AN ABSTRACT OF THE THESIS OF

Navid Lashkarian for the degree of Doctor of Philosophy in Electrical and Computer Engineering presented on June 7, 1999. Title: Optimum Equalization and Synchronization of Broadband Multicarrier Systems .

Abstract approved: Redacted for Privacy

Sayfe Kiaei

The application of information theory and digital signal processing techniques to digital communication has resulted in robust methods for reliable high speed data transmission over noisy channels environments. Among these methods, multicarrier systems have become a viable solution for exploiting maximum spectral efficiency over both wideband highly dispersive static and time-varying dynamic channels. However, there remains many problems in this field to maximize performance and efficiency of such systems. Among these problems, equalization and synchronization are the major focus of this thesis.

The contibution of this thesis on equalization is to optimize performance of equalizer from two different aspects: computational complexity and system capacity. A new fast equalization algorithm is developed to obtain near optimum setting of the equalizer under minimum mean square error criterion. Furthermore, a mathematical framework for analysis of maximum capacity equalization is developed. It is shown that the optimum equalization of multi-carrier system can be viewed as a constrained optimization problem where the constraint set exhibits identifiable geometric characteristic. Using these properties, a descent gradient search method for optimum equalization of multi-carrier systems is proposed.

In synchronization domain, the contribution of this thesis is the development and statistical analysis of a new class of cyclic based estimators: maximum likelihood and minimum variance unbiased and moment estimator for both symbol timing error and carrier frequency offset estimation of multicarrier systems.

Optimum Equalization and Synchronization of Broadband Multicarrier Systems

by

Navid Lashkarian

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Completed June 7, 1999
Commencement June 2000

Doctor of Philosophy thesis of Navid Lashkarian presented on June 7, 1999

APPROVED:

Redacted for Privacy

Major Professor, representing Electrical and Computer Engineering

Redacted for Privacy

Chair of the Department of Electrical and Computer Engineering

Redacted for Privacy

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Redacted for Privacy

Navid Lashkarian, Author

ACKNOWLEDGMENT

I am deeply grateful to my advisor, Professor Sayfe Kiaei for his endless support, guidance and encouragement. This is now an opportunity to acknowledge my debt to him for his generous support during my course of Ph.D program. I shall always remember the unique mentorship, detailed guidance and devoted support of sayfe during my stay at OSU. Working with sayfe and learning from him has been a true privilege for me. His comments and suggestions both clarified and enriched many parts of this document. I would also like to thank him for giving me a free hand to pursue my own interests.

Special thank goes to my graduate committee members, Dr. Gabor Temes, Dr. Wojtek kolodziej , Dr. Un-Ku Moon and Dr. William Warnes for serving on my qualifying and final defense exam.

During the course of my Ph.D program many friends contributed to making the years of graduate school enjoyable. I am indebted to my dear friends Siroos Karimpour, Khadijeh Mokhtari and Alireza Esteghlalian for their constant support and friendship. I owe a great sense of gratitude to my dear friend Amir Niknejad who shared and supported the enthusiasm throughout the last months of my Ph.D program. My appreciation is extended to my colleagues in Dearborn 206, Sumit Talwalkar and Amit Kumar Dutta for the invaluable technical insights and moral lessons that I received from them. Special thanks goes to my friends Masoud Haghi and Emad Bidari for their constant support and companionship through my life. I am greatful to my friends at MIT, Soosan Beheshti and Payman Kassaei for their true friendship. I treasured our philosophical discussions and poetic email exchange which alleviated my stress in the ups and downs that occurred during my Ph.D endeavor. The friendship of all my friends is one of my most valuable assets, and I feel blessed for such a privilege.

I would like to express my deep gratitude to all of my professors at University of Tehran, for their priceless encouragement and invaluable guidance during my undergraduate and graduate studies. I would also like to thank Mr. M Azimi and Mr. E. Javadi for their assistance on all the details involved for the past years of my life.

Finally, my deepest gratitude, love, affection and admiration goes, of course, to my parents Farah and Hassan and my lovely sister, Neda, who constantly showered me with their encouragement and love through my life. I was not the only one who suffered the hardship in attaining this goal, but my family equally suffered an enormous emotional loss in terms of my prolonged absence from home. Thus, they deserve special thanks for their patience. Without any hesitation, they all have inspired me with love of learning and academia. I am therefore, honored to dedicate this thesis to them.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

To my dear parents

Farah and Hassan

and my lovely sister

Neda

1. OPTIMUM EQUALIZATION AND SYNCHRONIZATION OF BROADBAND MULTI-CARRIER SYSTEMS

1.1 Introduction

With rapidly increasing demand for high speed data transmission in information highways, many new communication techniques has been emerged to support reliable data transmission over the media. Various transmission techniques have been proposed to facilitate the communication among users in applications ranging from internet access to cellular network. The choice of one over another depends on the application and performance requirements for the specific system in use.

Among these methods, multi-carrier modulation (MCM) has become a viable communication scheme for high speed data transmission over band-limited channels due to its high bandwidth efficiency [1]. The application of multi-carrier (MC) systems varies from high speed modems for asynchronous digital subscriber lines (ADSL) to digital audio and video broadcast (DAB/DVB) and wireless transceivers [2, 3, 4, 5].

The topic of efficient MC equalization and synchronization has recently received much attention in the area of digital communication [6, 7, 8, 9]. The purpose of this research is to optimize the performance of MCM systems in terms of computational complexity and system performance.

1.2 Motivation

Currently, there are two major trends for equalization of MC systems, namely minimum mean square error (MMSE) and maximum geometrical signal to noise ratio approach.

MMSE equalization of MC systems has received special attention during recent

years [27]. Due to its robust performance, this approach has become a dominant technique for equalization of MC systems in many practical systems. Despite its efficient performance, the computational complexity of MMSE method puts a burden in realization of this scheme for real time applications . In other words, the existing limitations in the processing resources of digital signal processing blocks present a major drawback in real-time realization of MMSE equalizers. Thus, it is desirable to develop efficient signal processing algorithms which provide a balance between computational complexity and performance of such schemes.

On the other hand, optimum equalization of MC systems is a new developed subject which is far from being mature as evidenced by small number of publications devoted to this subject. Although the mathematics and theory of MMSE equalization approach have provided a framework for analysis of such scheme, analysis of maximum capacity equalization remains a challenge in this field. More specifically, the existing theory with this respect fails to provide a unified solid foundation for analysis of such scheme. It is known that equalization of MC systems based on maximum capacity criterion results in a considerable data rate improvement in MC systems. However, this improvement is achieved at expense of solving a highly complex constrained optimization problem. One major concern, in this respect, is the lack of a closed form expression or even an iterative search approach for obtaining the optimal solution of the above problem.

Another interesting challenge in the area of MC systems is in the timing and frequency synchronization of such systems. Since these two problems, timing and offset frequency synchronization, directly affect the performance of MC systems by deteriorating the orthogonality among subcarriers, it is imperative to have efficient methods to estimate and eventually compensate these impairments in the system.

Throughout this thesis, an attempt has been made to address a few of these challenges, and solutions have been proposed to overcome some of these problems. It should also be mentioned that there are still many questions in this field requiring further research work.

1.3 Thesis Outline

This thesis deals with optimum equalization and synchronization of MC systems. Specifically, a framework for analysis of maximum capacity equalization of MC system is proposed. Furthermore, a new family of optimum estimators is developed utilizing the key concept of cyclic prefix.

In Chapter 2 we begin with a detailed description of the single carrier modulation and address its limitation for data transmission over bandlimited channels. Motivated by this shortcoming, we discuss the concept of parallel channel and explain how this approach can benefit data communication in achieving higher data rates. To make this thesis self-contained a brief overview of MC equalization and synchronization methods are also presented in Chapter 2. Analysis of this chapter provides a foundation for deriving novel equalization and synchronization methods in subsequent chapters.

The thesis is then divided into two parts. Part 1 covering Chapters 3 and 4 deals with equalization of MC systems while Part 2, covering Chapter 5, addresses the synchronization issue in MC systems.

Chapter 3 describes a new algorithm for computationally efficient equalization of multicarrier systems over time dispersive static channels. This efficient algorithm allows the equalizer to obtain near optimum solution through estimating the extreme eigenvector of a Hessian matrix. As an application, it is applied to equalization of discrete multi-tone systems. We conclude the chapter with a comparison between complexity and performance of the proposed algorithm against standard equalization methods.

Chapter 4 uses a geometrical method to provide a mathematical framework for analysis of optimum equalization of MC systems. As a result, a new efficient iterative algorithm for maximum capacity equalization of MCM is presented. Due to the versatility of this approach, it is then generalized to two main subclasses of equalizers namely, unit tap (decision feedback) and linear phase filters.

Chapter 5 uses the concept of cyclic prefix for estimation of symbol timing error and carrier frequency offset in MCM. Using probability distribution function of the received

data, a likelihood function for characterizing the effect of synchronization parameters in MC received symbol is presented. Based on this result, a ML estimator for joint estimation of synchronization parameters is derived. We then apply the concept of sufficient statistic to obtain a minimum variance unbiased estimation of carrier frequencty offset. To remove the effect of probability distribution function on the performance of the estimator, a moment estimator for this purpose is also proposed. As an application, the proposed methods are applied for synchronization of orthogonal frequency division modulation scheme.

Finally, Chapter 6 presents conclusions and considerations for further research work in this area.

2. BACKGROUND AND STATE-OF-THE-ART

To make this thesis self-contained, a brief introduction to the basic principles of MC system is given in this chapter. A background theory for analyzing single-carrier and multi-carrier modulation schemes is presented in section 2.1. Sections 2.2 and 2.3, provide an overview of the state-of-the-art equalization and synchronization techniques for MCM respectively.

2.1 Theoretical Background

This section provides an overview of single-carrier and multi-carrier modulation schemes. Section 2.1.1 addresses the single carrier modulation and its limitation in data transmission over bandlimited channels. The concept of parallel channel and channel partitioning methods are thoroughly investigated in sections 2.1.2 to 2.1.6.

2.1.1 Single Carrier Modulation

Capacity of band-limited Gaussian channel was first introduced and analyzed by Shannon in his original paper on information theory. In this paper, he showed that the maximum number of information units that can be transmitted over a band-limited noisy channel with bandwidth (W) can be obtained from

$$C = \log_m(1 + \frac{P}{NW}) \tag{2.1}$$

where P and N are the signal and noise power respectively. Intuitively, this number represents the maximum number of information units (bits) that can be transmitted per channel use, such that the probability of error event remains arbitrarily small (but not zero). In using the above equation, the frequency response of the channel over the entire bandwidth(W) is assumed to be flat. However, most practical communication channels,

such as twisted pair wire lines, exhibit a nonflat frequency response which results in a considerable data rate loss in data transmission [10].

A typical example is the copper phone line where the bandwidth is limited to 3300 Hertz . Assuming a typical signal to noise ratio of 25 dB, the capacity of this channel is limitted to 33000 bits per second. In practice, there are other factors such as cross talk, interference, echos, and non-flat frequency response which must be compensated for in order to achieve this capacity. Among the limitations imposed on the channel, non-flat frequency response is known to be the dominant factor in reducing the effective data rate of transmission system. Conceptually, spectrally shaped channels cause intersymbol interference (ISI) which eventually increases the noise power in detecting the data.

Equalization is the process of compensating a channel with non-flat frequency response into a frequency independent flat spectrum channel. In the time domain, the function of equalizer can be viewed as a system which conditions the channel in order to enforce the channel impulse response into a discrete impulse function. The system which performs this operation is known as equalizer. In many practical systems, equalization is performed by an appropriate filtering operation at the receiver.

Depending on the characteristic of the equalization system, several criteria can be considered for choosing the coefficient of the equalizer. Zero forcing (ZF) and minimum mean square error (MMSE) equalization are the most common schemes which are widely used for in practical systems [11]. In ZF scheme, equalizer taps are set such that the frequency response of the equalizer becomes the inverse of channel spectrum. Thus the frequency response of the overall system becomes flat, and there would be no ISI in the system. However, the effectiveness of this approach is limited to the class of minimum phase channels, where the existence of stable inverse systems is guarantied [12]. Also, it is known that ZF equalizers may result in considerable noise enhancement around singularities of the equalizer spectrum(poles of the transfer function).

In MMSE equalization, a set of previously received data and detected symbols are weighted by the equalizer taps to minimize the estimation error introduced by the noise and ISI [13]. To assure stability, the mean square error is used as a global function for

minimization purposes. In doing so, a set of feedback and feedforward filters are used at the receiver. MMSE equalization has the advantage of analytical tractability and mitigation of the noise-enhancement problem of the zero-forcing criterion. However, complexity of this approach is exponentially proportional to the effective memory of the channel impulse response.

Complexity of the equalizer is directly proportional to the effective memory of the channel. Consequently, as the effective length of channel impulse response increases, the complexity of these approaches becomes considerably high. Consequently, high bandwidth data transmission schemes require sophisticated equalization techniques along with highly complex equalizers. Thus, none of the standard equalization methods can be used as an effective technique for mitigating the ISI in these scenarios .

An alternative approach for coping this problem is the MCM [14]. In MCM, the channel spectrum is partitioned into a large number of orthogonal , parallel, and approximately memoryless channels. The number of input bits assigned to different subchannels are different, and depends on the signal to noise ratio in each subchannel. In the following section, the problem of optimum bit allocation for parallel channels is thoroughly analyzed.

2.1.2 Parallel Channel and Water-filling

Consider a set of k independent Gaussian channels, with a common power constraint. We intend to distribute the transmit power (P) among the subchannels ($[\ P_1 \quad P_2 \quad \cdots \quad P_k\ ]$) so as to maximize the overall capacity of the system. Also, it is assumed that contribution of the additive noise to the subchannels can be represented by the vector ($[\ N_1 \quad N_2 \quad \cdots \quad N_k\ ]$). A typical example of this scenario is the the set of k parallel QAM modulation schemes, where each subchannel can be treated as an AWGN channel with no memory as shown in Figure 2.1. To satisfy the power constraint, the optimization is performed subject to the constraint $\sum_i P_i = P$. By generalizing the concept of capacity

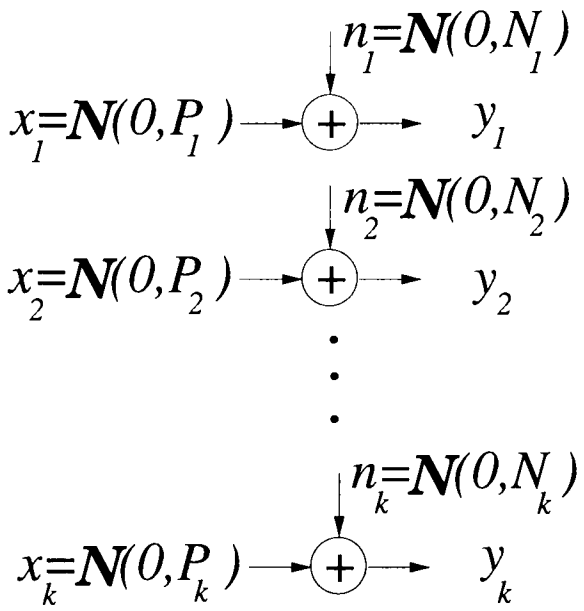


Figure 2.1: Parallel Gaussian Channel

given in section 2.1.1 to MCM, the overall capacity of the composite system can be written as a function of the individual power allocated to each channel. Therefore, the Lagrangian functional for the above constraint problem can be expressed as

$$J(P_1, P_2, \cdots, P_k) = \sum_{i=1}^{k} \frac{1}{2} log(1 + \frac{P_i}{N_i}) + \lambda(P - \sum_{i=1}^{k} P_i) \tag{2.2}$$

By taking the derivative of the Lagrangian with respect to each P_i, and setting it to zero, the optimum solution can be obtained from

$$\frac{1}{2} \frac{1}{P_i + N_i} + \lambda = 0 \tag{2.3}$$

or

$$P_i = v - N_i \tag{2.4}$$

Since P_i's must satisfy the non-negativity constraint, it might not be always possible to find a solution of this form. Therefore Kuhn-Tucker condition can be applied to verify that the solution

$$P_i = (v - N_i)^+ \tag{2.5}$$

is the assignment that maximizes capacity, where v is chosen so that

$$\sum_i (v - N_i)^+ = P \tag{2.6}$$

This solution is illustrated graphically in figure 2.2. The vertical levels indicate the noise levels in the various channels. As signal power is increased from zero, we allot the power to the channels with the lowest noise. When the available power is increased still further, some of the power is put into noisier channels. The process by which the power is distributed among the various bins is identical to the way in which water distributes itself in a vessel. Hence this is sometimes referred to as water-filling [15]. The water filling property is the principle theory of data rate maximization in high speed data transmission. While the water-filling energy allocation will indeed yield the optimal solution, it is often difficult to compute and tacitly assumes infinite granularity in constellation size, which is not realizable. One known finite-granularity multicarrier loading algorithm is the Hughes-Hartogs algorithm [16]. However, this algorithm is very slow for high speed applications, such as ADSL environment, where a large number of bits will be contained in each DMT symbol. A low complexity loading algorithm was also proposed in [17]. The key point in this approach is to distribute the rate among the subchannels according to the channel capacity. This approach maximizes the capacity for a given signal power. However, there are some applications in practice where the objective is to transmit a fixed data rate with a fixed power at the lowest error rate.

