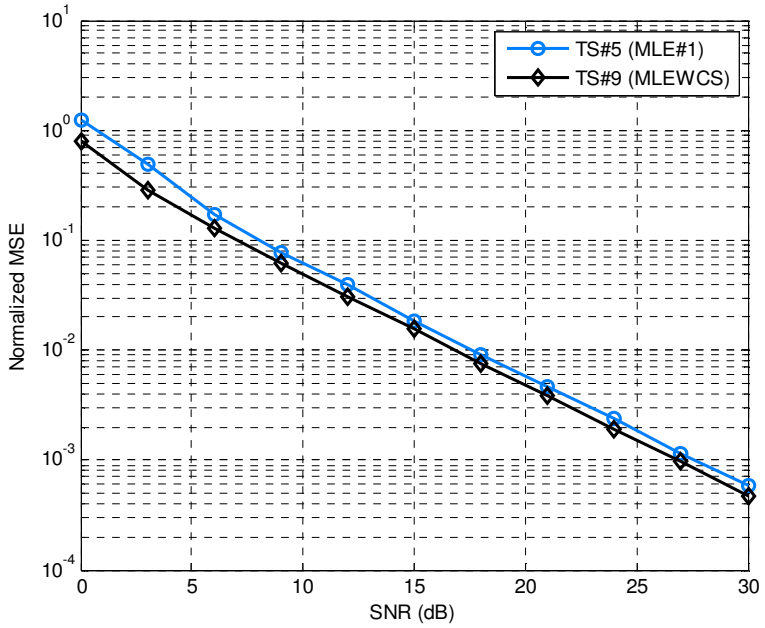


Figure 4-28: Normalized MSE for channel estimation with TS#5 (MLE#1) and TS#9 (MLEWCS) (three taps channel)

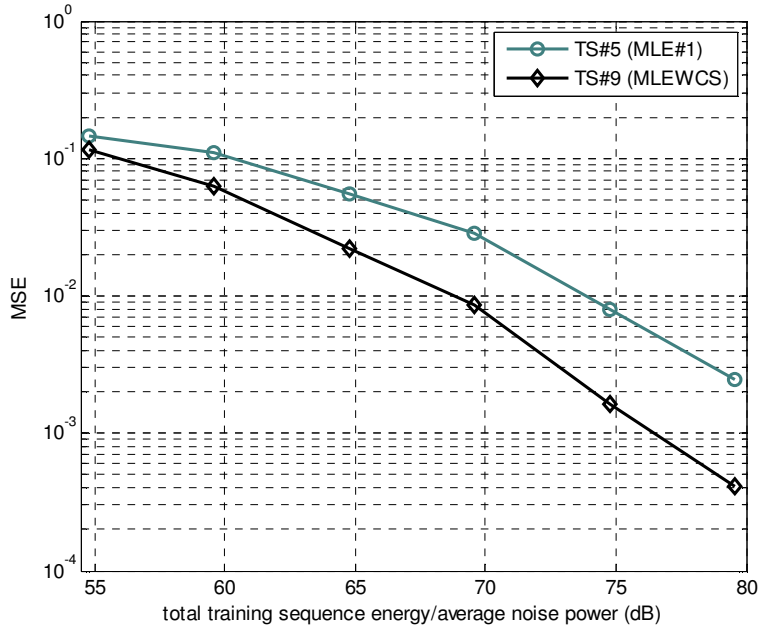


Figure 4-29: Comparison of the frequency offset estimation performances of TS#5 (MLE#1) and TS#9 (MLEWCS) in the change of total training sequence energy (three taps channel)

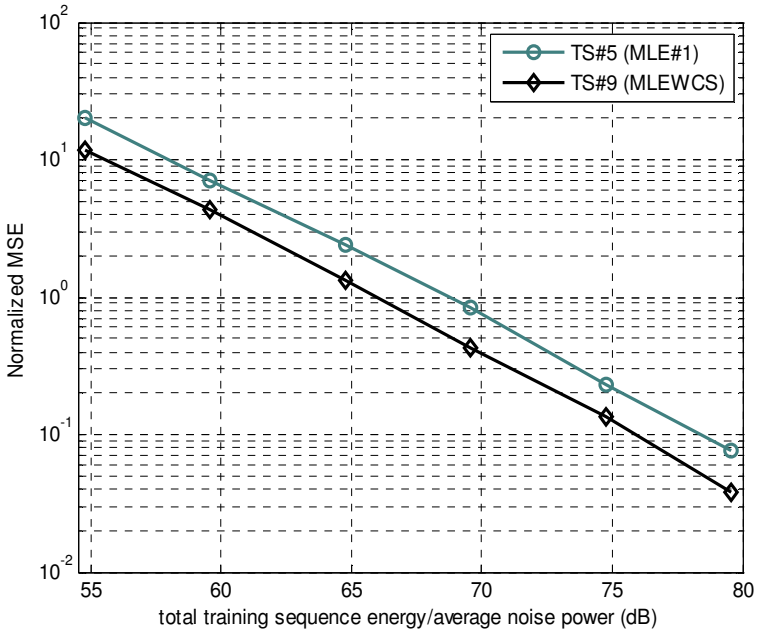


Figure 4-30: Comparison of the channel estimation performances of TS#5 (MLE#1) and TS#9 (MLEWCS) in the change of total training sequence energy (three taps channel)

4.3.2 Comparison of the Performances of MLE#2 and MLEWCS

In order to compare the estimation performances of MLE#2 and MLEWCS Simulation-10 has been realized. As training sequences TS#6 in Table 4-2 and TS#8 in Table 4-3 have been used.

4.3.2.1 Simulation-10

In Simulation-10, TS#6 has been applied with MLE#2 while TS#8 has been used with MLEWCS. As in Simulation-4 the last seven bits of TS#6 have been transmitted as precursors. As communication channel an eight taps channel impulse response sequences has been used. The true frequency offset values have been

chosen from the range of [-0.05, 0.05] due to the narrow estimation range of MLE#2.

In Figure 4-31, the MSE curves for the frequency offset estimation with MLE#2 and MLEWCS are illustrated. Form the inspection of the figure, it is obviously seen that MLEWCS is superior to MLE#2 in frequency offset estimation nearly for all levels of SNR. Although providing such a better performance, MLEWCS reaches the MSE curve in the figure with a larger computational complexity using nearly the same signal energy as in MLE#2 (due to the transmission of nearly the same number of training bits). This can be accepted as a trade off between MLE#2 and MLEWCS.

In Figure 4-32, normalized MSE curves for channel estimation with MLE#2 and MLEWCS are plotted. Examining the figure, it can be said that the MLEWCS outperforms the MLE#2 in channel estimation.

The frequency offset and channel estimation performances of MLE#2 and MLEWCS in the change of training sequence energy are compared in Figure 4-33 and Figure 4-34. From the analyze of the figures, it can be said that MLEWCS shows a better performance in frequency offset and channel estimation with spending the same amount training sequence energy.

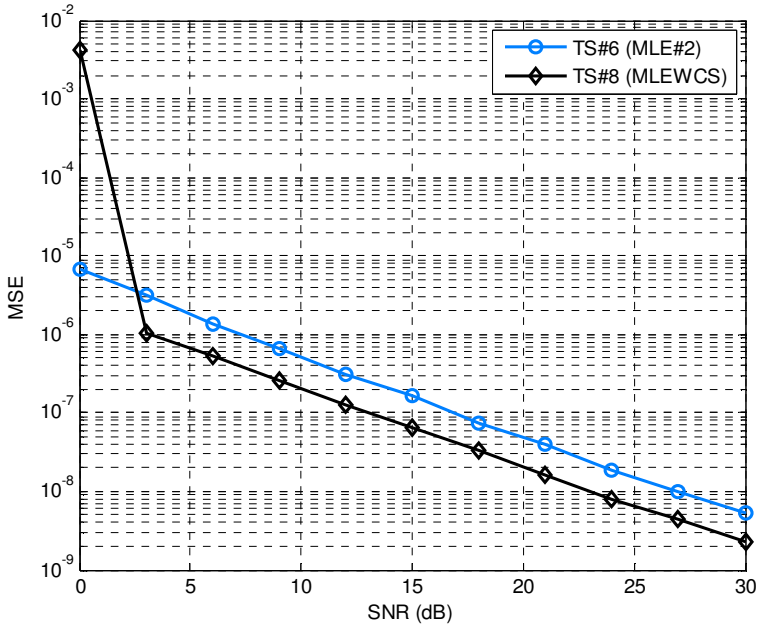


Figure 4-31: MSE for frequency offset estimation with TS#6 (MLE#2) and TS#8 (MLEWCS) (eight taps channel)

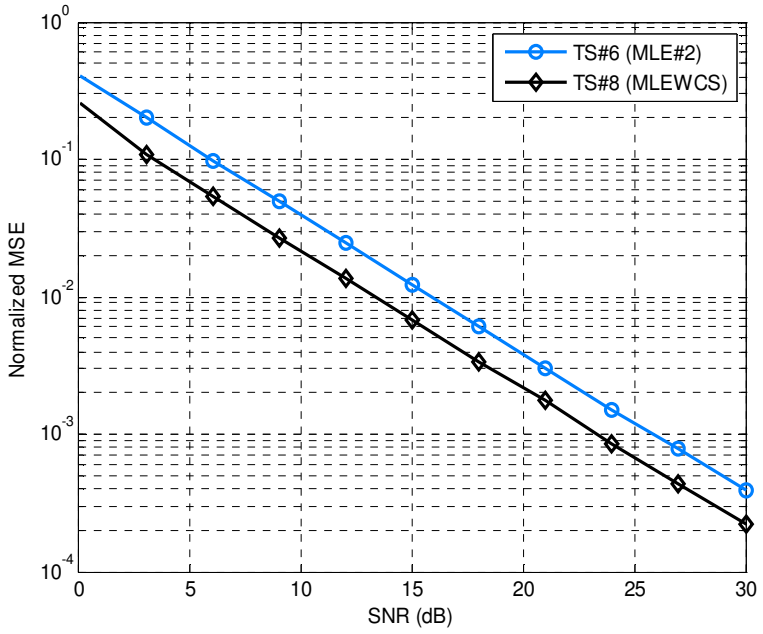


Figure 4-32: Normalized MSE for channel estimation with TS#6 (MLE#2) and TS#8 (MLEWCS) (eight taps channel)