In [18, 19], authors use the maximum rate loading criteria to allocate bits among the subcarriers.

Channel partitioning technique has received a special attention during the recent years [20]. As a result, several forms of MCM have been developed depending on the choice of the modulating and demodulating vectors. In the following chapter, we recapitulate the general theory of channel partitioning and provide a unified framework for analyzing such systems.

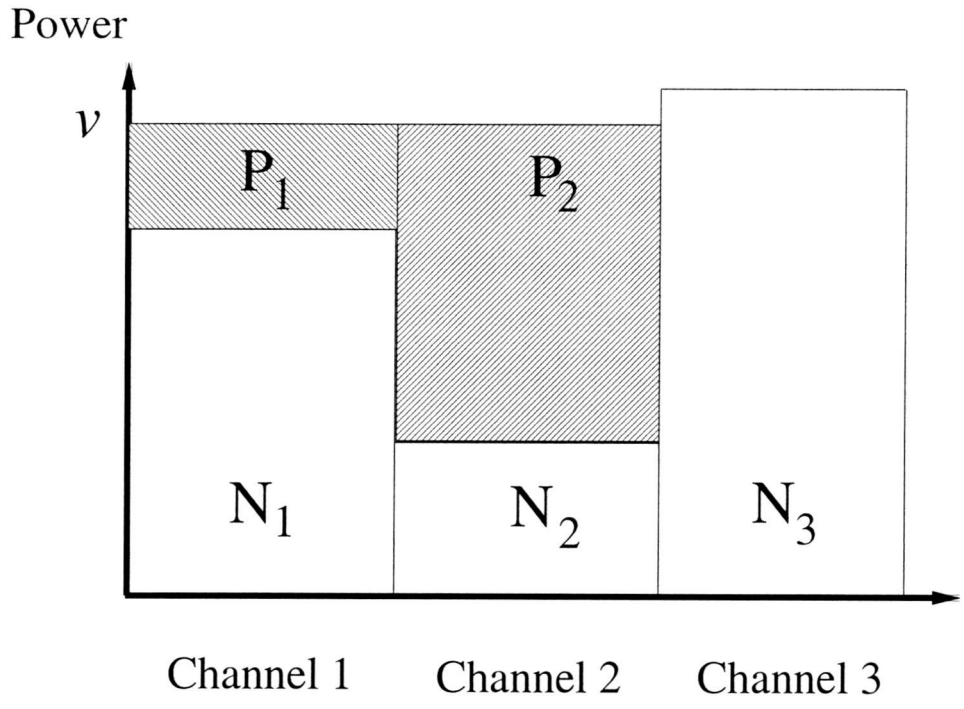


Figure 2.2: Typical Water-Filling Distribution

2.1.3 Vector Space Approach

Consider a block of N information symbols which is used for data transmission. By virtue of vector space concept, this N-tuple vector can be represented as a linear combination of N orthogonal basis functions as

$$x_k(t) = \sum_{n=1}^{N} x_n \phi_n(t) \tag{2.7}$$

Consequently, a succession of such transmission with symbol period of T can be represented as

$$x(t) = \sum_{k} \sum_{n=1}^{N} x_{n,k} \phi_n(t - kT) \tag{2.8}$$

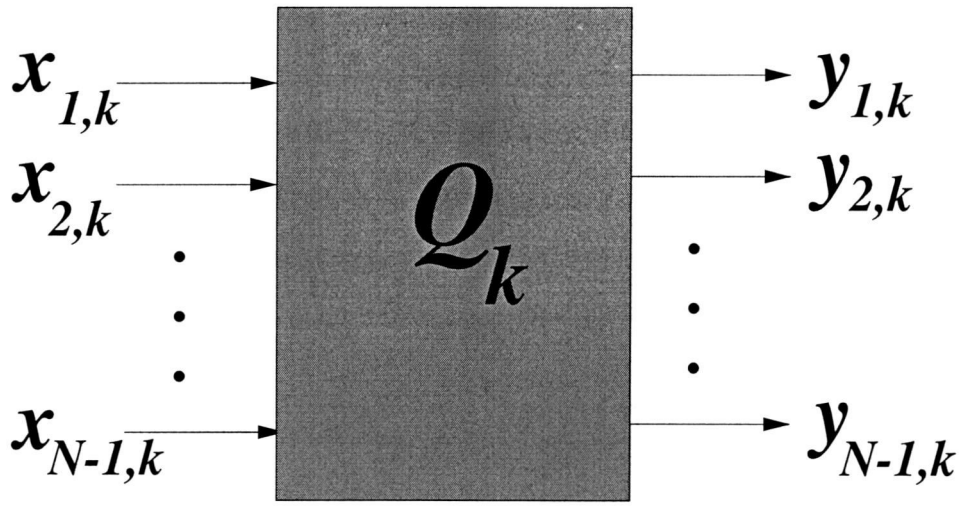


Figure 2.3: Block diagram of MMSE Filter

Considering this function as an input to a channel with impulse response $h(t)$, the received signal at the output of channel can be represented as a linear convolution according to

$$y(t) = h(t) * x(t) = \sum_{k} \sum_{n=1}^{N} x_n \phi_n(t - kT) * h(t) = \sum_{n=1}^{N} x_{n,k} \phi_n(t) \tag{2.9}$$

In order to detect the modulated vector, a set of N matched filters matched to pulse responses $\phi_n(t)$ are used at the receiver. Thus, the maximum likelihood estimate of the transmitted signal can be obtained by maximizing the signal to noise ratio at the output of each filter. In doing so, there are N^2 describing functions

$$q[m,n](t) = \frac{p_n(t) * p_m^*(-t)}{||p_n||\ ||p_m||} \tag{2.10}$$

that characterize the modulation and demodulation process. Consequently, the above modulation scheme can be viewed as a multiple-input/multiple-output system in which the discrete-time I/O transfer function can be viewed as an NxN time-varying matrix $\mathbf{Q}$ where the entries of the matrix have the following representation

$$\mathbf{Q}_k[m,n] = q[m,n](kt) \tag{2.11}$$

To avoid ISI in data transmission, the characteristic matrix $\mathbf{Q}$ should satisfy the

Nyquist condition

$$\mathbf{Q}_k[m,n] = I \tag{2.12}$$

where I is the identity matrix. There are different ways for selecting the basis functions to satisfy the equation 2.12. A nonoptimum choice of these functions would remove the orthogonality among the describing functions. This phenomena can be viewed as the orthogonality distortion caused by nonideal transmit pulse shapes in PAM or QAM transmission. In the following section, we investigate an optimum set of basis functions which result in an ISI-free data transmission.

2.1.4 Modal Modulation

Given an impulse response $h(t)$ the corresponding autocorrelation function can be defined as

$$r(t) = h(t) * h^*(-t) \tag{2.13}$$

The set of corresponding eigenfunctions for this function is defined as functions $\phi_n(t)$ which satisfy the following relationship

$$\int_{T/2}^{-T/2} r(t-\tau)\phi_n(\tau) = \rho_n\phi_n(t)d\tau \tag{2.14}$$

$$\int_{T/2}^{-T/2} \phi_i(t)\phi_j^*(t)dt = \delta_{ij} \tag{2.15}$$

One way of selecting the basis functions is to assign each basis function to a eigenfunction of the channel autocorrelation function. The block diagram of such system is depicted in Figure 2.4. Using this assumption, we can express the received signal $r(t)$ as

$$r(t) = h(t) * x(t) + n(t) \tag{2.16}$$

where $n(t)$ is the additive white Gaussian noise with power of $\frac{N}{2}$. The channel matched filter $(h^*(T-t))$ output is given by

$$y(t) = \sum_{n=1}^{N} x_n[r(t) * \phi_n(t)] + \tilde{n}(t) \tag{2.17}$$

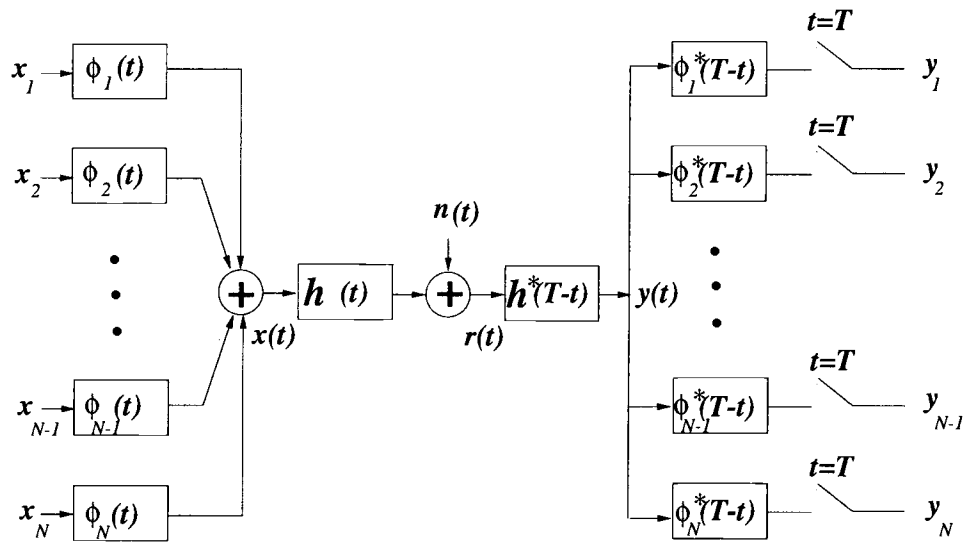


Figure 2.4: Block diagram of MMSE Filter

where $\tilde{n}(t)$ is the response of channel matched filter to the input noise $n(t)$. Through equation 2.14, the above expression can be written as

$$y(t) = \sum_{n=1}^{N} x_n \rho_n \phi_n(t) + \tilde{n}(t) \tag{2.18}$$

The receiver matched filters ($\phi_1(t)$,$\phi_2(t)$,$\cdots$, $\phi_1(t)$) consist of N filters each matched to the corresponding eigenfunction. Therefore, the sampled noise output at the output of receiver matched filter has the following form

$$\tilde{n}_i = \int_{-T/2}^{T/2} n(t-\tau)\phi_i(t)dt \tag{2.19}$$

and the autocorrelation of this sequence can be written as

$$E[\tilde{n}_i \tilde{n}_j^*] = \rho_i \delta_{ij} \frac{N}{2} \tag{2.20}$$

Thus, this particular choice of basis functions partitions the channel into a set of N independent subchannels to which the results of water-filling property can be applied. However, this particular selection requires prior knowledge of channel impulse response and may vary from one channel to another. Such functions would be difficult to implement

exactly in practice. In the next section, we address a particular technique which provides an efficient partitioning method which is independent of channel impulse response.

2.1.5 Discrete-time Channel Partitioning

Consider a communication channel in which impulse response can be modeled as a FIR filter $h = \{h_0, h_1, \cdots, h_v\}$. For a block of N_f received samples, the discrete-time input/output relation of channel can be cast in matrix form as follows

$$\begin{bmatrix} \mathbf{y}[N_f-1] \\ \mathbf{y}[N_f-2] \\ \vdots \\ \mathbf{y}[0] \end{bmatrix} = \begin{bmatrix} h[0] & h[1] & \cdots & h[v] & 0 & \cdots & 0 \\ 0 & h[0] & h[1] & \cdots & h[v] & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & h[0] & \cdots & h[v-1] & h[v] \end{bmatrix} \begin{bmatrix} \mathbf{x}[N-1] \\ \mathbf{x}[N-2] \\ \vdots \\ \mathbf{x}[-v] \end{bmatrix}$$

or more compactly,

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{2.21}$$

where $\mathbf{y}$, $\mathbf{x}$ and $\mathbf{n}$ are the vectors of output, input and noise respectively. The $N\text{x}(N+v)$ matrix has a singular value decomposition (SVD) of

$$\mathbf{P} = \mathbf{F}\begin{bmatrix} \Lambda & \vdots & \mathbf{0}_{\mathbf{N,v}} \end{bmatrix}\mathbf{M}^* \tag{2.22}$$

where the matrices $\mathbf{F}$ and $\mathbf{M}$ are unitary matrices ($\mathbf{F}^*\mathbf{F} = I_N$ and $\mathbf{M}^*\mathbf{M} = I_{N+v}$) of size $N\text{x}\ N$ and $(N+v)\text{x}\ (N+v)$ respectively. If the input vector is generated carefully according to

$$\mathbf{x} = \mathbf{M}\begin{bmatrix} X_{N-1} & X_{N-2} & \vdots & X_0 & 0 & \vdots & 0 \end{bmatrix}^* \tag{2.23}$$

then the modulated signal can be perfectly reconstructed at the receiver as

$$\mathbf{Y} = \mathbf{F}^*\mathbf{F}\begin{bmatrix} \Lambda & \vdots & \mathbf{0}_{\mathbf{N,v}} \end{bmatrix}\mathbf{M}^*\mathbf{M}\mathbf{x} + \mathbf{F}^*\mathbf{n} = \Lambda\mathbf{X} + \mathbf{N} \tag{2.24}$$

Therefore equalization at the receiving end can be performed efficiently using hadamard division operation. This operation is known as frequency domain equalization (FEQ) in

MC systems. The main disadvantage of this scheme is in the relatively high overhead associated with computing SVD. However, as is explained in subsequent section, an appropriate coding technique can be applied prior to transmitting data over channel in order to reduce the complexity of SVD.

2.1.6 Channel Partitioning For Discrete Multitone Systems

Discrete multitone(DMT) system is a special case of vector coding which provides a reduced complexity channel partitioning through adding some restriction to the data. This scheme, forces the transmit symbol to have $x_{-k} = x_{N-k}$ for $k = [1, \cdots, v]$. Such repeating of the last v samples at the beginning of the symbol is called **cyclic prefix**. With cyclic prefix, the input/output channel description can be cast in the following form

$$\begin{bmatrix} \mathbf{y}[N_f-1] \\ \mathbf{y}[N_f-2] \\ \vdots \\ \mathbf{y}[0] \end{bmatrix} = \begin{bmatrix} h[0] & h[1] & \cdots & h[v] & 0 & \cdots & 0 \\ 0 & h[0] & h[1] & \cdots & h[v] & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & h[0] & \cdots & h[v-1] & h[v] \\ h[v] & 0 & \cdots & 0 & h[0] & \cdots & h[v-1] \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ h[1] & \cdots & h[v] & 0 & \cdots & 0 & h[0] \end{bmatrix} \begin{bmatrix} \mathbf{x}[N-1] \\ \mathbf{x}[N-2] \\ \vdots \\ \mathbf{x}[0] \end{bmatrix} \tag{2.25}$$

or in compact form

$$\mathbf{y} = \tilde{\mathbf{H}}\mathbf{x} + \mathbf{n} \tag{2.26}$$

Due to the circular characteristic of the channel description matrix, it can be decomposed into products of Fourier matrices and a diagonal matrix as

$$\tilde{\mathbf{H}} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^* \tag{2.27}$$

In the above equation, $\mathbf{U}$ is the Fourier matrix, and $\mathbf{\Lambda}$ is a diagonal matrix, in which the diagonal elements are the N_f-point DFT of the first row of matrix $\tilde{\mathbf{H}}$, where N_f is the size of matrix $\tilde{\mathbf{H}}$. Due to the unitary characteristic of matrix $\mathbf{U}$ A, ($\mathbf{U}\mathbf{U}^* = \mathbf{U}^*\mathbf{U} = \mathbf{I}$)

the decoding and encoding procedure can be performed efficiently using FFT and IFFT transform as follows

$$\mathbf{Y} = \mathbf{U}^*\mathbf{y} = \mathbf{U}^*\tilde{\mathbf{H}}\mathbf{x} + \mathbf{n} = \mathbf{\Lambda}\mathbf{U}^*\mathbf{x} + \tilde{\mathbf{n}} \tag{2.28}$$

Therefore if the input signal is chosen to satisfy

$$\mathbf{x} = \mathbf{U}^*\mathbf{X} \tag{2.29}$$

then the demodulated vector at the receiver can be represented as

$$\mathbf{Y} = \mathbf{\Lambda}\mathbf{X} + \tilde{\mathbf{n}} \tag{2.30}$$

in which $\tilde{\mathbf{n}} = \mathbf{U}^*\mathbf{n}$.