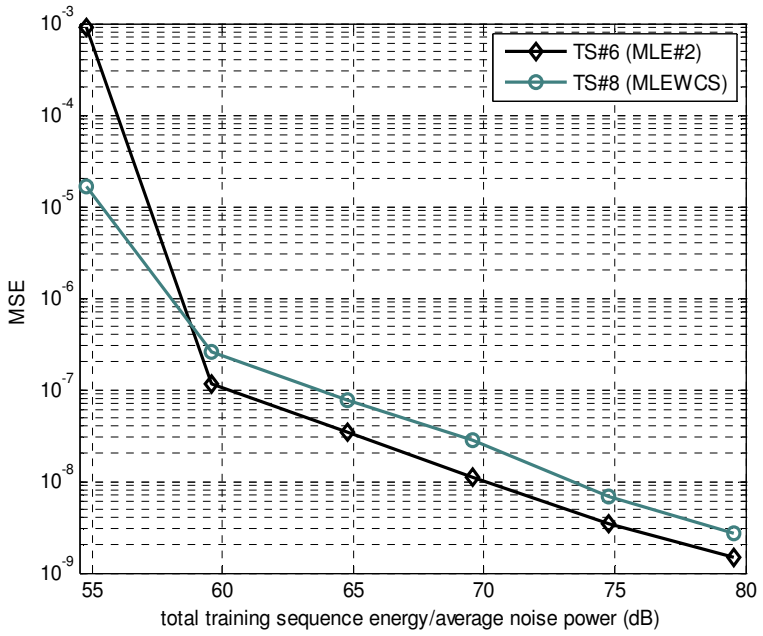


Figure 4-33: Comparison of the frequency offset estimation performances of TS#6 (MLE#2) and TS#8 (MLEWCS) in the change of total training sequence energy (eight taps channel)

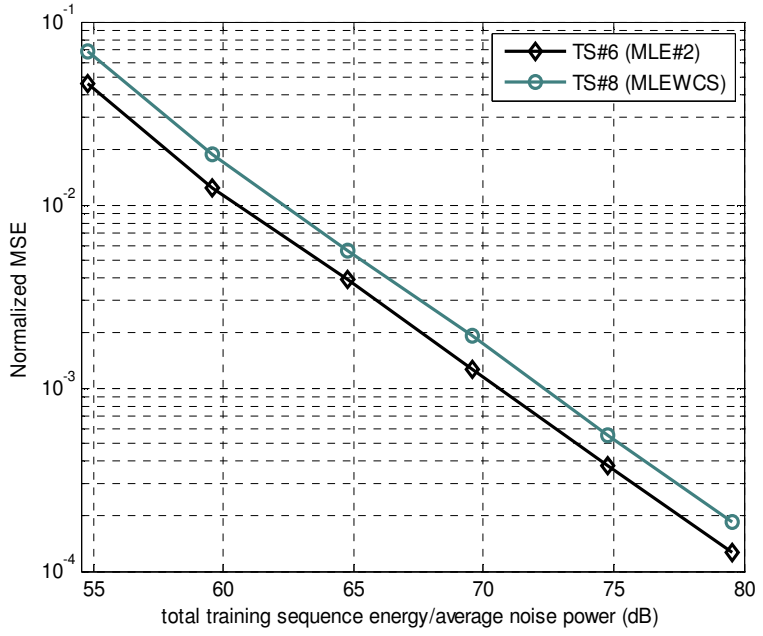


Figure 4-34: Comparison of the channel estimation performances of TS#6 (MLE#2) and TS#8 (MLEWCS) in the change of total training sequence energy (eight taps channel)

4.3.3 Comparison of the Performances of AHE and MLEWCS

For comparison of performances of AHE and MLEWCS in frequency offset and channel estimation Simulation-11 has been run.

4.3.3.1 Simulation-11

In Simulation-11, TS#6 in Table 4-2 and TS#8 in Table 4-3 have been used with AHE and MLEWCS respectively. As communication channel an eight gains channel model has been applied. The frequency offset values have been sleceted from the range of [-0.05, 0.05] due to the narrow estimation range of AHE.

A graph comparing the frequency offset estimation MSE curves of AHE and MLEWCS is given in Figure 4-35. Analyzing the figure it is observed that MLEWCS outperforms the AHE nearly for all SNR levels. As in Simulation-10, the MLEWCS reaches this better MSE with a larger computational complexity compared with AHE. This can be thought as a trade off between the use of AHE and MLEWCS.

The normalized MSE curves for channel estimation with AHE and MLEWCS are displayed in Figure 4-36. As in Simulation-5, the channel estimation with AHE has been fulfilled substituting the estimated value of frequency offset provided by AHE into Eq. (3-23). From the inspection of the figure it can be said that, MLEWCS exhibits a better performance than AHE.

In Figure 4-37 and Figure 4-38, the frequency offset and channel estimation performances of AHE and MLEWCS are compared for the change of total training sequence energy used. From the figures, it is obviously seen that MLEWCS outperforms the AHE in both of frequency offset and channel estimation with spending the same amount training sequence energy.

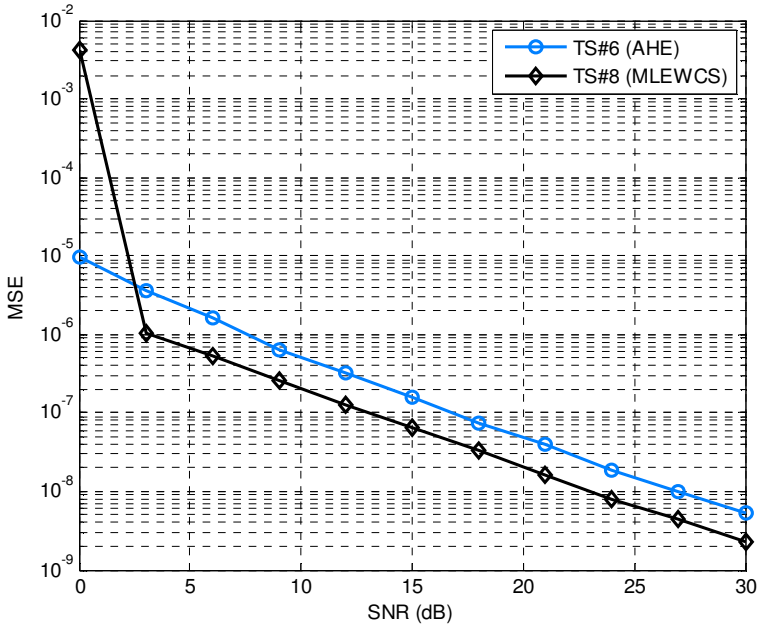


Figure 4-35: MSE for frequency offset estimation with TS#6 (AHE) and TS#8 (MLEWCS) (eight taps channel)

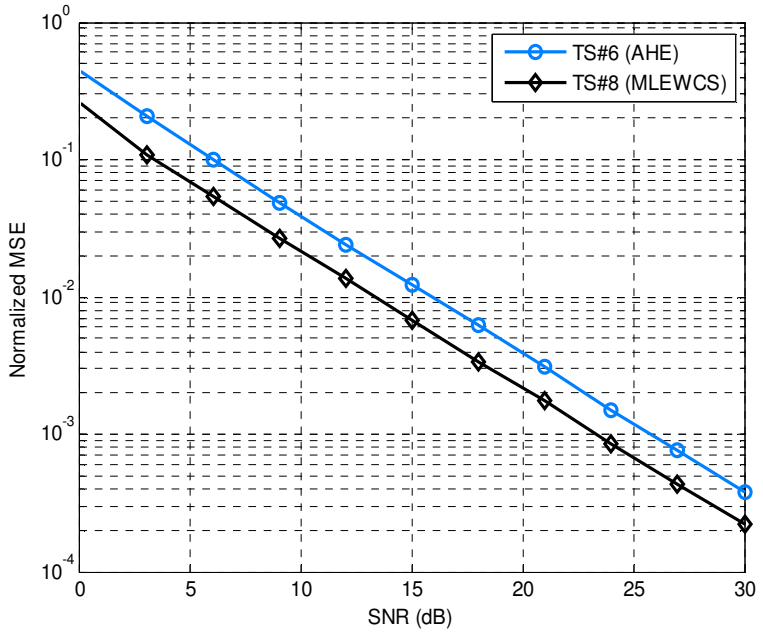


Figure 4-36: Normalized MSE for channel estimation with TS#6 (AHE) and TS#8 (MLEWCS) (eight taps channel)

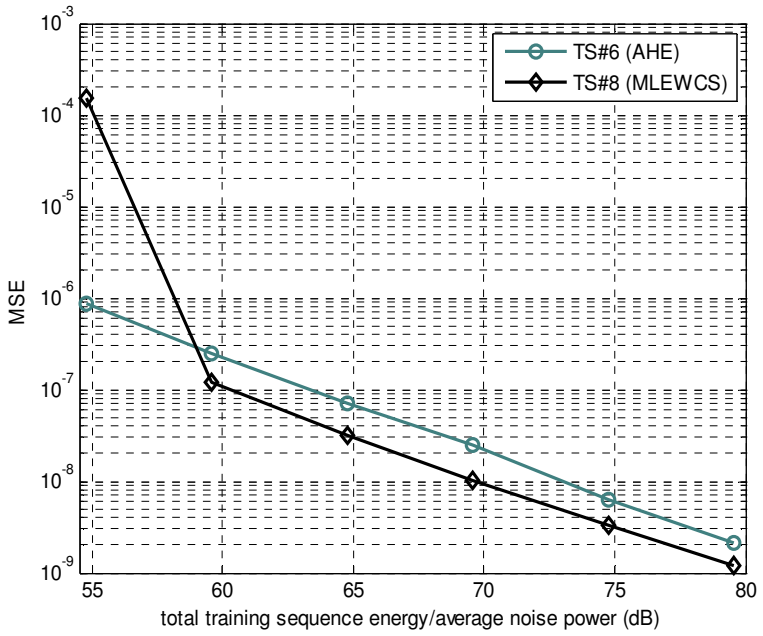


Figure 4-37: Comparison of the frequency offset estimation performances of TS#6 (AHE) and TS#8 (MLEWCS) in the change of total training sequence energy (eight taps channel)

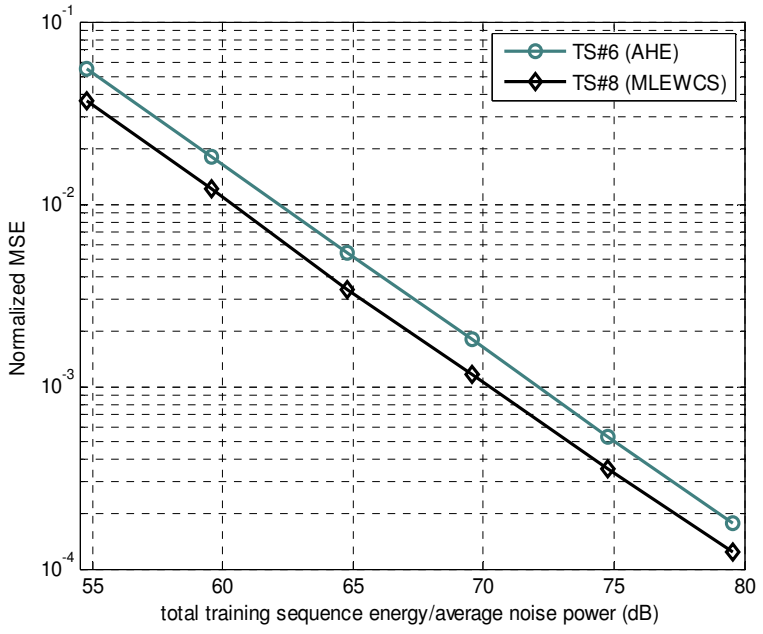


Figure 4-38: Comparison of the channel estimation performances of TS#6 (AHE) and TS#8 (MLEWCS) in the change of total training sequence energy (eight taps channel)

4.4 Performances of MLEWCS in MIMO Communication Systems

In order to observe the performance of MLEWCS in MIMO communication systems Simulation-12 has been carried out. In the simulation, TS#8 given in Table 4-3 has been used as training purposes.