Figures 2.5 and 2.6 depict the block diagram of DMT transmitter and receiver respectively. From the above equation, it this clear that appending cyclic prefix to the modulated symbol provides an efficient way for computing the singular values of the channel description matrix. Using the circular characteristic of the channel description matrix, the SVD operation can be performed with $N_f log_2 N_f$ operations (FFT operation). Due to the additional restriction of cyclic prefix, performance of the DMT systems is always upper-bounded by the performance of vector coding. When $N_f >> v$, the difference between DMT and VC would be small. However, in twisted pair wirelines, this assumption is violated and extra processing is needed to enforce the highly time dispersive channel into relatively short channel. This operation, known as impulse response shortening, is the topic of the following section [21, 22]

2.2 Overview of Equalization Methods

In this section we address the impulse response shortening problem and recapitulate the existing methods for training the time domain equalizer (TEQ) for DMT systems. According to the previous discussion on channel partitioning, TEQ plays an important rule in performance of DMT systems. As long as the TEQ effectively shortens the channel, the

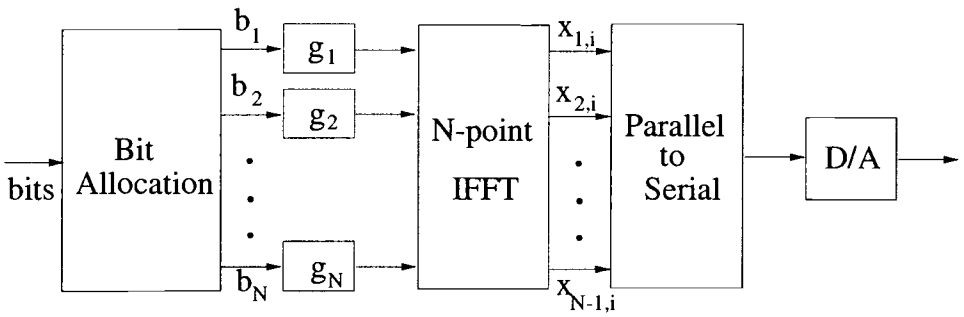


Figure 2.5: Block diagram of DMT transmitter

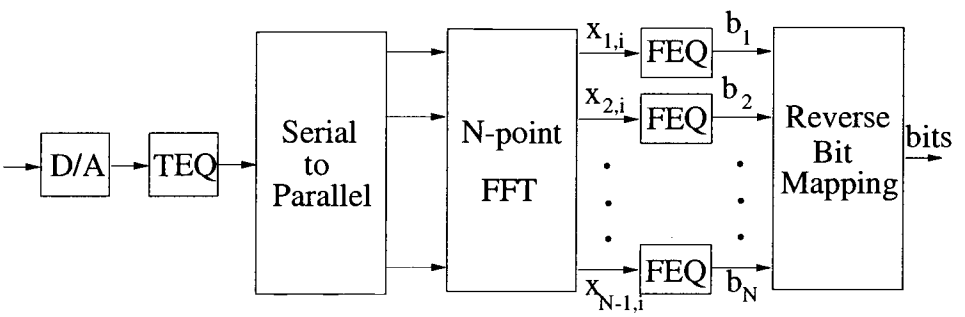


Figure 2.6: Block diagram of DMT receiver

system is free from ISI, which in turn guarantees that the system is free from intercarrier interference (ICI).

Equalization of multicarrier was originally discussed by Weinstein and Ebbert [23]. However, the proposed approach only encounters the distortion due to the cochannel interference and completely ignores the interchannel interference. Equalization of MC signal that has passed through a distorting channel was first addressed by Hirosaki [24].

Pioneering efforts for practical equalization of MC systems were started by Chow et. al. [25]. In this paper, authors characterize the channel by a rational transfer function in the form of $\frac{a(z)}{p(z)}$ where $a(z)$ and $p(z)$ are the polynomial of finite degree (FIR). In addition, the equalizer is set to $p(z)$, the numerator of the transfer function. Consequently, the

equalized channel was characterized by the FIR filter $a(z)$. Due to the high computational complexity involved with the channel identification procedure, this method is inefficient for channels with relatively high memory. Soon after, an LMS-like adaptive algorithm was proposed by Chow et. al. which facilitated the training procedure for setting the TEQ taps. A salient feature of this algorithm is that it jointly adapts the decision delay and equalizer coefficients in each iteration. However, as stated by the authors later, the algorithm exhibits anomalous behaviour such as increasing the number of iterations. Although it is adhoc and not globally convergent, this approach is considered as the most practical scheme for setting the equalizer in many practical systems. Later, in a complementary paper, authors proposed a DFE-like filtering technique and formulated the problem into a quadratic optimization problem [26]. It was shown in this paper, that the DFE may not necessarily result in minimum mean square error and in optimizing the performance of TEQ various causality scenarios have to be considered . Later, the challenge for removing these imperfections resulted in a comprehensive paper by Al-Dharir and Cioffi, wherein a solid framework for analyzing the MMSE equalization was proposed [27]. Fundamentally, this approach provides the optimum setting for a finite length input aided equalizer by minimizing the MSE between the equalized channel impulse response and a target impulse response (TIR). Also, this approach devises a search procedure for computing the optimum delay associated with TIR which outperforms the previous search technique proposed for this purpose [26]. Due to the high relevancy of this subject to the material developed in chapters 3 and 4, this theory is thoroughly investigated in section 2.2.1.

Although efficient and mathematically solid, MMSE equalization has a high computational complexity which puts a burden in its application for real time systems. To remove this impediment, a few fast versions of this approach were proposed in [28, 29]. In this approach, a periodic input signal is used for training the equalizer. Due to this periodicity, the autocorrelation matrix of the received vector appear to be circulant and therefore its inverse can be obtained efficiently using DFT operation. Also note that the proposed method imposes the unit tap constraint on the equalizer. There is also a recent

interest in application of multirate filter banks for impulse response shortening problem [30]. Recently, the topic of optimum equalization of multicarrier systems has received special attention by many researchers. Al-Dhahir et. al., showed that although MMSE equalization is the most popular equalization technique due to its tractability and adaptability, it does not optimize the criterion in conjunction with DMT systems [7, 8]. Having established this property, a new criterion was proposed and the optimum equalization of DMT systems was formulated as a constrained optimization problem. This subject is the topic of section 2.2.2. In the next section we reformulate the MMSE equalization problem and consider various approaches in setting the equalizer taps.

2.2.1 MMSE Equalization

The optimum setting of MMSE-DFE equalizers for ML detection have already been investigated in many papers [27], [31]. In order to provide a sufficient background for the following chapters, we recapitulate the theory of MMSE equalization.

Consider an additive Gaussian noise channel in which the received signal $y(t)$ is constructed by modulating a series of input symbols x_l by a channel response $h(t)$,

$$y(t) = \sum_{n=-\infty}^{\infty} x_p h(t-nT) + n(t) \tag{2.31}$$

in which $\{x_l\}$ is a sequence of iid zero mean random variables with power σ_x^2. The noise $n(t)$ is assumed to be zero mean Gaussian and independent of the input sequence. The channel response represents the combined effect of the transmit and receive filter as well as channel impulse response. By resorting to the concept of fractionally spaced equalization, the input/output relation for the discrete time equivalent channel can be represented in the following form

$$\mathbf{y}_k = \sum_{l=-\infty}^{l=\infty} \mathbf{h}_l x_{k-l} + \mathbf{n}_k \tag{2.32}$$

where the output, input and noise vectors are defined as

$$\mathbf{y}_k^* = \begin{bmatrix} y[k+1-1/M] & y[k+1-2/M] & \cdots & y[k] \end{bmatrix}$$

$$\mathbf{h}_k^* = \begin{bmatrix} h[k+1-1/M] & h[k+1-2/M] & \cdots & h[k] \end{bmatrix}$$

$$\mathbf{n}_k^* = \begin{bmatrix} n[k+1-1/M] & n[k+1-2/M] & \cdots & n[k] \end{bmatrix}$$

In the above expressions, the sequence of samples $y[k]$, $h[k]$ and $n[k]$ are obtained through periodic sampling of the functions $y(t)$, $h(t)$ and $n(t)$ with period of T/M respectively. The use of vector channel impulse response coefficients assumes a poly-phase representation of the channel. In other words, $\mathbf{h}_k$ is an M-dimensional column vector that contains M phases of the kth symbol period and M is the oversampling factor.

Consider a block of N_f output samples. Successive using of equation 2.31 provides a set of difference equation in a matrix form as following

$$\begin{bmatrix} \mathbf{Y}_k \\ \mathbf{X}_{k-\Delta} \end{bmatrix} = \begin{bmatrix} \mathbf{H} \\ \mathbf{S} \end{bmatrix} \mathbf{X}_k + \mathbf{N}_k \tag{2.33}$$

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}_0 & \mathbf{h}_1 & \cdots & \mathbf{h}_v & 0 & \cdots & 0 \\ 0 & \mathbf{h}_0 & \mathbf{h}_1 & \cdots & \mathbf{h}_v & 0 & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{h}_0 & \mathbf{h}_1 & \cdots & \mathbf{h}_v \end{bmatrix} \tag{2.34}$$

$$\mathbf{S} = \begin{bmatrix} \mathbf{0}_{(N_b+1,\Delta)} & \mathbf{I}_{(N_b+1,N_b+1)} & \mathbf{0}_{(N_b+1,N_f+v-N_b-1)} \end{bmatrix} \tag{2.35}$$

$$\tag{2.36}$$

where,

$$\mathbf{X}_{k-\Delta}^* = \begin{bmatrix} x_{k-\Delta} & x_{k-\Delta-1} & \cdots & x_{k-\Delta-N_b} \end{bmatrix} \tag{2.37}$$

$$\mathbf{X}_k^* = \begin{bmatrix} x_k & x_{k-1} & \cdots & x_{k-\Delta-N_f-v+1} \end{bmatrix} \tag{2.38}$$

$$\mathbf{N}_k^* = \begin{bmatrix} \mathbf{n}_k^* & \mathbf{n}_{k-1}^* \cdots & \mathbf{n}_{k-N_f+1}^* \end{bmatrix} \tag{2.39}$$

$$\mathbf{Y}_k^* = \begin{bmatrix} \mathbf{y}_k^* & \mathbf{y}_{k-1}^* & \cdots & \mathbf{y}_{k-N_f+1}^* \end{bmatrix} \tag{2.40}$$

As shown in Figure (2.7), MMSE equalizer consists of two filters, a fractionally spaced feedforward filter $\mathbf{w}$ with MN_f taps, also known as time domain equalizer (TEQ),

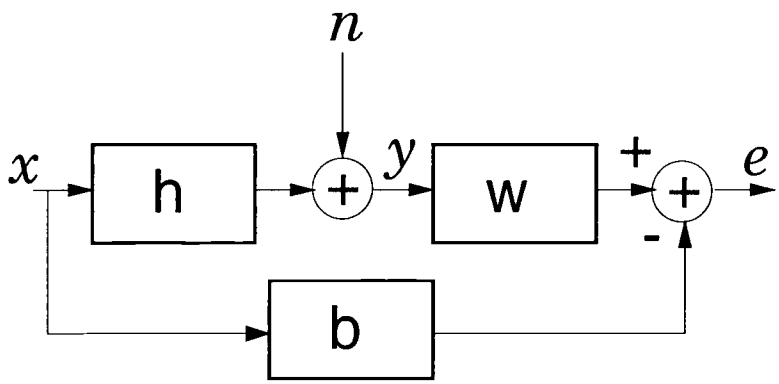


Figure 2.7: Block diagram of MMSE Filter

and a symbol spaced feedback filter with $N_b + 1$ taps, known as target impulse response (TIR). For a given value of filter lengths N_b and N_f the objective is to compute the MN_f coefficients of the feedforward equalizer, w, and the $N_b + 1$ taps of the feedback filter b such that the mean square of the error sequence $e_k = z_k - r_k$ is minimized. In doing so, we compute the residual error according to

$$e = \mathbf{w}^*\mathbf{Y}_k - \mathbf{b}^*\mathbf{X}_{k-\Delta} \tag{2.41}$$

in which $\mathbf{w}$ and $\mathbf{b}$ are the vectors of feedforward and feedback filter coefficients respectively,

$$\mathbf{b}^* = \begin{bmatrix} b_0 & b_1 & \cdots & b_{N_b} \end{bmatrix}$$
$$\mathbf{w}^* = \begin{bmatrix} w_1 & w_2 & \cdots & w_{N_f.M} \end{bmatrix}$$

Therefore the mean square error can be represented as

$$\mathbf{E}[e_k^2] = \mathbf{b}^*\mathbf{R}_{xx}\mathbf{b} - \mathbf{b}^*\mathbf{R}_{xy}\mathbf{w} + \mathbf{w}^*\mathbf{R}_{yy}\mathbf{w} - \mathbf{w}^*\mathbf{R}_{yx}\mathbf{b}$$

In the above expression, $\mathbf{R}_{xy} \triangleq \mathbf{E}[\ \mathbf{X}_{k-\Delta}\mathbf{Y}_k^*] = \mathbf{R}_{yx}^*$, $\mathbf{R}_{xx} \triangleq \mathbf{E}[\ \mathbf{X}_{k-\Delta}\mathbf{X}_{k-\Delta}^*]$, $\mathbf{R}_{yy} \triangleq \mathbf{E}[\ \mathbf{Y}_k\mathbf{Y}_k^*]$. By invoking the orthogonality principle ($\mathbf{b}^*\mathbf{R}_{xy} = \mathbf{w}^*\mathbf{R}_{yy}$), the mean square error can be written in the form of

$$\mathbf{E}[e^2] = \mathbf{b}^* \mathbf{R}_{x|y} \mathbf{b} \tag{2.42}$$

in which

$$\mathbf{R}_{x|y} = \mathbf{R}_{xx} - \mathbf{R}_{xy} \mathbf{R}_{yy}^{-1} \mathbf{R}_{yx} \tag{2.43}$$

Equation 2.42 is in the form of quadratic function which can be minimized using standard Newton's method. In order to avoid converging to the trivial solution an additional constraint is imposed on the Feedback filter.

1- Unit Tap Constraint (UTC)

In this approach, ith element of the feedback filter b is forced to unity. This condition is imposed through adding the constraint $1 - \mathbf{u_i}^* \mathbf{b} = \mathbf{0}$ to the original optimization problem, in which $\mathbf{u_i}$ is the i'th unit vector. In light of this fact, the optimization problem is formulated as,

$$\mathbf{b} = \arg \min_{\mathbf{b}} \; \mathbf{b}^* \mathbf{R}_{x|y} \mathbf{b} + \lambda (1 - \mathbf{u_i}^* \mathbf{b}) \tag{2.44}$$

Using Lagrange functional, the optimum vector $\mathbf{b}_{opt}$ which provides the global minimum of the above function can be obtained from

$$\mathbf{b} = \frac{\mathbf{R}_{x|y}^{-1} \mathbf{u_i}}{\mathbf{u_i}^* \mathbf{R}_{\mathbf{x|y}}^{-1} \mathbf{u_i}} \tag{2.45}$$

and the minimum mean square error is accordingly found from

$$\mathbf{MSE_{UTC}} = \frac{1}{\mathbf{R}_{x|y}^{-1}(i,i)} \tag{2.46}$$

2- Unity Energy Constraint (UEC)

In this approach, feedback filter satisfies the unit energy constraint. This condition is satisfied through adding the unit energy constraint to the optimization problem which results in a objective function as follows,

$$\mathbf{b} = \arg \min_{\mathbf{b}} \; \mathbf{b}^* \mathbf{R}_{x|y} \mathbf{b} + \lambda (1 - \mathbf{b}^* \mathbf{b}) \tag{2.47}$$

Similarly, the optimum vector $\mathbf{b}_{opt}$ which provides the global minimum solution of the above objective function is the form of

$$\mathbf{b} = \arg \min_{\mathbf{b}} \frac{\mathbf{b}^* \mathbf{R}_{\mathbf{x}|\mathbf{y}} \mathbf{b}}{\mathbf{b}^* \mathbf{b}} \tag{2.48}$$

The solution to this problem is the eigenvector corresponding to the minimum eigenvalue of matrix $\mathbf{R}_{x|y}$. Therefore, the minimum mean square error obtained from using this approach would be the minimum eigenvalue of matrix $\mathbf{R}_{x|y}$ as

$$\mathbf{MSE_{UEC}} = \lambda_{\mathbf{min}}(\mathbf{R}_{\mathbf{x}|\mathbf{y}}) \tag{2.49}$$

Thus, the optimum solution may be obtained from inverting the Hessian matrix $\mathbf{R}_{x|y}$ followed by applying power iteration algorithm. A close look at equation 2.48 reveals that the MMSE-UEC problem can be viewed as minimizing the Rayleigh quotient of matrix $\mathbf{R}_{x|y}$. Later, we will use this property to obtain a fast algorithm for computing a near optimum solution for MMSE-UEC approach.