4.4.1 Simulation-12

In order to realize the Simulation-12, a two transmit and one receive antenna communication system has been modeled. As communication medium two eight taps channel models have been used. Communication channels have been modeled according to the model given in Section 4.1. The frequency offset values to be estimated have been selected from the range of [-0.48,0.48].

MSE curve for frequency offset estimation with MLEWCS is given in Figure 4-39. For comparison purposes minimum CRB achieved in 1000 simulation runs is also illustrated in the same figure. As it is expected MSE curve follows the CRB for intermediate and high SNR levels and exhibits a similar performance to the applications in SISO communication systems.

Normalized MSE curves for channel estimation with MLEWCS are given in Figure 4-40 and Figure 4-41. In the figures ideal normalized MSEs in the case of ideal frequency offset recovery are also illustrated. As it is seen in the figures, normalized MSE curves follow their ideal curves in all SNR levels.

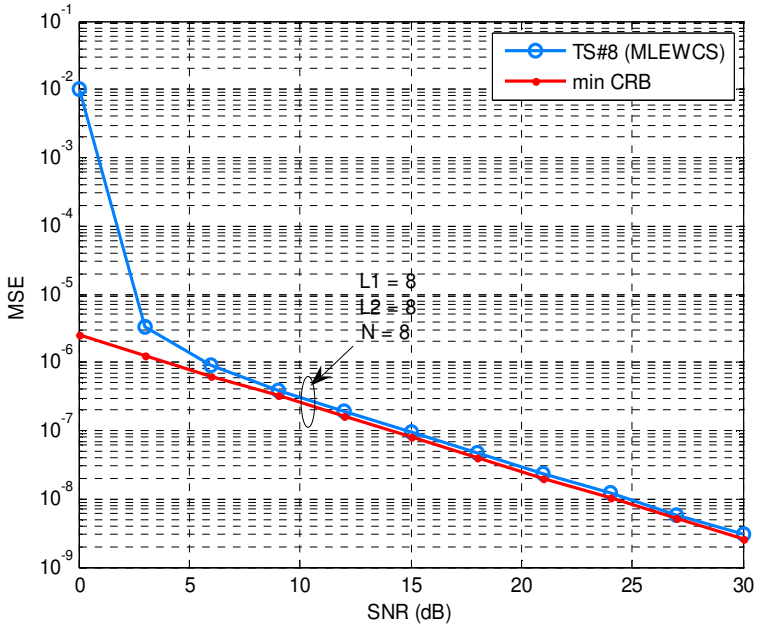


Figure 4-39: MSE for frequency offset estimation with TS#8 (MLEWCS) (eight taps channel)

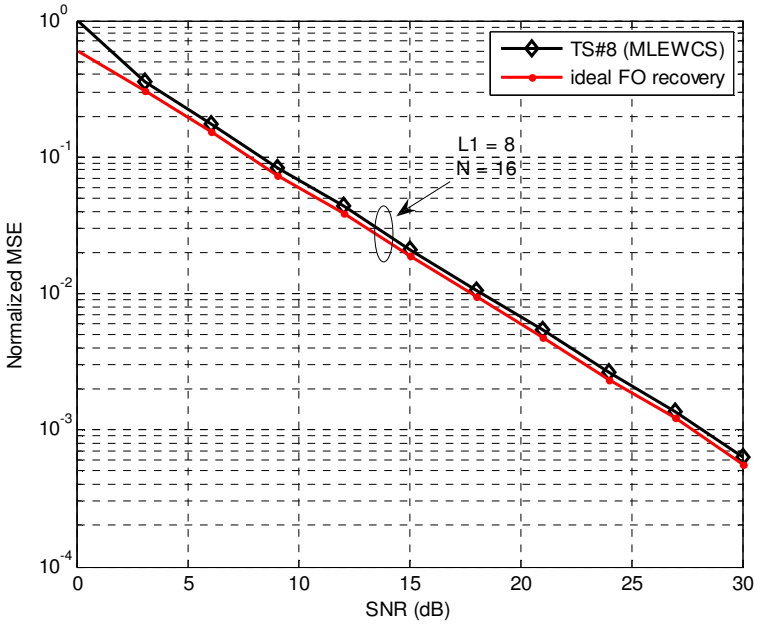


Figure 4-40: Normalized MSE for frequency offset estimation with TS#8 (MLEWCS) (MIMO systems, Channel#1) (eight taps channel)

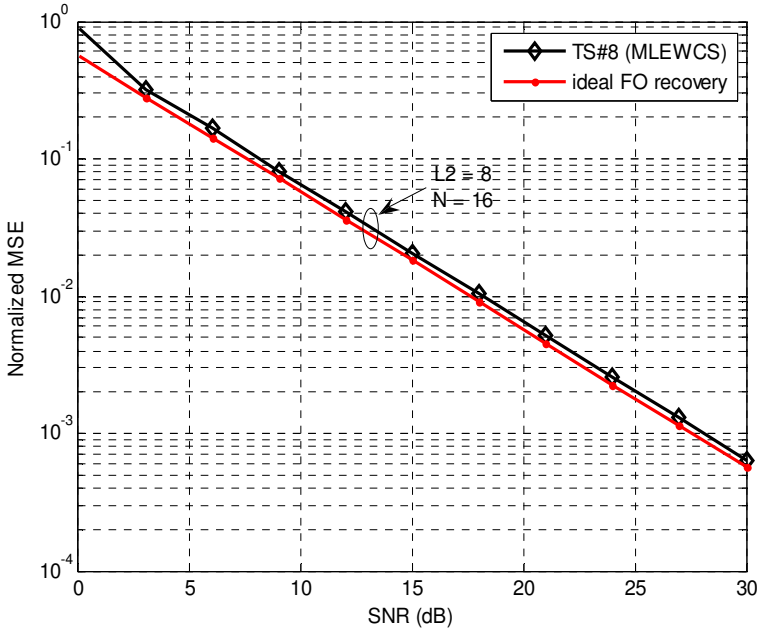


Figure 4-41: Normalized MSE for frequency offset estimation with TS#8 (MLEWCS) (MIMO systems, Channel#2) (eight taps channel)

CHAPTER 5

CONCLUSION

In this thesis methods of joint frequency offset and channel estimation have been examined. The classical problem of conventional channel estimation forms a special of the problem examined. A review of channel estimation methods has also been given.