In both approaches, the optimum equalizer settings for the feedforward filter $\mathbf{w}$ is obtained from solving Wiener-Hopf equation as

$$\mathbf{w} = \mathbf{R}_{yy}^{-1} \mathbf{R}_{yx} \mathbf{b} \tag{2.50}$$

This unified approach requires one-time inversion of an MN_f-dimensional matrix $\mathbf{R}_{yy}$ to compute $\mathbf{R}_{x|y}$. For the special case where the input spectrum has a white spectrum with twosided power spectral density of σ_x^2, the cross correlation matrices $\mathbf{R}_{xy}$ and $\mathbf{R}_{yx}$ can be directly obtained from the channel description matrix as

$$\begin{aligned} \mathbf{R}_{xy} &= \mathbf{E}[\, \mathbf{X}_{k-\Delta} \mathbf{Y}_k^*] = \mathbf{E}[\, \mathbf{X}_{k-\Delta} (\mathbf{H}_k \mathbf{X} + \mathbf{N}_k)^*] \\ &= \mathbf{E}[\, \mathbf{X}_{k-\Delta} \mathbf{X}^* \mathbf{H}_k^*] = \mathbf{E}[\, \mathbf{X}_{k-\Delta} [\, \mathbf{X}_a^* \, \mathbf{X}_{k-\Delta}^* \, \mathbf{X}_b^*] \, \mathbf{H}^*] = \mathbf{S} \, \mathbf{H}^* \end{aligned}$$

wherein the vectors $\mathbf{X}_a$, $\mathbf{X}_b$ and the matrix $\mathbf{S}$ is defined as

$$\mathbf{X}_a^* = \begin{bmatrix} x_k & x_{k-1} & \cdots & x_{k-\Delta+1} \end{bmatrix}$$

$$\mathbf{X}_b^* = \begin{bmatrix} x_{k-\Delta-N_b-1} & x_{k-\Delta-N_b-2} & \cdots & x_{k-N_f-v+1} \end{bmatrix}$$

$$\mathbf{S} = \begin{bmatrix} \mathbf{0}_{N_b+1,\Delta} & \mathbf{I}_{N_b+1,N_b+1} & \mathbf{0}_{N_b+1,s} \end{bmatrix}$$

In the above expression, Δ is the decision delay involved with the feedback filter and $s \triangleq N_f + v - N_b - 1$. Matrices $\mathbf{I}$, $\mathbf{0}$ and $\mathbf{R}_{nn}$ represent the identity, zero, and noise autocorrelation matrices, respectively. Using the new expression for crosscorrelation matrices, the Hessian matrix $\mathbf{R}_{x|y}$ for the white input sequence would be in the form of,

$$\mathbf{R}_\Delta \triangleq \sigma_x^2 \mathbf{I}_{N_b+1} - \sigma_x^4 \mathbf{S}^* \mathbf{H}^* \mathbf{R}_{yy}^{-1} \mathbf{H} \mathbf{S}^* = \sigma_x^2 \mathbf{I}_{N_b+1} - \sigma_x^4 \mathbf{H}_\Delta^* \mathbf{R}_{yy}^{-1} \mathbf{H}_\Delta \tag{2.51}$$

where $\mathbf{H}_\Delta \triangleq \mathbf{H}\mathbf{S}^*$. From the above expression one can conclude that changing the decision delay would result in different objective functions which changes the MSE. Therefore, the optimum value for the parameter Δ is obtained from performing exhaustive search on all values of $\Delta = 0, 1, \cdots, N_f + v$ and selecting the one which provides the minimum value of the objective function given in 2.51. It can be shown that the MMSE error under UEC results in a lower value in comparison to the MSE provided by UTC approach [27].

Figure 2.8 shows the result of the MMSE approach for both UTC and UEC constraints on a typical twisted pair wireline, namely $9Kft$ 26 AWG channel.

Although the problem of finding optimum setting of coefficients in FIR MMSE-DFE has been addressed thoroughly, this technique is still not attractive for real time applications due to the following reasons

- Computational complexity

  Due to the fairly complex matrix operations associated with this MMSE equalization approach, , this scheme is still not attractive for real time applications. In Chapter 3 we propose an efficient technique which removes this impediment in practical realization of this technique.

- Suboptimum performance

  The MMSE equalization method obtains the setting of the equalizer by minimizing the mean square error in impulse response shortening . However, it is known that

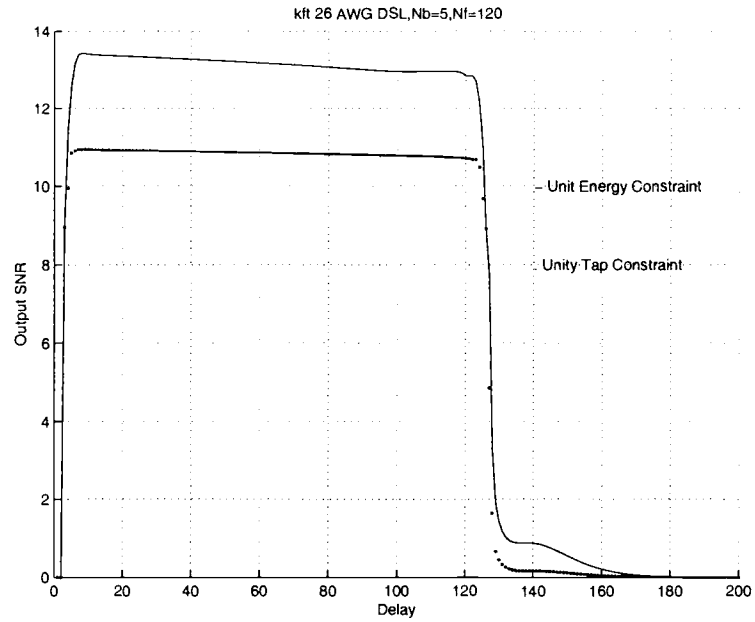


Figure 2.8: Comparison between MMSE-UEC and MMSE-UTC equalization approach for 9Kft 26 gauge AWGN channel

this criterion does not result in the best system performance in conjunction with MC systems. In other words, the capacity obtained from using the MMSE equalization approach may not necessarily result in maximum data rate for MC system. The problem of optimum equalization of DMT system is the subject of the following section.

2.2.2 Maximum Capacity Equalization

The problem of optimum equalization of MC systems is investigated in this section. Consider a MC system consisting of $\bar{N}$ independent and equally spaced subchannels. By virtue of equation 2.1, the capacity of the composite system can be expressed as

$$b_{DMT} = \sum_{i=1}^{\bar{N}} log_2(1 + \frac{SNR_i}{\Gamma}) \tag{2.52}$$

wherein SNR_i and Γ are the signal to noise ratio and coding gain of the i'th subchannel respectively. By factoring the *log* terms inside the sum expression, we can reformulate equation 2.52 into the form

$$b_{DMT} = \bar{N} log_2(1 + \frac{SNR_{geom}}{\Gamma}) \tag{2.53}$$

where the term SNR_{geom} is the geometrical signal to noise ratio defined by

$$SNR_{geom} = \Gamma(\left[\prod_{i=1}^{\bar{N}}(1 + \frac{SNR_i}{\Gamma})\right]^{\frac{1}{\bar{N}}} - 1) \tag{2.54}$$

To simplify the analytical derivations, it is assumed that all the subchannels contribute in the data transmission. In other words, all of the available bandwidth is assumed to be used by the MC system. Neglecting the "+1" and "-1" terms in the above expressions, the SNR_{geom} can be approximated by

$$SNR_{geom} = \left[\prod_{i=1}^{\bar{N}}(SNR_i)\right]^{\frac{1}{\bar{N}}} \tag{2.55}$$

The noise power in each subchannel can be decomposed into two terms. The first term is due to the residual error associated with the impulse response shortening problem (mean square error) . The remainder is due to the AWGN introduced by channel and receiver noise. Assuming the contribution of MSE resulted from impulse response shortening is negligible comparing to the AWGN term, the geometrical signal to noise ratio can be reformulated as

$$SNR_{geom} = S_x \left[\prod_{i=1}^{\bar{N}}(\frac{|B_i|^2}{S_n[i]|W_i|^2})\right]^{\frac{1}{\bar{N}}} \tag{2.56}$$

where $S_n[i]$, B_i and W_i are the noise power, DFT of filter $\mathbf{b}$ and DFT of filter $\mathbf{w}$ in i^{th} bin respectively. For fixed bandwidth,maximizing the achievable bit rate would be the same as maximizing the SNR_{geom}. Thus, the optimization problem can be formulated as

$$[\mathbf{b}, \mathbf{w}] = \arg\max_{\mathbf{b},\mathbf{w}} \left[\log \prod_{i=1}^{\bar{N}}(\frac{|B_i|^2}{S_n[i]|W_i|^2})\right] \tag{2.57}$$

To reduce the complexity of consecutive stages as follows

Step 1 The optimum solution for TIR ise obtained from solving the following constraint optimization problem

$$\mathbf{b} = \arg\max_{\mathbf{b}} \sum_{i=1}^{\bar{N}} \log \mathbf{b}^* \mathbf{G}^i \mathbf{b} \tag{2.58}$$

where $\mathbf{G}^i \triangleq \mathbf{g}^i \mathbf{g}^{i*}$, and $\mathbf{g}^i$ is the i^{th} Fourier basis vector given by

$$\mathbf{g}^i \triangleq \left[\begin{array}{cccc} 1 & e^{-j\frac{2\pi i}{N}} & \cdots & e^{-j\frac{2\pi i N_b}{N}} \end{array} \right]^*$$

Step 2 The Wiener-Hopf equation given in 2.50 is used to obtain the optimum setting of the TEQ.

The above optimization problem was originally proposed in [27]. In this paper, authors used sequential quadratic programming (SQP) method in order to obtain the optimum setting of the TEQ. However, no specific iterative method for equalizer training was proposed for this problem. Motivated by this shortcoming, a new efficient iterative approach for the above constraint optimization problem is developed in Chapter 4.

2.3 Overview of Synchronization Methods

The principle disadvantage of MC system is in its sensitivity to improper synchronization environment introduced by nonideal receiver. More specifically, due to improper sampling of the received signal, MC symbol is subject to the timing delay. Moreover, an offset in the carrier frequency of demodulator at the receiving end may cause loss of orthogonality among subcarriers. These deleterious effects puts an impediment in realization of MC systems for practical purposes.

Synchronization, is the process of adapting the receiver to the symbol timing error and carrier frequency offset introduced by the channel delay and improper sampling operation at the receiver. The effect of frequency offset on the performance of MC systems was originally analyzed by Moose [32]. It was shown in this paper that the signal to noise

ratio at the output of MC receiver is lower-bounded by

$$SNR \geq \frac{E_c}{N_0} \frac{\{\sin \pi\varepsilon/(\pi\varepsilon)\}^2}{1 + .5947(E_c/N_0)(sin\pi\epsilon)^2} \tag{2.59}$$

where E_c, N_n and ϵ are the signal power and noise power and carrier frequency offset respectively. Pollet et. al. analyzed the sensitivity of MC system to carrier frequency offset and provided an expression for computing the probability of error [33]. Eduardo et. al. expanded this concept to OFDM communication and derived the BER for binary OFDM systems.

Categorically, the previously proposed methods for synchronization of multicarrier systems can be classified into two main subclasses, namely minimum mean square error (MMSE) and ML estimators. In MMSE approach, estimator uses the side information provided by the reference signal (pilot tones) in order to minimize a cost function associated with the synchronization parameters [34, 35]. A salient feature of this approach is that no probabilistic assumptions are made in regard to the data. Although MMSE estimators usually result in a tractable (globally stable) and easy to implement realization, no optimality criterion (probabilistic or statistical) is associated with these estimators. Also, since part of the transmitted information is allocated to the reference pilots, the bandwidth efficiency of these methods is lower in comparison to the nonpilot schemes.

On the other hand, ML estimators provide the estimate of unknown parameter subject to the minimum probability of error criterion [36], [37], [32]. Although not exactly efficient, ML estimators are asymptotically minimum variance unbiase (MVU), i.e. their variance attains that of MVU estimator as length of the data record goes to infinity. However, due to the physical constraints, systems with infinitely long data records are not feasible for implementation purposes.

In [32], authors use retransmission technique in order to reveal the frequency offset parameter in the likelihood function of the received signal. Due to the redundancy introduced by repeating the data block, the spectral efficiency is dropped by a factor of two. To avoid this imperfection, a new ML estimator based on cyclic prefix (CP) was introduced in [37]. In this approach, the side information provided by the CP is used to

obtain the likelihood function for joint estimation of symbol timing error and frequency offset in MC systems.

Our research revealed that the likelihood function proposed in [37] does not globally characterize the observation vector over the entire range of the timing offset. Consequently, the ML estimator proposed based on this likelihood function would result in considerable performance loss over a finite range of timing offset interval.

Motivated by the suboptimum performance of this estimator, a new likelihood function for joint estimation of carrier frequency offset and symbol timing error of multicarrier systems is introduced in Chapter 5. Based on this result, a new optimum ML estimator for the joint estimation problem is also presented. In an attempt to reduce the variance of ML estimator, we also investigate a new class of MVU estimators for frequency offset estimation of multicarrier systems. It is shown that there exists but one function of sufficient statistic which provides the MVU estimate of the frequency offset. The proposed estimator provides a closed form expression as a function of data statistic. Consequently, the new method does not suffer from converging to multiple local minima, the problem which arises in ML technique with nonconvex loglikelihood function [36]. Synchronization of multicarrier system is thoroughly addressed in Chapter 5.

3. FAST ALGORITHM FOR FINITE LENGTH MMSE EQUALIZERS WITH APPLICATION TO DISCRETE MULTITONE SYSTEMS

3.1 Summary

This section presents a *new*, fast algorithm for finite-length minimum mean square error (MMSE) equalizers. The research exploits asymptotic equivalence of Toeplitz and circulant matrices to estimate Hessian matrix of a quadratic form. Research shows that the Hessian matrix exhibits a specific structure. As a result, when combined with the Rayleigh minimization algorithm, it provides an efficient method to obtain the global minimum of constrained optimization problem. A salient feature of this algorithm is that extreme eigenvector of the Hessian matrix can be obtained without direct computation of the matrix. In comparison to the previous methods, the algorithm is more computationally efficient and highly parallelizable, which makes the algorithm more attractive for real time applications. The algorithm is applied for equalization of discrete multitone (DMT) systems for asynchronous digital subscriber line (ADSL) applications.

3.2 Introduction

In design of adaptive filters for signal processing applications, various optimality criteria can be used to obtain the optimum setting of the adaptive filter. However, MMSE is considered to be the most tractable technique which guarantees existence and uniqueness of global optimum solution. The problem of finite-length MMSE filtering has already been investigated in many literatures [27],[31]. In [31], author applies the notion of MMSE filtering for system identification problems. Partial equalization of spectrally shaped channels is another fruitful application of MMSE filtering in communication and signal process-

ing. Specifically, given a highly dispersive channel of length v, the objective is to design a finite-length time domain equalizer (TEQ) to force the effective channel into a much shorter filter known as target impulse response (TIR). In general, the optimum solution to this problem is obtained from computing the global minimum of a quadratic function. Due to the inherent potential of quadratic forms to converge to the trivial solution, an energy boosting constraint is applied to the problem. Among the feasible constraint sets, unit energy constraint (UEC) and unit tap constraint (UTC) have found more applications in communication systems. In principle, decision feedback equalization (DFE) can be categorized as a special class of MMSE equalizers under UTC. A fast algorithm for MMSE equalizers has already been proposed in [28]. However, the study conducted in [28] provided the optimum solution subject to UTC. In this paper, a new fast iterative algorithm for computing optimum setting of MMSE-UEC equalizers is presented. Also note that equalization under UEC provides better SNR in comparison to UTC [27]. The method makes use of asymptotic equivalence of circulant and Toeplitz matrices to obtain a closed form expression for the Hessian matrix. Additionally, we show that any quadratic form can be computed efficiently using the discrete Fourier transform (DFT) operation. When combined with the Rayleigh minimization algorithm, it provides a fast algorithm for computing coefficients of TIR and TEQ . The algorithm provides the solution after $N_b + 1$ iterations and requires $O(N_f log_2(N_f))$ operations/iteration where N_f and $N_b + 1$ are the length of TEQ and TIR, respectively. The rest of this paper is organized as follows. In section 3.3 an overview of MMSE approach is presented. In section 4.3 few properties of Hessian matrix are derived. Based on these derivations, a new iterative algorithm for MMSE-UEC is proposed. The complexity of algorithm is compared against the standard matrix inversion method. Finally, in section 3.6 the algorithm is applied to impulse response shortening of DMT systems.

3.3 MMSE equalization

This section presents an overview of the MMSE equalization problem.[1] Block diagram of the equalizer studied in this paper is depicted in figure 3.1. The channel response is modeled as a discrete time FIR filter, expressed by $\mathbf{h} = \{\mathbf{h}[0], \mathbf{h}[1], \cdots, \mathbf{h}[v]\}$ where v is the channel spread. The channel response represents the combined effect of the transmit and receive filters as well as the channel impulse response. Input is an independent identically distributed random sequence with power of σ_x^2. In MMSE approach, equalizer taps are set such that the residual error between output of TIR and TEQ filters is minimized in the mean square sense. MMSE equalization can be viewed as a quadratic optimization problem in which the optimum settings for the TIR and TEQ filters are obtained from the following equations

$$\mathbf{b}_{opt} = \arg \min_{\mathbf{b}} \mathbf{b}^* \mathbf{R}_\Delta \mathbf{b} \tag{3.1}$$

$$\mathbf{w} = \sigma_x^2 {\mathbf{R}_{yy}}^{-1} \mathbf{H}_\Delta \mathbf{b}_{opt} \tag{3.2}$$

where $\mathbf{R}_\Delta$ is the Hessian matrix given in

$$\mathbf{R}_\Delta \triangleq \sigma_x^2 \mathbf{I}_{N_b+1} - \sigma_x^4 \mathbf{H}_\Delta^* \mathbf{R}_{yy}^{-1} \mathbf{H}_\Delta \tag{3.3}$$

and

$$\mathbf{R}_{yy} = \sigma_x^2 \mathbf{H}\mathbf{H}^* + \mathbf{R}_{nn} \tag{3.4}$$

$$\mathbf{H}_\Delta = \mathbf{H}\mathbf{S}^* \tag{3.5}$$

$$\mathbf{H} = \begin{bmatrix} h[0] & h[1] & \cdots & h[v] & \cdots & 0 \\ 0 & h[0] & h[1] & \cdots & h[v] & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & h[0] & h[1] & \cdots & h[v] \end{bmatrix} \tag{3.6}$$

$$\mathbf{S} = \begin{bmatrix} \mathbf{0}_{N_b+1,\Delta} & \mathbf{I}_{N_b+1,N_b+1} & \mathbf{0}_{N_b+1,\kappa} \end{bmatrix} \tag{3.7}$$

[1] Throughout the paper, symbols $\odot$, $*$, $\Re$ and $\tilde{\cdot}$, represent element by element vector multiplication, linear convolution, real and Fourier transform operations respectively. Also matrices and vectors are represented by upper-case and lower-case bold characters, respectively.

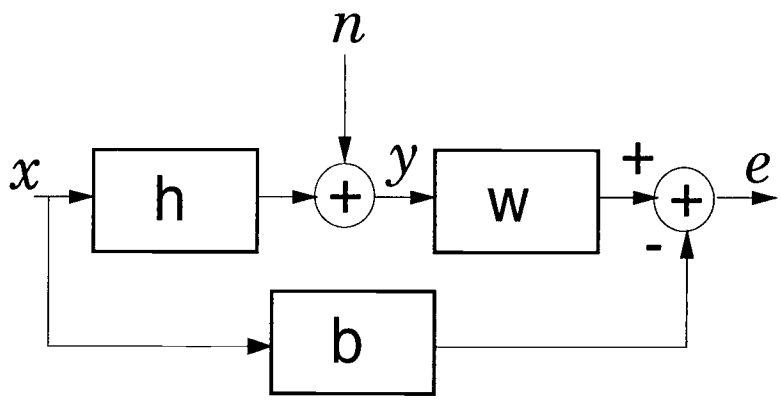


Figure 3.1: Block Diagram of MMSE Equalizer

In the above equations, Δ is the decision delay involved with the TIR and $\kappa \triangleq N_f + v - N_b - 1$. Also matrices $\mathbf{I}$,$\mathbf{0}$ and $\mathbf{R}_{nn}$ represent the identity,zero and noise autocorrelation matrices, respectively. Recall that in the presence of UEC, the optimization problem given in (4.1) can be viewed as minimization of Rayleigh quotient of matrix $\mathbf{H}_\Delta$. With this idea in mind, we apply the iterative algorithm proposed in [38] to obtain extreme eigenvalue (vector) of the Hessian matrix. Basically, this algorithm applies the conjugate-gradient method to find the minimum of second order approximation of Rayleigh quotient. Rayleigh minimization algorithm is widely used in subsapce tracking problems [39]. This method, although iterative in nature, is guaranteed to converge after at most $N_b + 1$ iterations. Robustness to matrix condition number is another noticeable feature of this algorithm [40]. Although the Rayleigh minimization algorithm does not exploit matrix inversion , it requires frequent computation of quadratic forms. In the next section we describe an efficient method for computing the quadratic terms, which eventually leads us to a fast algorithm for computing coefficients of MMSE-UEC equalizers.

3.4 Iterative Algorithm

The first step in reducing the complexity of the algorithm is to provide an efficient method for computing the inverse of the autocorrelation matrix. In doing so, we approximate the Toeplitz matrix

$$\mathbf{R}_{yy} = Tpltz(r_{yy}[0], r_{yy}[1], \quad \cdots \quad , r_{yy}[N_f - 1]) \tag{3.8}$$

with its asymptotic equivalent, known as circulant matrix [41]

$$\mathbf{C}_{yy} = Tpltz(r_{yy}[0], r_{yy}[1], \cdots, r_{yy}[N_f/2], \tag{3.9}$$

$$r_{yy}[N_f/2 - 1], \cdots, r_{yy}[1]) \tag{3.10}$$

Note that the argument of the $Tpltz$ operator is the first row of the Toeplitz matrix. It is known that the matrices $\mathbf{R}_{yy}$ and $\mathbf{C}_{yy}$ defined in 3.8 and 3.9 satisfy the condition

$$\lim_{N_f \to \infty} |\mathbf{R}_{yy} - \mathbf{C}_{yy}| = 0 \tag{3.11}$$

where the weak norm is defined as [42]

$$|\mathbf{A}| \triangleq (n^{-1} trace[\mathbf{A}^*\mathbf{A}])^{\frac{1}{2}} = (n^{-1} \sum_{k=0}^{N_f-1} \sum_{j=0}^{N_f-1} |a_{k,j}|^2) \tag{3.12}$$

A distinct advantage of circulant matrix is that it can be decomposed into the product of Fourier matrices and a diagonal matrix as given by

$$\mathbf{C}_{yy} = \mathbf{U}_{N_f} \, \Psi \, \mathbf{U}_{N_f}^* \tag{3.13}$$

where

$$\Psi = diag(\psi[0] \quad \psi[1] \quad \cdots \quad \psi[N_f - 1]) \tag{3.14}$$

$$\mathbf{U}_{N_f}[k,l] = \frac{1}{\sqrt{N_f}} e^{-j\frac{2\pi kl}{N_f}} \quad k,l = 0,1,...N_f - 1 \tag{3.15}$$

$$\psi[k] = \sum_{m=0}^{N_f-1} \mathbf{c}[m] e^{-j\frac{2\pi km}{N_f}} \quad k = 0,1,...N_f - 1 \tag{3.16}$$

and $\mathbf{c}[j]$ is the j'th element of the first row of matrix $\mathbf{C}_{yy}$. Using orthogonal properties of Fourier matrices, we can estimate the Hessian matrix as

$$\mathbf{R}_\Delta = {\sigma_x}^2 \mathbf{I}_{N_b+1} - {\sigma_x}^4 \mathbf{H}_\Delta^* \mathbf{U}_{N_f}^* \, \mathbf{\Psi}^{-1} \, \mathbf{U}_{N_f} \, \mathbf{H}_\Delta \tag{3.17}$$

This closed form expression appears to attain fruitful properties as we will illustrate shortly.

Property 1:

Given an arbitrary pair of vectors $\mathbf{p}$ and $\mathbf{q}$ of length $N_b + 1$, the quadratic term $\mathbf{p}^*\mathbf{R}_\Delta\mathbf{q}$ can be expressed as

$$\mathbf{p}^*\mathbf{R}_\Delta\mathbf{q} = \mathbf{p}^*({\sigma_x}^2 \mathbf{I}_{N_b+1} - {\sigma_x}^4 \mathbf{H}_\Delta^* \mathbf{U}^* \, \mathbf{\Psi}^{-1} \, \mathbf{U} \, \mathbf{H}_\Delta)\mathbf{q} \tag{3.18}$$

Define a dummy vector $\mathbf{c} \triangleq {\sigma_x}^2 \mathbf{H}_\Delta q$ as

$$\begin{bmatrix} \mathbf{c}[0] \\ \mathbf{c}[1] \\ \vdots \\ \mathbf{c}[N_f-1] \end{bmatrix} = \begin{bmatrix} h[\Delta] & h[\Delta+1] & \cdots & h[\Delta+N_f-1] \\ h[\Delta-1] & h[\Delta] & \cdots & h[\Delta+N_f-2] \\ \vdots & \ddots & \ddots & \ddots \\ h[\Delta-N_f+1] & \cdots & h[\Delta-1] & h[\Delta] \end{bmatrix} \begin{bmatrix} \mathbf{q}[0] \\ \mathbf{q}[1] \\ \vdots \\ \mathbf{q}[N_f-1] \end{bmatrix} \tag{3.19}$$

Due to the circular property of matrix $\mathbf{H}_\Delta$ this vector can be written as linear convolution of two vectors as expressed by

$$\mathbf{c}\,[n] \;=\; {\sigma_x}^2 \sum_{l=0}^{N_f-1} \mathbf{h}\,[\Delta - n + l]\,\mathbf{q}\,[l] = {\sigma_x}^2 \mathbf{g}\,[n] * \mathbf{q}\,[n] \tag{3.20}$$

where $\mathbf{g}\,[n] \triangleq \mathbf{h}\,[-n+\Delta]$. The term ${\sigma_x}^2\mathbf{U}\mathbf{H}_\Delta\mathbf{q} = \tilde{\mathbf{c}}$ is simply the Fourier transform of vector $\mathbf{c}$ and can be computed efficiently as

$$\sigma_x^2 \mathbf{U}\mathbf{H}_\Delta\mathbf{q} = \tilde{\mathbf{c}} = \sigma_x^2 \tilde{\mathbf{g}} \odot \tilde{\mathbf{q}} \tag{3.21}$$

Note that in the above expression, we have assumed that the length of TEQ exceeds that of TIR filter ($N_f > N_b + 1$). In applications in which this constraint can not be tolerated, the long TEQ filter can be well approximated by a pole-zero filter with fewer coefficients

[43]. Applying equation (3.21) and Parseval's equality to equation (3.18) results in a closed form expression for the quadratic term $\mathbf{p}^*\mathbf{R}_\Delta\mathbf{q}$ as given by

$$\mathbf{p}^*\mathbf{R}_\Delta\mathbf{q} \;=\; \sum_{k=0}^{N_f-1} \tilde{\mathbf{z}}\,[k]\,\tilde{\mathbf{p}}^*\,[k]\,\tilde{\mathbf{q}}\,[k] \tag{3.22}$$

where the new vector $\tilde{\mathbf{z}}$ is defined as

$$\tilde{\mathbf{z}}[k] \triangleq {\sigma_x}^2 - {\sigma_x}^4 \frac{\tilde{\mathbf{g}}[k]\tilde{\mathbf{g}}^*[k]}{\psi[k]} \tag{3.23}$$

This closed form expression given in (3.22) suggests performing the Rayleigh minimization algorithm in the frequency domain. In doing so, we need to represent the Fourier transform of vector $\mathbf{R}_\Delta\mathbf{q}$ as a function of vector $\tilde{\mathbf{q}}$. Wishing to avoid performing the above operation in the time domain, we propose an efficient method which performs the above operation using DFT.

Property 2:

For a vector $s \triangleq \mathbf{R}_\Delta\mathbf{q}$, the i'th element can be represented as

$$s\,[i] = \mathbf{e}_i\mathbf{R}_\Delta\mathbf{q} = \frac{1}{\sqrt{N_f}} \sum_{k=0}^{N_f-1} \tilde{\mathbf{z}}\,[k]\,\tilde{\mathbf{q}}\,[k]\,e^{\,j\frac{2\pi ki}{N_f}} \tag{3.24}$$

where $\mathbf{e}_i$ is the i'th unit vector of length N_b+1. In deriving the above equation, we have used the closed form expression given in (3.22). Equation (3.24) appears to be the i'th element of IDFT of vector $\tilde{\mathbf{z}}\odot\tilde{\mathbf{q}}$. Hence, the vector $\mathbf{s}$ can be obtained from the first N_b+1 elements of the $IDFT(\tilde{\mathbf{z}}\odot\tilde{\mathbf{q}})$. Consequently, Fourier transform of the vector $\mathbf{R}_\Delta\mathbf{q}$ can be obtained by performing DFT operation on the vector $\mathbf{s}$. These two properties along with Parseval provide us an efficient algorithm as we will explain in the subsequent section.

3.5 Fast Algorithm

- Initialization:

  Starting from an arbitrary normalized vector $\tilde{\mathbf{b}}^0$, compute the minimum eigenvalue estimate, residual error and descent direction according to

$$\lambda^0 \;=\; \sum_{k=0}^{N_f-1} \tilde{\mathbf{z}}\,[k]\,\tilde{\mathbf{v}}^0[k] \tag{3.25}$$

$$\tilde{\mathbf{r}}^0 = \lambda^0\tilde{\mathbf{b}}^0 - \tilde{\mathbf{s}}^0 \quad \tilde{\mathbf{p}}^0 = \tilde{\mathbf{r}}^0 \quad \tilde{\mathbf{v}}^0[k] = |\tilde{\mathbf{b}}^0[k]|^2 \tag{3.26}$$

- Iteration:

  For $i = 0 \cdots N_b$ compute the TIR frequency response as

$$\tilde{\mathbf{b}}^{i+1} = \tilde{\mathbf{b}}^i + \mu^i\tilde{\mathbf{p}}^i \quad \mu^i = \frac{-B + \sqrt{B^2 - 4CD}}{2D} \tag{3.27}$$

$$D = \rho_b^i\rho_c^i - \rho_a^i\rho_d^i \ , \ B = \rho_b^i - \lambda^i\rho_d^i \ , \ C = \rho_a^i - \lambda^i\rho_c^i \tag{3.28}$$

$$\rho_a{}^i = \sum_{k=0}^{N_f-1} \tilde{\mathbf{z}}[k]\,\tilde{\mathbf{d}}_1[k] \quad , \quad \rho_b{}^i = \sum_{k=0}^{N_f-1} \tilde{\mathbf{z}}[k]\,\tilde{\mathbf{d}}_2[k] \tag{3.29}$$

$$\rho_c{}^i = \sum_{k=0}^{N_f-1} \tilde{\mathbf{d}}_1[k] \quad , \quad \rho_d{}^i = \sum_{k=0}^{N_f-1} \tilde{\mathbf{d}}_2[k] \tag{3.30}$$

$$\tilde{\mathbf{d}}_1[k] = (\tilde{\mathbf{p}}^i[k])^*\tilde{\mathbf{b}}^i[k] \quad , \quad \tilde{\mathbf{d}}_2[k] = (\tilde{\mathbf{p}}^i[k])^*\tilde{\mathbf{p}}^i[k] \tag{3.31}$$

Compute the minimum eigenvalue estimate, residual error, descent direction and normalized TIR vector according to

$$\lambda^{i+1} = \frac{1}{\tau^{i+1}} \sum_{k=0}^{N_f-1} \tilde{\mathbf{z}}[k]\,\tilde{\mathbf{v}}^i[k] \ , \ \tilde{\mathbf{r}}^{i+1} = \frac{\lambda^{i+1}\tilde{\mathbf{b}}^{i+1} - \tilde{\mathbf{s}}^{i+1}}{\tau^{i+1}} \tag{3.32}$$

$$\tilde{\mathbf{p}}^{i+1} = \tilde{\mathbf{r}}^{i+1} + \beta^i\tilde{\mathbf{p}}^i \quad , \quad \tilde{\mathbf{b}}^{i+1} = \frac{\tilde{\mathbf{b}}^{i+1}}{\tau^{i+1}} \tag{3.33}$$

where

$$\beta^i = \frac{\sum_{k=0}^{N_f-1} \tilde{\mathbf{z}}[k]\ (\tilde{\mathbf{r}}^{i+1}[k])^*\tilde{\mathbf{p}}^i[k] + (\|r^{i+1}\|^2)(\rho_c{}^i + \mu^i\rho_d{}^i)}{\rho_b{}^i - \lambda^{i+1}\,\rho_d{}^i} \tag{3.34}$$

$$\tilde{\mathbf{v}}^{i+1}[k] = \frac{\tilde{\mathbf{v}}^i[k]}{(\tau^i)^2} + (\mu^i)^2\tilde{\mathbf{d}}_2[k] + 2\mu^i\Re(\tilde{\mathbf{d}}_1[k]) \tag{3.35}$$

$$\tau^{i+1} = \textstyle\sum_{k=0}^{N_f-1}(\tilde{\mathbf{v}}^{i+1}[k]) \tag{3.36}$$

Upon computing the optimum setting for TIR, TEQ's coefficients are obtained from equation (4.2). It is also worthwhile to remark that the term $\sigma_x^2\mathbf{H}_\Delta\mathbf{b}$ in equation (4.2) can be computed efficiently using equation (3.21). Table 3.1 compares the computational complexity of the proposed method against the standard matrix inversion method.