Joint channel and frequency estimation methods are in general composed of two stages. In the first stage, the frequency offset is estimated. In the second stage, the received data is demodulated with the estimated frequency offset to remove the effects of carrier frequency mismatch and then channel estimation methods are applied.

In the literature, there are maximum likelihood (ML) based methods treating both channel response and frequency offset as unknown deterministic quantities. These methods involve a grid search to find the ML estimate of frequency offset. In ML estimation is non-linear in the frequency offset variable; determination of the optimal point is significantly difficult. As noted in this thesis, the cost functions has many local maximas and minimas therefore a dense grid search is required to reliably estimate the frequency offset. Once the frequency offset is calculated, the channel response can be estimated easily using least square solution for linear set of equations which the ML estimate of channel response.

In addition to the ML based methods, there are ad-hoc methods which are designed for joint channel and frequency estimation. The ad-hoc technique described in [5]

requires a simple division of consecutive matched-filter peaks for the frequency offset estimation. It has been shown that this method is very effective at high SNR.

In this thesis we compare the methods described using different training sequences. Some of the training sequences used require non-zero amplitude precursors, while some others do not have such a requirement. We present a comparison of training sequences in terms of

a) Mean square estimation error versus receiver SNR

b) Mean square estimation error versus total energy allocated for training

In addition to this comparison, we have also used complementary sequences as the training sequences. The complementary sequences have been used for training in the literature. However, to the best of our knowledge they have not been used in joint frequency and channel response estimation. According to our numerical simulations, the complementary sequences resulted in the least mean square estimation error in both receiver SNR and allocated energy sense.

The thesis also includes some extensions such as the formulation of multiple frequency offset estimation and an extension to multi-input multi-output joint channel and frequency estimation. Both of these extensions can be valuable future research direction. Furthermore, we have not compared the delay in estimation of channel and frequency offset parameters. A future research direction can be the comparisons of the schemes with respect to their delay. Another research direction is to compare various periodic training sequences of the same total length (N), but with different periods (P).

REFERENCES

[1] *S.N. Crozier, D.D. Falconer, S.A. Mahmoud, "Least sum of squared errors (LSSE) channel estimation" IEE Proceedings, vol 138, no. 2, pp. 371-378, Aug. 1991*

[2] *C. Tellambura, M. G. Parker, Y. Jay Guo, "Optimal sequences for channel estimation using discrete Fourier transform techniques," IEEE Transactions on Communications, vol. 47, no. 2, pp. 230-238, Feb. 1999*

[3] *W. Chen, U. Mitra, "Training sequence optimization: comparisons and an alternative criterion," IEEE Transactions on Communications, vol. 48, no. 12, pp. 1987-1991, Dec. 2000*

[4] *O. Rousseaux, G. Leus, P. Stoica, M. Moonen, "Gaussian maximum-likelihood channel estimation with short training sequences," IEEE Transactions on Wireless Communications, vol. 4, no. 6, pp. 2945-2955, Nov. 2005*

[5] *M. Morelli, U. Mengali, "Carrier-frequency estimation for transmissions over selective channels," IEEE Transactions on Communications, vol. 48, no. 9, pp. 1580-1589, Sep. 2000*

[6] *P. Spasojevic, C. N. Georghiades, "Complementary sequences for ISI channel estimation," IEEE Transactions on Information Theory, vol. 47, no. 3, pp. 1145-1152, March 2001*

[7] *S. Wang, A. Abdi, "Aperiodic complementary sets of sequences-based MIMO frequency selective channel estimation," vol. 9, no. 10, pp. 891-893, Apr. 2005*

[8] *T. Cui, C. Tellambura, "Joint channel and frequency offset estimation and training sequence design for MIMO systems over frequency selective channels," IEEE Communication Society Globecom 2004, pp. 2344-2348*

[9] *S. Ahmed, S. Lambotharan, A. Jakobson, J. A. Chambers, "MIMO frequency-selective channels with multiple frequency offsets: estimation and detection techniques," IEE Proc.-Common, vol. 152, no. 4, pp. 489-493, Aug. 2005*

[10] *S. Ahmed, S. Lambotharan, A. jakobson, A. Chambers, "Parameter estimation and equalization techniques for communication channels with multipath and multiple frequency offsets," IEEE Transactions on Communications, vol. 53, no. 2, pp. 219-223, Feb. 2005*

[11] *M. J. E. Golay, "Multislit spectroscopy," J. Opt. Soc. Amer.,vol. 39, pp. 437-444, 1949*

[12] *M. J. E. Golay, "Complementary Series," IEEE Transactions on Information Theory, vol. IT-7, pp. 82-87, Apr. 1961*

[13] *M. G. Parker, K. G. Peterson, C. Tellambura, "Golay Complementary Sequences," , Jan. 2004*