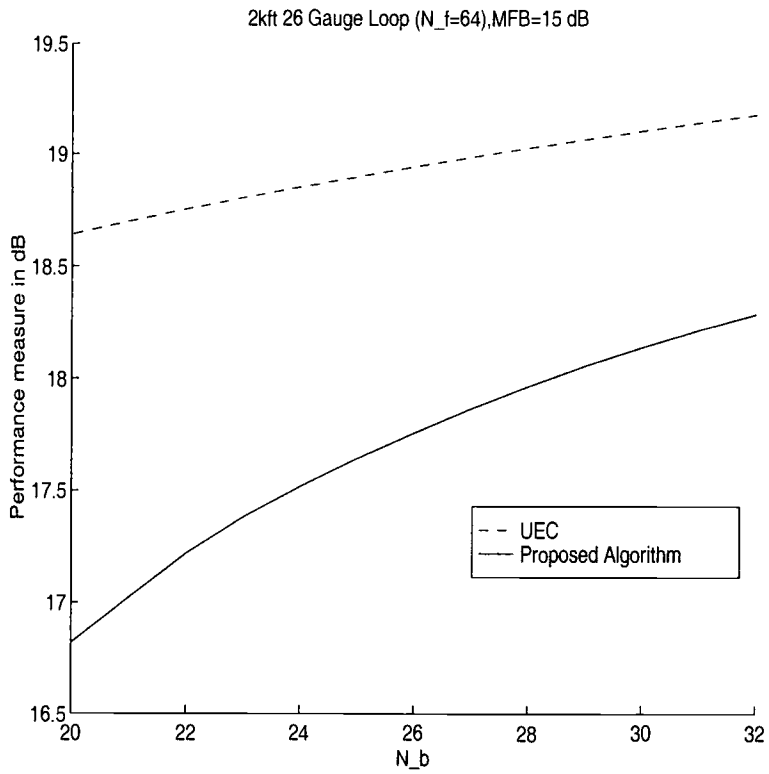


Figure 3.2: Performance with Different TIR Filter Lengths

3.6 Simulations and Performance Evaluation of the Algorithm

In this section we apply the proposed algorithm for equalization of DMT in ADSL environment. A series of simulations are performed on 2 kft, 26 guage (AWG) wire line sampled at 2.208 MHz. The power spectral density of near-end crosstalk (NEXT) noise is generated by exciting the NEXT coupling filter $|H_x(f)|^2 = k_{NEXT} f^{3/2}$ by a white Gaussian noise with power of 10mW. Unless specified, k_{NEXT} is fixed to 10^{-13}. Also there is an AWGN with power of -30dBm across the two sided spectral bandwidth. Decision delay is set to the optimum delay obtained from MMSE-UEC .Unless specified, transmit power is set such that the matched filter bound ($\mathbf{MFB} = ||h||^2\sigma_x^2/\sigma_n^2$) of 15 dB is achieved at the receiving point. As a performance measure, we compute signal power

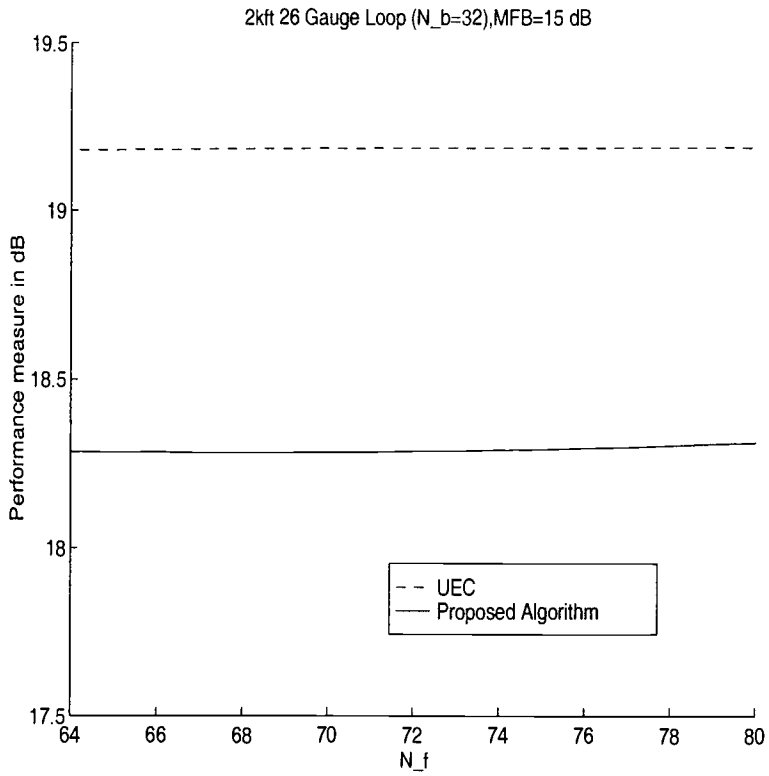


Figure 3.3: Performance with Different TEQ Lengths

to MSE ($SNR = \sigma_x^2/\mathbf{b}^* R_\Delta \mathbf{b}$) to evaluate the performance of the methods. Throughout the simulations, performance of the proposed method is compared against MMSE-UEC method. The following points can be inferred from the plots.

- The gap between exact solution and proposed algorithm reduces as the length of the TIR filter increases (Figure 3.2). This is due to the fact that the Rayleigh minimization algorithm provides more exact solutions as dimension of the Hessian matrix increases.

- As long as N_f is large enough to satisfy the asymptotic equivalence of Toeplitz and circulant matrices, the proposed algorithm provides a robust solution for various values of N_f.

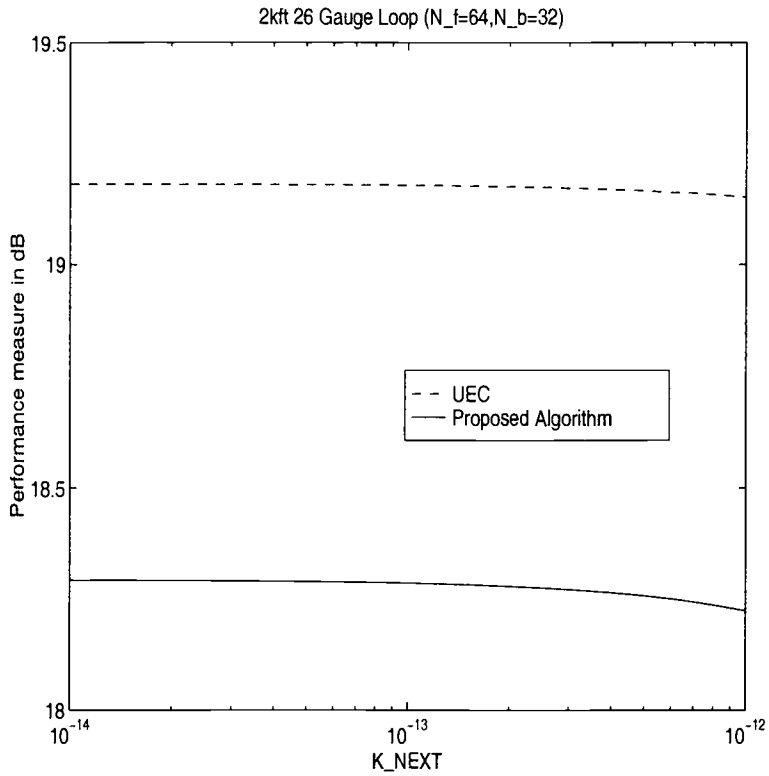


Figure 3.4: Performance as a Function of K_{NEXT}

- Performance of the algorithm is not influenced by the spectrum of the noise. As is shown in Figure (3.4) Signal to MSE is constant over a large range of K_{NEXT}.
- Signal to MSE is a linear function of signal power (MFB). This is a favorable characteristic, as there would be no limitation on the dynamic range of transmit power.

3.7 Concluding Remarks

We have developed a novel fast algorithm as a straightforward application of Rayleigh minimization approach to solving the optimum MMSE-UEC equalization problem. Its structure was chosen to allow the use of the DFT operation which makes the algorithm highly parallelizable. The proposed method can be customized to provide a balance be-

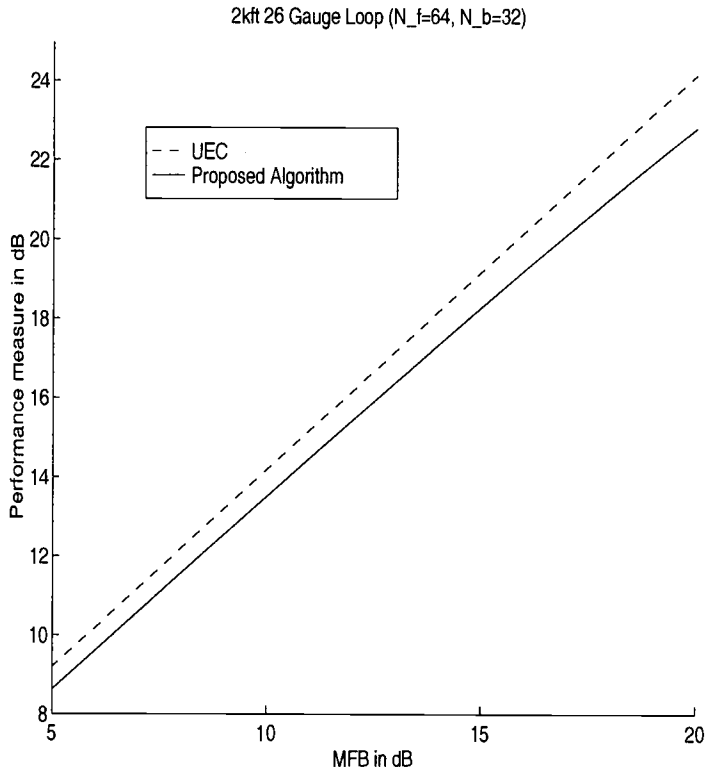


Figure 3.5: Performance as a Function of MFB

tween performance and computational complexity. Simulation results in this paper show that the numerical complexity in the minimum eigenvector estimation can be reduced considerably by exploiting the proposed algorithm, without significant loss in the performance.

	Power iteration	Proposed algorithm
Hessian matrix Computation	$N_f^2(N_b+1)$ $+N_f(N_b+1)^2$	—
Min. eigen. Computation	$O((N_b+1)^3)$	$O(N_f log_2 N_f)$ per iteration
Sensitivity to condition No.	Highly sensitive	Robust
Other Features	Requires matrix inversion	Parallelizable, requires large N_f

Table 3. : Comparison Between Proposed Method and Standard Power iteration Algorithm

4. OPTIMUM EQUALIZATION OF MULTICARRIER SYSTEMS: A UNIFIED GEOMETRIC APPROACH

This section presents a *new*, iterative equalization algorithm that maximizes capacity for discrete multitone (DMT) systems. The research modifies a previously proposed criterion and applies an appropriate transformation to map the constraint set into a proper region. The resulting constraint set exhibits an identifiable geometric characteristic, which provides an efficient method for obtaining the optimal solution. Using the gradient projection method in conjunction with projection onto convex sets (POCS) provides us with an iterative search algorithm which facilitate the search direction. We also generalize the approach to two important subclasses of equalizers, namely linear phase and unit tap filters. An fundamental limit on the performance of the approach is also derived. In comparison with the previous methods, the proposed equalization algorithm is less computationally complex, more robust, and geometrically intuitive. Simulation experiments confirm the validity of the proposed method for practical purposes.

4.1 Introduction

Since the introduction of channel capacity by Shannon, there has been considerable interest to maximize the bit rate through the communication channels. Multicarrier transmission systems exploit several parallel quadrature amplitude modulation (QAM) blocks to transmit data reliably over highly dispersive channels. As a subclass of multicarrier systems, discrete multitone (DMT) systems provide an efficient method for partitioning the channel into a set on nonoverlapping orthogonal subchannels. DMT systems generate the transmit sequence by performing a discrete Fourier transform (DFT) operation on a block of data. Prior to sending the data on the channel, a portion of generated sequence, known as cyclic prefix (CP), is prepended to the modulated symbol. The cyclic prefix makes the channel-description matrix circulant, thus the orthogonal set of Fourier basis

vectors can be applied to find its associated eigenvalues [44]. The cyclic prefix introduces redundancy in data transmission, therefore reduces the effective data rate of a digital transmission system.

A short CP would reduce the performance loss of data transmission and is favorable in many real time applications. However the length of the cyclic prefix is lower bounded by the effective length of the channel [45]. In many practical channels, such as digital subscriber loops, the effective length of the channel is large, which results in a considerable performance loss due to adding the cyclic prefix. The solution is to force a long channel impulse response to a short filter, which reduces the performance loss introduced by adding the cyclic prefix. At the receiving end, the DMT system uses a finite impulse response (FIR) filter ($\mathbf{w}$), known as a time domain equalizer (TEQ). TEQ forces the effective channel to a much shorter filter, known as the target impulse response (TIR). In setting the coefficients of TEQ, several criteria have been considered and investigated. Chow *et al.* provided an adaptive LMS algorithm for setting the coefficients of the TEQ [46]. Although it is simple in structure, the algorithm is not robust and globally optimum. Following this work, Al-Dhahir and Cioffi proposed a unified robust method that provides the optimum solution of the impulse response shortening problem based on minimum mean square error (MMSE) criterion [27]. Later, in a comprehensive study performed by the same authors, it was found that the solution obtained from MMSE approach may not necessarily optimize the performance (capacity or margin) of the DMT system. Based on this observation, a new objective function was defined which outperformed the setting obtained from the MMSE approach. Nevertheless, no specific algorithm was proposed to obtain the optimal setting of the equalizer. This chapter presents a new iterative algorithm for obtaining the optimum setting of the TEQ. The research makes use of the gradient projection method to obtain the descent direction for the gradient method. As a result, when combined with the POCS technique, the stationary point obtained from the algorithm converges to an optimum point. POCS is a powerful technique which has found a wide spread applications in set theoretic signal processing algorithms [47]. The rest of this chapter is organized as follows. Section 4.2 presents an overview of equalizer training approaches for

DMT systems. In section 4.3 a new algorithm for training the DMT equalizer based on maximum capacity criterion is presented. Section 4.4 addresses the unit tap constraint on the optimization problem. Optimization under linear phase constraint is thoroughly investigated in section 4.5. An upper-bound on the performance of the system is derived in section 4.6. Finally in section 4.7 the algorithm is applied to equalization of DMT systems in order to assess the performance of proposed method for asynchronous digital subscriber line (ADSL) applications [4].