[14] *C. Tellambura, Y. J. Guo, S. K. Barton, "Channel estimation using aperiodic binary sequences," IEEE Communications Letters, vol. 2, no. 5, pp. 140-142*

[15] *E. Larsson, P. Stoica, G. Ganesan, "Space-Time Block Coding for Wireless Communications," Cambridge University Press, 2003*

[16] *H. Wishwanatan, R. Krishnamoorty, "A frequency offset estimation technique for frequency-selective fading channels," IEEE Communications Letters, vol. 5, no. 4, pp. 166-168, Apr. 2001*

[17] *S. M. Kay, "Fundamentals of Statistical Signal Processing Volume 1: Estimation Theory" Englewood Cliffs, Prentice-Hall, 1993*

[18] *N. Levanon, E. Mozeson, "Radar Signals," John Wiley, 2004*

[19] *J. G. Proakis, "Digital Communications," McGraw-Hill, 2001*

[20] *L. Tong, B. M. Sadler, M. Dong, "Pilot-assisted wireless transmissions: general model, design criteria, and signal processing," IEEE Signal Processing Magazine, vol. 21, pp. 12-25, Nov. 2004*

[21] *H. Liu, U. Tureli, "A high-efficiency carrier estimator for OFDM communications," IEEE Communications Letters, vol. 2, no. 4, pp. 104-106, Apr. 2004*

[22] *U. Tureli, H. Liu, M. D. Zoltowski, "OFDM blind carrier offset estimation: ESPRIT," IEEE Transactions on Communications, vol. 48, no. 9, pp. 1459-1461, Sep. 2000*

APPENDIX A

PROOF OF THE MINIMUM MSE FOR LSSE [1]

Let $\boldsymbol{R}$ be N-by-N positive definite Hermitian matrix and $\boldsymbol{R}_d$ be the diagonal matrix containing only the diagonal elements of $\boldsymbol{R}$. Our aim is to prove the inequality in Eq. (A-1).

$$tr\{\boldsymbol{R}^{-1}\} \geq tr\{\boldsymbol{R}_d^{-1}\} \tag{A-1}$$

Applying the Cholesky factorization, $\boldsymbol{R}$ can be defined as

$$\boldsymbol{R} = \boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^H \tag{A-2}$$

where $\boldsymbol{L}$ is a lower triangular matrix 1's in the diagonal and $\boldsymbol{D}$ is a positive definite diagonal matrix with elements d_i. The inverse of R can be defined as

$$\boldsymbol{R}^{-1} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^H \tag{A-3}$$

where $\boldsymbol{U}^H = \boldsymbol{L}^{-1}$. Here $\boldsymbol{L}^{-1}$ is a matrix which has the same matrix form as $\boldsymbol{L}$ and $\boldsymbol{U}$ is an upper diagonal matrix with 1's in diagonal. Let $l_{i,j}$'s are elements of $\boldsymbol{L}$, the ith diagonal element of $\boldsymbol{R}$ can be bounded as in Eq. (A-4).

$$a_{i,i} = d_i + \sum_{j=1}^{i-1} \left|l_{i,j}\right|^2 d_j \geq d_i \tag{A-4}$$

Using Eq. (A-4) and Eq. (A-3) we can write the inequality in Eq. (A-5).

$$a_{i,i} = \frac{1}{d_i} + \sum_{j=i+1}^{N} \left|u_{i,j}\right|^2 \frac{1}{d_j} \geq \frac{1}{d_i} \geq \frac{1}{a_{i,i}} \tag{A-4}$$

Hence, $tr\{\boldsymbol{R}^{-1}\}$ takes its minimum value when $\boldsymbol{R}$ is perfectly diagonal. This is the proof of the minimum MSE for LSSE channel estimation method.

APPENDIX B

TRAINING SEQUENCE DESIGN FOR LSSE CHANNEL ESTIMATION METHOD [1]

One of the measuring method to measure the quality of channel estimation is Signal to-Estimation Error Ratio (SER). The SER is given by

$$SER = \frac{|\boldsymbol{h}|^2}{\sigma_w^2 tr(\boldsymbol{R}_{xx}^{-1})} \tag{B-1}$$

where $\boldsymbol{h}$ is channel impulse response, σ_w^2 is variance of AWGN and $tr(\boldsymbol{R}_{xx}^{-1})$ is trace of inverse autocorrelation matrix of training sequence.

By relating the SER to Signal-to-Noise Ratio (SNR) defined the normalized SER is achieved.

$$SER' = \frac{SER}{SNR} \tag{B-2}$$

where SNR is defined by Eq. (B-3)

$$SNR = \frac{|\boldsymbol{x}[n]|^2 |\boldsymbol{h}|^2}{\sigma_w^2} \tag{B-3}$$

where $\boldsymbol{x}[n]$ is the vector defined by Eq. (B-4)

$$x[n] = [x[n]\, x[n-1] \ldots x[n-L+1]]^T \tag{B-4}$$

Inserting Eq. (B-3) into Eq. (B-2) the normalized SER is obtained as

$$SER' = (|\boldsymbol{x}[n]|^2 tr(\boldsymbol{R}_{xx}^{-1}))^{-1} \tag{B-5}$$

The normalized SER is an important measure to compare the estimation performance of different training sequences. As a general rule the performance of a training sequence is better if SER' corresponds to it is higher.

The upper bound for normalized SER can be found and used to find the training sequences for optimum channel estimation. As a characteristic of LSSE channel estimation method, two different types of training sequences can be designed.