4.2 State of the Art

This section presents an overview of various time domain equalization methods for DMT systems as shown in figure 4.1. Throughout this paper, the symbols t , $*$ and $\tilde{.}$, represent transpose, Hermitian transpose and Fourier transform operations, respectively. Matrices and vectors are represented by upper-case and lower-case bold characters, respectively. The channel response is modeled as a discrete time finite impulse response (FIR) filter, expressed by $h = \{h_0, h_1, \cdots, h_v\}$ where v is the channel spread. The channel response represents the combined effect of the transmit and receive filters as well as the channel impulse response. The input signal x, is an independent identically distributed random sequence with power of σ_x^2. As explained earlier, several objective functions can be used to optimize the performance of the TEQ. Among the existing methods, MMSE is known to be the most tractable technique for impulse response shortening problem [27]. Several methods based on this approach have been proposed [29]. In this approach, the optimum equalizer taps are computed to minimize the mean square error between output of the TIR and TEQ filters. MMSE equalization can be viewed as a quadratic optimization problem in which the optimum settings for the TIR and TEQ filters are obtained from the following equations

$$\mathbf{b} = \arg \min_{\mathbf{b}} \mathbf{b}^* R_\Delta \mathbf{b} \tag{4.1}$$

$$\mathbf{w} = \sigma_x^2 \mathbf{R}_{yy}{}^{-1} \mathbf{H}_\Delta \mathbf{b} \tag{4.2}$$

where $\mathbf{R}_\Delta$ is the Hessian matrix given in

$$\mathbf{R}_\Delta \triangleq \sigma_x^2 \mathbf{I}_{N_b+1} - \sigma_x^4 \mathbf{H}_\Delta^* \mathbf{R}_{yy}^{-1} \mathbf{H}_\Delta \tag{4.3}$$

and

$$\mathbf{R}_{yy} = \sigma_x^2 \mathbf{H}\mathbf{H}^* + \mathbf{R}_{nn} \tag{4.4}$$

$$\mathbf{H}_\Delta \triangleq \mathbf{H}\,\mathbf{S}^*$$

$$\mathbf{H} = \begin{bmatrix} h_0 & h_1 & \cdots & h_v & \cdots & 0 \\ 0 & h_0 & h_1 & \cdots & h_v & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & h_0 & h_1 & \cdots & h_v \end{bmatrix}$$

$$\mathbf{S} = \begin{bmatrix} \mathbf{0}_{N_b+1,\Delta} & \mathbf{I}_{N_b+1,N_b+1} & \mathbf{0}_{N_b+1,s} \end{bmatrix}$$

in which N_f and $N_b + 1$ are the length of TEQ and TIR filters, respectively. Also, Δ is the decision delay involved with the TIR and $s \triangleq N_f + v - N_b - 1$. Matrices $\mathbf{I}$, $\mathbf{0}$ and $\mathbf{R}_{nn}$ represent the identity, zero, and noise autocorrelation matrices, respectively.

To avoid converging to the trivial solution, further constraint is imposed on the optimization problem (4.1). The Unit energy constraint (UEC) requires the norm of TIR filter to be equal to one ($\mathbf{b}^*\mathbf{b} = 1$), and the unit tap constraint (UTC) forces one of the taps in the TIR to be unity ($\mathbf{b}[k] = 1 \mid k \in \{0, 1, \cdots, N_b\}$). Further investigations on optimizing the performance of DMT systems determined that the equalizer setting obtained by using the MMSE criterion would not necessarily result in the best geometrical signal to noise (SNR_{geom}) ratio

$$SNR_{geom} \triangleq \left[\prod_{i=1}^{\bar{N}} (1 + SNR_i) \right]^{\frac{1}{\bar{N}}} \tag{4.5}$$

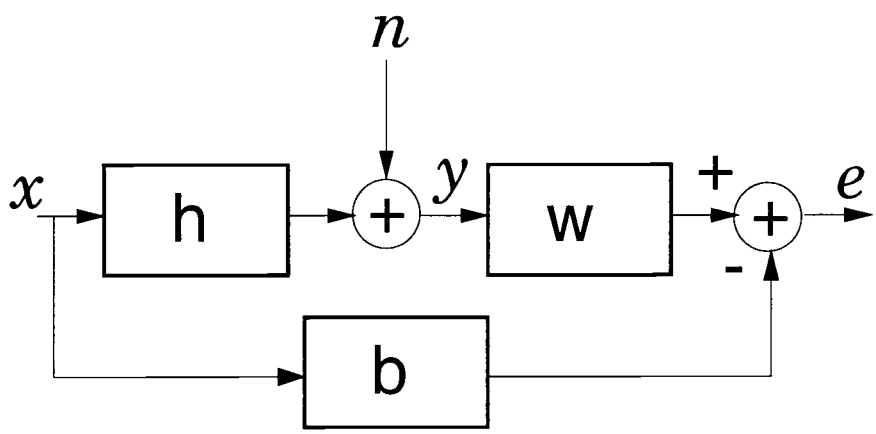


Figure 4.1: Block diagram of MMSE equalizer

for DMT systems, where $\bar{N}$ is the number of effective used subchannels [8]. A new criterion for setting the coefficients of the TEQ equalizer to maximize the (SNR_{geom}) was proposed in [7] ,[8]. According to this criterion, the coefficients of the equalizer are obtained to maximize the number of bits transmitted per DMT symbol, as expressed by

$$Bits_{DMT} = \sum_{i=1}^{\bar{N}} log_2(1 + \frac{SNR_i}{\Gamma_i}) \tag{4.6}$$

where SNR_i and Γ_i are the signal to noise ratio and the coding gain for the i'th subchannel respectively. Throughout our analysis it is assumed that Γ_i is constant over the entire subchannels. Optimization under unequal coding gain has been investigated in [....].

Following the approach proposed in [7], it is straightforward to show that the optimum setting for TIR filter, which results in the maximum capacity criterion, is found by solving a dual constrained optimization problem as expressed by

$$\mathbf{b}_{opt} = \arg \max_{\mathbf{b}} \sum_{i=1}^{N} log_2 \mathbf{b}^* \mathbf{G}^i \mathbf{b} \tag{4.7}$$

$$s.t. \quad \mathbf{C}_1 : \quad \mathbf{b}^* \mathbf{R}_\Delta \mathbf{b} \leq \epsilon^2 \tag{4.8}$$

$$\mathbf{C}_2 : \quad \mathbf{b}^* \mathbf{b} = 1 \tag{4.9}$$

In equation (4.7), matrix $\mathbf{G}^i$ is defined as

$$\mathbf{G}^i \triangleq \mathbf{g}^i \mathbf{g}^{i*}$$

where $\mathbf{g}^i$ is the i^{th} Fourier basis vector given by

$$\mathbf{g}^i \triangleq \left[\begin{array}{cccc} 1 & e^{-j\frac{2\pi i}{N}} & \cdots & e^{-j\frac{2\pi i N_b}{N}} \end{array} \right]^*$$

After obtaining the optimum setting for TIR, the optimum setting for TEQ is obtained by solving the Wiener-Hopf equation (4.2). The optimization problem given in (4.7) does not have a closed form solution. In [7] and [8], authors use standard optimization software [48] tools in order to solve the above optimization problem .In the next section, we will present a new iterative gradient search algorithm, for obtaining the optimum solution of the problem given in 4.7.

4.3 Proposed Iterative Gradient Search Algorithm

As explained in the previous section, the optimum equalization of DMT can be obtained by solving the constrained optimization problem given in 4.7. As depicted in figure 4.2, the constraint set in problem (4.7) is the intersection of two regions. The first region $\mathbf{C}_1 : \quad \{\mathbf{b} \in R^{N_b+1} \mid \mathbf{b}^*\mathbf{R}_\Delta\mathbf{b} \leq \epsilon^2\}$ represents a closed set on the Euclidean space R^{N_b+1}. Geometrically, the set $\mathbf{C}_1$ represents an ellipsoid in R^{N_b+1}. Because of the positive definite property of the matrix $\mathbf{R}_\Delta$, this constraint set forms a convex set on the Hilbert space. However, the unit energy constraint, $\mathbf{C}_2 : \quad \{\mathbf{b} \in R^{N_b+1} \mid \mathbf{b}^*\mathbf{b} = 1\}$, represents a region on the surface of a unit radial sphere that lacks convexity. In order to exploit the potential advantage of POCS, we remove the UEC from the constraint set. Unlike the MMSE approach, we can remove the UEC from the constraint set because origin is not among the local minimums (maximums) of the objective function and no energy boosting constraint is needed in order to avoid converging to the trivial solution. However, upon obtaining the global minimum, the solution vector can be normalized in order to satisfy the UEC. This scaling would not affect the geometrical signal-to-noise

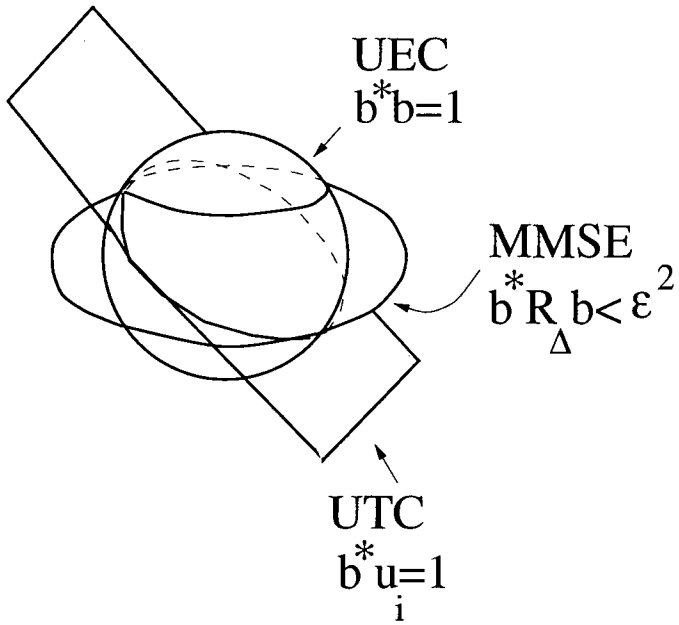


Figure 4.2: Geometrical Representation of the Algorithm

(SNR_{geom}) profile, as the TEQ coefficients would be scaled accordingly. Consequently, the mean square error and the additive noise contribution would be scaled by the same factor. Using the convexity property of the constraint set $\mathbf{C}_1$, along with a suitable iterative descent algorithm can lead us to a stationary point. We considered the gradient projection method in order to find the feasible direction at each iteration.

Due to the positive definite property of matrix $\mathbf{R}_\Delta$, any $N_b + 1$-dimensional vector $\mathbf{b}$ can be represented as linear combination of $N_b + 1$ orthonormal eigenvectors of matrix $\mathbf{R}_\Delta$ given by

$$\mathbf{b} = \alpha_0 \mathbf{v}_0 + \alpha_1 \mathbf{v}_1 + \cdots + \alpha_{N_b} \mathbf{v}_{N_b} \qquad \alpha_0, \alpha_1, \cdots, \alpha_{N_b} \in R \tag{4.10}$$

where $\mathbf{v}_m$ and λ_m are the m^{th} normalized eigenvector and associated eigenvalue of matrix $\mathbf{R}_\Delta$ which satisfy

$$\mathbf{R}_\Delta \mathbf{v}_i = \lambda_i \mathbf{v}_i \quad \text{and} \quad \mathbf{v}_i^* \mathbf{v}_j = \delta[i-j]$$

By substituting equation (4.10) into equation (4.7) the objective function can be expressed

as

$$\begin{aligned} f &= \sum_{i=1}^{N} log_2((\sum_{j=0}^{N_b} \alpha_j \mathbf{v}_j)^* \mathbf{G}^i (\sum_{j=0}^{N_b} \alpha_j \mathbf{v}_j)) \\ &= \sum_{i=1}^{N} log_2[\alpha_0 \;\; \alpha_1 \;\; \cdots \alpha_{N_b}] \mathbf{Q}^i [\alpha_0 \;\; \alpha_1 \;\; \cdots \alpha_{N_b}]^* \end{aligned}$$

where the new matrix $\mathbf{Q}^i$ is defined as

$$\mathbf{Q}^i \triangleq \begin{bmatrix} \mathbf{v}_0^* \mathbf{G}^i \mathbf{v}_0 & \mathbf{v}_0^* \mathbf{G}^i \mathbf{v}_1 & \cdots & \mathbf{v}_0^* \mathbf{G}^i \mathbf{v}_{N_b} \\ \mathbf{v}_1^* \mathbf{G}^i \mathbf{v}_0 & \mathbf{v}_1^* \mathbf{G}^i \mathbf{v}_1 & \cdots & \mathbf{v}_1^* \mathbf{G}^i \mathbf{v}_{N_b} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{v}_{N_b}^* \mathbf{G}^i \mathbf{v}_0 & \mathbf{v}_{N_b}^* \mathbf{G}^i \mathbf{v}_1 & \cdots & \mathbf{v}_{N_b}^* \mathbf{G}^i \mathbf{v}_{N_b} \end{bmatrix}$$

Due to the properties of matrix $\mathbf{G}^i$, its $(m,n)^{th}$ entry can be computed efficiently as

$$\mathbf{Q}^i[m,n] = \tilde{\mathbf{v}}_m^*[i] \tilde{\mathbf{v}}_n[i] \tag{4.11}$$

By virtue of equation (4.10), the constraint set $\mathbf{C}_1$ can be written as

$$\begin{aligned} \mathbf{b}^* \mathbf{R}_\Delta \mathbf{b} &= (\sum_{j=0}^{N_b} \alpha_j \mathbf{v}_j)^* \mathbf{R}_\Delta (\sum_{j=0}^{N_b} \alpha_j \mathbf{v}_j) \\ &= \lambda_0 {\alpha_0}^2 + \lambda_1 {\alpha_1}^2 + \cdots + \lambda_{N_b} {\alpha_{N_b}}^2 \leq \epsilon^2 \end{aligned}$$

Using this transformation, the optimization problem given in (4.7) subject to the constraint set $\mathbf{C}_1$ can be written as

$$\begin{aligned} \bar{\alpha}_{opt} &= \arg \min_{\bar{\alpha}} \sum_{i=1}^{N} log_2 \frac{1}{\bar{\alpha}^* \mathbf{Q}^i \bar{\alpha}} \\ \text{s.t } \mathbf{C}_1 : \;& {\alpha_0}^2 \lambda_0 + {\alpha_1}^2 \lambda_1 + \cdots + {\alpha_{N_b}}^2 \lambda_{N_b} \leq \epsilon^2 \end{aligned}$$

where $\bar{\alpha}$ is the projection vector given by

$$\bar{\alpha} \triangleq \begin{bmatrix} \alpha_0 & \alpha_1 & \cdots & \alpha_{N_b} \end{bmatrix}^*$$

The principle drawback of the gradient projection method is the substantial overhead for computing the projection at each iteration. However, as we will address shortly,

the canonical property of this constraint set enables us to perform the projection in an efficient way. The main idea with regard to the gradient projection method is that, in each iteration, a feasible direction is obtained by taking a step along negative gradient followed by a projection onto the constraint set given by

$$\bar{\alpha}^{k+1} = \left[\bar{\alpha}^k - s^k \nabla f(\bar{\alpha}^k)\right]^+$$

Here $[.]^+$ denotes the projection onto constraint set $\mathbf{C}_1$, s^k is a positive step-size, and ∇f is the gradient of the objective function given by

$$\nabla f = \frac{-1}{\ln 2} \sum_{i=1}^{N} \frac{(\mathbf{Q}^i + \mathbf{Q}^{i^t})\bar{\alpha}}{\bar{\alpha}^* \mathbf{Q}^i \bar{\alpha}}$$

There are several step size selection procedures for the gradient projection method. In order to simplify the search direction, we consider a constant step size $s^k = s$. It is possible to show that the limit points of a sequence generated by the gradient projection with a constant step size are stationary, provided s is sufficiently small and the gradient satisfies a Lipschitz condition [49].