- Non-zero amplitude precursor
- Zero amplitude precursor

B.1 Non-Zero Amplitude Precursor Case

For non-zero amplitude precursor preamble all elements have same amplitude. With a well chosen non-zero amplitude precursor the autocorrelation matrix $\boldsymbol{R}_{xx}$ is nearly diagonal. For this case diagonal elements are given by

$$r_{xx}[l,l] = M|\boldsymbol{x}[n]|^2 \quad l = 0, 1, \dots, L-1 \tag{B-6}$$

where M is the length of training sequence part of the preamble.

If autocorrelation matrix $\boldsymbol{R}_{xx}$ is perfectly diagonal then the diagonal elements of inverse autocorrelation matrix $\boldsymbol{R}_{xx}^{-1}$ is obtained as

$$r_{xx}^{-1}[l,l] = \frac{1}{r_{xx}[l,l]} = \frac{1}{M|\boldsymbol{x}[n]|^2} \quad l = 0, 1, \dots, L-1 \tag{B-7}$$

Inserting Eq. (B-7) into Eq. (B-5) the upper bound for normalized SER is found.

$$SER' \leq \frac{M}{L} \tag{B-8}$$

In study had been performed for [1], a computer search for finding the best preamble sequences had been run by using normalized SER as a quality measure. The best preamble sequences had been found in this search are given in Table B-1.

Table B-1: Best binary preamble sequences designed by using LSSE channel estimation including non-zero amplitude precursors

Channel Length, L	**Total Preamble Length, $N+L-1$**					
	$\mathbf{2L-1}$	$\mathbf{2L}$	$\mathbf{2L+1}$	$\mathbf{2L+2}$	$\mathbf{2L+3}$	$\mathbf{2L+4}$
2	1*	1	2*	2	5*	5
3	2	4*	4	9	12	24*
4	8*	8	17	27	58*	58
5	10	2E	4F	58	B1	162
6	5C	B0	162	162	45E	530
7	13A	48E	4F5	123B	650	14C1
8	91D	123B	2476	C2EF	2981	2981
9	476E	48ED	8BCE	8D04	5302	CAF9

In Table B-1, given preambles provides a BPSK signaling characteristics and convenient for any MPSK system. In table the preambles are given in hexadecimal format. Each preamble contains $L-1$ precursor and enough amount of zeros to get suitable preamble length. The preambles indicated by * have perfectly diagonal autocorrelation matrix.

B.2 Zero Amplitude Precursor Case

If chosen preamble is zero amplitude the diagonal elements of autocorrelation matrix $\boldsymbol{R}_{xx}$ are given by

$$r_{xx}[l,l] = (M-l)|\boldsymbol{x}[n]|^2 \quad l = 0, 1, \dots, L-1 \tag{B-9}$$

If autocorrelation matrix is perfectly diagonal, the diagonal elements of inverse autocorrelation matrix are obtained as

$$r_{xx}^{-1}[l,l] = \frac{1}{r_{xx}[l,l]} = \frac{1}{(M-l)|\boldsymbol{x}[n]|^2} \quad l = 0, 1, \dots, L-1 \tag{B-10}$$

Inserting Eq. (B-10) into Eq. (B-5), the upper bound for normalized SER is found and expressed as

$$SER' \leq \left(\sum_{l}^{L-1} \frac{1}{M-l} \right)^{-1} \tag{B-11}$$

The estimation performance of preambles including zero amplitude precursor is worse than the performance of preambles in non-zero amplitude precursor case. The performance of zero amplitude precursor preambles approaches to performance of non-zero amplitude precursor preambles when the length of training sequence in preamble is much bigger than the length of precursor. The reason for that is transmitting the less energy in zero amplitude precursor case.

In Table B-2, the preambles had been found in [1] are given. The preambles in Table B-2 include zero amplitude precursors. The preambles have BPSK signaling characteristics and shown in hexadecimal format. The preambles include enough "-1" to achieve correct training sequence length, M, and "0" to achieve correct preamble length, $M + L - 1$.

Table B-2: Best binary preamble sequences designed by using LSSE channel estimation including zero amplitude precursors

Channel Length, L	**Total Preamble Length, $N + L - 1$**					
	$\mathbf{2L - 1}$	$\mathbf{2L}$	$\mathbf{2L + 1}$	$\mathbf{2L + 2}$	$\mathbf{2L + 3}$	$\mathbf{2L + 4}$
2	0	1*	1	2*	2	5*
3	0	1	2	4	8	9
4	0	1	8	11	11	58
5	0	1	1D	3B	68	C2
6	0	1	1D	3B	68	13D
7	0	1	48	91	13D	13D
8	0	1	E2	16E	397	6C2
9	0	1	E2	1C4	4C2	4C2

In Table B-2, the preambles indicated by * are preambles which have perfectly diagonal autocorrelation matrix.

APPENDIX C

ASYMPTOTIC PROPERTY OF THE MLE [17, p. 167]

Let $p(\boldsymbol{h}; v)$ is probability density function (PDF) of the parameter $\boldsymbol{h}$. If $p(\boldsymbol{h}; v)$ satisfies some "regularity" conditions, then MLE of the parameter v is asymptotically distributed in accordance with distribution in Eq. (C-1).

$$\mathcal{N}(0, \mathcal{F}^{-1}(v)) \tag{C-1}$$

In Eq. (C-1), $\boldsymbol{\mathcal{F}}(v)$ is the Fisher information matrix calculated with true value of v.

The regularity conditions which $p(\boldsymbol{h}; v)$ should be satisfies can be given as below.

1. The derivative of the log-likelihood function derived by parameters $(\boldsymbol{h}; v)$ should exist.
2. The Fisher information matrix $\boldsymbol{\mathcal{F}}(v)$ should be non-zero.

When regularity conditions are satisfied, the MLE becomes asymptotically unbiased and asymptotically achieves the CRLB. Therefore, MLE can be defined as asymptotically efficient and asymptotically optimal.