Next we derive the projection onto the convex set. Given a point $\bar{\alpha}_i \in R^{N_b+1}$, the projection of this point onto the set would be a point in the set such that it minimizes the distance $\|\bar{\alpha} - \bar{\beta}\|$ among all the points inside the set. In light of this fact, projection of a point $\bar{\alpha}_i \notin \mathbf{C}_1$ would be on the boundary of the set. Also each point inside the constraint set would satisfy the constraint and would be projected onto itself. Therefore the projection operator is defined as follows

$$[\alpha]^+ = \begin{cases} \bar{\alpha} & \text{if } \bar{\alpha} \in \mathbf{C}_1 \\ \bar{\beta} & \text{if } \bar{\alpha} \notin \mathbf{C}_1 \end{cases}$$

where $\bar{\beta} = \begin{bmatrix} \beta_0 & \beta_1 & \cdots & \beta_{N_b} \end{bmatrix}^*$ is a point in $\mathbf{C}_1$ which satisfies the constraint with equality

$$\mathbf{C}_1 = \{\bar{\beta} \in R^{N_b+1} \mid {\beta_0}^2\lambda_0 + {\beta_1}^2\lambda_1 + \cdots + {\beta_{N_b}}^2\lambda_{N_b} = \epsilon^2\}$$

To find this point we construct the Lagrange functional

$$\begin{aligned} J(\bar{\beta},\gamma) &= \|\beta-\alpha\|^2+\gamma\left[\sum_{i=0}^{N_b}\lambda_i{\beta_i}^2-\epsilon^2\right] \\ &= \sum_{i=0}^{N_b}\left[(\beta_i-\alpha_i)^2+\gamma(\lambda_i{\beta_i}^2)\right]+\gamma\epsilon^2 \end{aligned}$$

By taking the partial derivative of $J(\beta,\gamma)$ with respect to particular β_i, and setting it to zero, we obtain

$$\frac{\partial J(\beta,\gamma)}{\partial\beta_i}=2(\beta_i-\alpha_i)+2\gamma\lambda_i\beta_i=0$$

$$\beta_i=\frac{\alpha_i}{1+\gamma\lambda_i} \tag{4.12}$$

Also taking the partial derivative of the Lagrangian functional with respect to Lagrange multiplier γ and setting it to zero provides the following

$$\frac{\partial J(\bar{\beta},\gamma)}{\partial\gamma}=\sum_{i=0}^{N_b}(\beta_i)^2\lambda_i-\epsilon^2=0 \tag{4.13}$$

Substituting equation 4.12 in the above equation provides

$$\psi(\gamma)\triangleq\frac{\partial J(\bar{\beta},\gamma)}{\partial\gamma}=\sum_{i=0}^{N_b}(\frac{\alpha_i}{1+\gamma\lambda_i})^2\lambda_i-\epsilon^2=0 \tag{4.14}$$

Clearly equation (4.14) is a non-linear equation in γ . It can be shown that starting from $\gamma^0=0$ the iterates generated by Newton's method

$$\gamma^{k+1}=\gamma^k-\frac{\psi(\gamma^k)}{\psi'(\gamma^k)} \tag{4.15}$$

would always lead us to the unique positive solution of this equation, which results in a projection vector $\bar{\beta}$ that has a smaller distance to $\bar{\alpha}$ than that furnished by use of any other root [50]. Upon computing the Lagrange multiplier γ from the equation (4.15), the projection vector is found through equation (4.12).

A few additional remarks regarding to the effect of initial condition are appropriate. First we point out that the objective function lacks the convexity property. Therefore, the stationary point obtained from exploiting this algorithm is dependent upon the choice of initial condition. However, a proper choice of initial condition can lead us to a stationary point close to the optimal solution. A feasible initial condition can be the solution obtained from MMSE-UEC approach as given in equation (4.1). As a second comment, we also point out that in order for the algorithm to converge to a stationary point, the initial condition should be set so that the starting point satisfies the inequality constraint (feasible point). In the following sections, we investigate the effect of further constraints such as UTC and Linear phase on the performance of of the TEQ.

4.4 Unit Tap Constraint

In some applications, it is desirable to impose a UTC on the TIR filter. This constraint forces k^{th} tap of the TIR filter to unity. Decision feedback equalization is a special case of UTC with $k = 0$. As mentioned before, UTC forces the k^{th} element of TIR filter to unity. Using equation (4.10), we can represent k^{th} element of TIR as linear combination of k^{th} elements of orthonormal eigenvectors of matrix $\mathbf{R}_{\Delta}$. Consequently the UTC set ($\mathbf{C}_3$) can be formulated as

$$\mathbf{C}_3 = \{\bar{\beta} \in R^{N_b+1} \quad | \quad \sum_{j=0}^{N_b} \beta_j \mathbf{v}_j[k] = 1\} \tag{4.16}$$

The above equation conforms a hyperplane in R^{N_b+1} which is both closed and convex. In general, the projection operator is a vector $\bar{\beta}$ which minimizes the Lagrangian functional, i.e.

$$J(\bar{\beta}, \psi) \quad = \quad \|\bar{\beta} - \bar{\alpha}\|^2 + \psi(\sum_{j=0}^{N_b} \beta_j \mathbf{v}_j[k] - 1) \tag{4.17}$$

To obtain this point we compute the partial derivative of the above term with respect to ψ and set it to zero which provides

$$\frac{\partial J(\bar{\beta}, \psi)}{\partial \beta_i} \quad = \quad 2(\beta_i - \alpha_i) + \psi \mathbf{v}_i[k] = 0$$

Solving the above equation with respect to β_i provides the projection operator as

$$\beta_i = \alpha_i - \frac{\psi}{2}\mathbf{v}_i[k] \tag{4.18}$$

where ψ is the parameter which minimizes the term given in (4.17) as ,i.e.

$$\frac{\partial J(\bar{\beta}, \psi)}{\partial \psi} = \sum_{j=0}^{N_b} \beta_j \mathbf{v}_j[k] - 1 = 0$$

Substituting equation (4.18) into the above term results in a closed form expression for the parameter ψ as expressed by

$$\psi = 2\frac{\sum_{j=0}^{N_b} \alpha_j \mathbf{v}_j[k] - 1}{\sum_{j=0}^{N_b} \mathbf{v}_j^2[k]}$$

The above equation along with (4.18) results in the projection operator as

$$\beta_i = \alpha_i - \frac{\sum_{j=0}^{N_b} \alpha_j \mathbf{v}_j[k] - 1}{\sum_{j=0}^{N_b} \mathbf{v}_j^2[k]}\mathbf{v}_i[k] \tag{4.19}$$

4.5 Linear Phase Constraint

In optimizing the performance of TEQ, the effect of phase distortion was not considered. In order to remove the phase distortion, linear phase constraint must be imposed on the TIR filter. This would add another constraint set to the previous problem, which increases the complexity of the problem. However the linear phase constraint is the intersection of hyperplanes in Euclidean space R^{N_b+1} that is both closed and convex. Therefore, projection onto convex sets can be extended to provide the optimum solution under linear phase constraint. Following, we provide the projection operator for linear phase filters type (I) and (III) [12].

4.5.1 Linear Phase Type I

Linear phase type I satisfies the even symmetry property as given by

$$\mathbf{b}[k] = \mathbf{b}[N_b - k] \quad k = 0, 1, \cdots, \frac{N_b}{2} - 1$$

Using equation (4.10), the above constraint can be written as a function of orthonormal eigenvectors of matrix $\mathbf{R}_\Delta$ as expressed by

$$\sum_{j=0}^{N_b} \beta_j \mathbf{v}_j[k] = \sum_{j=0}^{N_b} \beta_j \mathbf{v}_j[N_b - k] \qquad k = 0, 1, \cdots, \frac{N_b}{2} - 1$$

Therefore linear phase type (I) constraint set is the intersection of $\frac{N_b}{2}$ hyperplanes in R^{N_b+1} given by

$$\mathbf{C}_4 = \{\bar{\beta} \in R^{N_b+1} \quad | \quad \sum_{j=0}^{N_b} \beta_j \rho_j[k] = 0 \qquad , k = 0, 1, \cdots, \frac{N_b}{2} - 1\}$$

whose j^{th} component of the normal vector corresponding to the n^{th} hyperplane has the form of

$$\rho_j[n] \triangleq \mathbf{v}_j[n] - \mathbf{v}_j[N_b - n] \tag{4.20}$$

The projection vector β represents a point on the intersection of these $\frac{N_b}{2}$ planes which minimizes the following Lagrangian functional

$$J(\bar{\beta}, \bar{\psi}) \quad = \quad \|\beta - \alpha\|^2 + \sum_{k=0}^{\frac{N_b}{2}-1} \psi_k \sum_{j=0}^{N_b} \beta_j \rho_j[k] \tag{4.21}$$

where $\bar{\psi} \triangleq [\ \psi_0 \quad \psi_1 \quad \cdots \quad \psi_{\frac{N_b}{2}-1}\ ]^*$. By computing partial derivative of the above term with respect to vector β and setting it to zero we obtain the projection operator as

$$\beta_i = \alpha_i - \frac{1}{2} \sum_{k=0}^{\frac{N_b}{2}-1} \psi_k \rho_i[k] \qquad i = 0, 1, \cdots, N_b \tag{4.22}$$

In the above term, ψ_k is the k^{th} element of the vector $\bar{\psi}$ which minimizes the term given in (4.21). Taking the partial derivative of $J(\bar{\beta}, \bar{\psi})$ withrespect to particular ψ_n and setting it to zero, we obtain

$$\frac{\partial J(\bar{\beta}, \bar{\psi})}{\partial \psi_n} = \sum_{j=0}^{N_b} \beta_j \rho_j[n] = 0 \qquad n = 0, 1, \cdots, \frac{N_b}{2} - 1$$

substituting equation (4.22) into the above term yields

$$\sum_{k=0}^{\frac{N_b}{2}-1} \psi_k \frac{1}{2} \sum_{j=0}^{N_b} \rho_j[n] \rho_j[k] = \sum_{j=0}^{N_b} \alpha_j \rho_j[n]$$

The above term is a set of $\frac{N_b}{2}$ simultaneous linear equations which can be cast in a matrix form as following

$$\Theta\bar{\psi} = \mathbf{p} \tag{4.23}$$

where the matrix Θ and vector $\mathbf{p}$ are defined as

$$\Theta[n,k] \triangleq \sum_{j=0}^{N_b} \rho_j[n]\rho_j[k] \quad n,k = 0,1,\cdots,\frac{N_b}{2}-1 \tag{4.24}$$

$$\mathbf{p}[n] \triangleq 2\sum_{j=0}^{N_b} \alpha_j\rho_j[n] \quad n = 0,1,\cdots,\frac{N_b}{2}-1 \tag{4.25}$$

There are two approaches to compute the vector $\bar{\psi}$. The first method is to apply the idea of POCS and compute the projection using successive projection on $\frac{N_b}{2}$ hyperplanes. An alternative method is to compute the vector $\bar{\psi}$ through inverting the matrix Θ as given by

$$\bar{\psi} = \Theta^{-1}\mathbf{p} \tag{4.26}$$

4.5.2 Linear Phase Type III

Linear phase filters type III satisfy the even antisymmetry property as expressed by

$$\mathbf{b}[n] = \begin{cases} -\mathbf{b}[N_b - n] & n = 0,\cdots,\frac{N_b}{2}-1 \\ 0 & \frac{N_b}{2} \end{cases}$$

Likewise linear phase type (I), the first constraint is the intersection of hyperplanes in R^{N_b+1}. Also, the second constraint can be viewed as another hyperplane in R^{N_b+1} which encounters origin and can be represented as

$$\mathbf{b}[\frac{N_b}{2}] = \sum_{j=0}^{N_b} \beta_j \mathbf{v}_j[\frac{N_b}{2}] = 0$$

Using the above equations, the Lagrangian functional can be written as

$$J(\bar{\beta},\bar{\psi}) = \|\beta-\alpha\|^2 + \sum_{k=0}^{\frac{N_b}{2}-1} \psi_k \sum_{j=0}^{N_b} \beta_j\tau_j[k] + \psi_{\frac{N_b}{2}} \sum_{j=0}^{N_b} \beta_j \mathbf{v}_j[\frac{N_b}{2}]$$

where the variables $\tau_j[n]$ and $\bar{\psi}$ are defined as

$$\tau_j[n] \triangleq \mathbf{v}_j[n] + \mathbf{v}_j[N_b - n] \tag{4.27}$$

$$\bar{\psi} \triangleq \begin{bmatrix} \psi_0 & \psi_1 & \cdots & \psi_{\frac{N_b}{2}} \end{bmatrix}^* \tag{4.28}$$

Taking the partial derivative of equation (4.27) with respect to β_i and setting it to zero provides the projection operator as follows

$$\beta_i = \alpha_i - \frac{1}{2}\sum_{k=0}^{\frac{N_b}{2}-1} \psi_k \tau_i[k] - \frac{1}{2}\psi_{\frac{N_b}{2}} \mathbf{v}_i[\frac{N_b}{2}] \tag{4.29}$$

Similarly taking the partial derivate of equation (4.27) and setting it to zero results in

$$\frac{\partial J(\bar{\beta}, \bar{\psi})}{\partial \psi_n} = \begin{cases} \sum_{j=0}^{N_b} \beta_j \tau_j[n] = 0 & n = 0, \cdots, \frac{N_b}{2} - 1 \\ \sum_{j=0}^{N_b} \beta_j \mathbf{v}_j[\frac{N_b}{2}] = 0 & n = \frac{N_b}{2} \end{cases}$$

Substituting equation (4.29) into the above equations results in a set of $\frac{N_b}{2}+1$ simultaneous equations as follows

$$\sum_{k=0}^{\frac{N_b}{2}-1} \psi_k \sum_{j=0}^{N_b} \tau_j[n]\tau_j[k] + \psi_{\frac{N_b}{2}} \sum_{j=0}^{N_b} \tau_j[n]\mathbf{v}_j[\frac{N_b}{2}] = 2\sum_{j=0}^{N_b} \alpha_j \tau_j[n]$$

$$\sum_{k=0}^{\frac{N_b}{2}-1} \psi_k \sum_{j=0}^{N_b} \mathbf{v}_j[\frac{N_b}{2}]\tau_j[k] + \psi_{\frac{N_b}{2}} \sum_{j=0}^{N_b} \mathbf{v}_j[\frac{N_b}{2}]^2 = 2\sum_{j=0}^{N_b} \alpha_j \mathbf{v}_j[\frac{N_b}{2}]$$

The above equations can be cast in a matrix form as

$$\mathbf{\Gamma}\psi = \mathbf{r} \tag{4.30}$$

where the matrix $\mathbf{\Gamma}$ and vector $\mathbf{r}$ are defined as

$$\mathbf{\Gamma}[m,n] \triangleq \begin{cases} \sum_{j=0}^{N_b} \tau_j[m]\tau_j[n] & m,n = 0, \cdots, \frac{N_b}{2} - 1 \\ \sum_{j=0}^{N_b} \tau_j[n]\mathbf{v}_j[\frac{N_b}{2}] & m = \frac{N_b}{2} \;\&\; n = 0, \cdots, \frac{N_b}{2} \\ \sum_{j=0}^{N_b} \tau_j[m]\mathbf{v}_j[\frac{N_b}{2}] & n = \frac{N_b}{2} \;\&\; m = 0, \cdots, \frac{N_b}{2} \\ \sum_{j=0}^{N_b} \mathbf{v}_j[\frac{N_b}{2}]^2 & m = n = \frac{N_b}{2} \end{cases}$$

$$\mathbf{r}[n] \triangleq \begin{cases} 2\sum_{j=0}^{N_b} \tau_j[m]\alpha_j & n = 0, 1, \cdots, \frac{N_b}{2} - 1 \\ 2\sum_{j=0}^{N_b} \tau_j[m]\alpha_j \mathbf{v}_j[\frac{N_b}{2}] & n = \frac{N_b}{2} \end{cases}$$

The lagrangian multiplier vector can be obtained through successive iteration of POCS algorithm or simply through inverting the matrix Γ as follows

$$\psi = \mathbf{\Gamma^{-1}r} \tag{4.31}$$

4.6 Upper-bound on Performance

Assuming the optimization problem given in (4.7) finely maximizes the objective function given in (4.6), an upper-bound on the total number of bits transmitted in one DMT symbol can be obtained as follows

$$\begin{aligned} \sum_{i=1}^{N} log_2 \mathbf{b^* G^i b} &= log_2 \prod_{i=1}^{N} |\mathbf{b} \cdot \mathbf{g^i}|^2 \\ \overset{1}{\leq} log_2 \prod_{i=1}^{N} \|\mathbf{b}\|^2 \|\mathbf{g}^i\|^2 &\overset{2}{\leq} log_2 (\frac{\epsilon^2 (N_b+1)^2}{\lambda_{min}(\mathbf{R}_\Delta)})^N \end{aligned} \tag{4.32}$$

Note that the first inequality follows from the Cauchy-Schwartz inequality while the second inequality is obtained from applying Rayleigh inequality as expressed by

$$\lambda_{min}(\mathbf{R}_\Delta) \leq \frac{\mathbf{b^* R_\Delta b}}{\|\mathbf{b}\|^2} \leq \lambda_{max}(\mathbf{R}_\Delta)$$

where λ_{min} (λ_{max}) is the minimum (maximum) eigenvalue of matrix $\mathbf{R}_\Delta$.

The above expression shows that increasing quadratic inequality constant (QIC) ϵ^2, would result in a larger upper-bound for the objective function. On the other hand, a smaller value for QIC, would cause the dual constraint problem given in (4.7) to better approximate maximization of the original objective function described in (4.6). This fact is consolidated through computer simulation in the subsequent section.

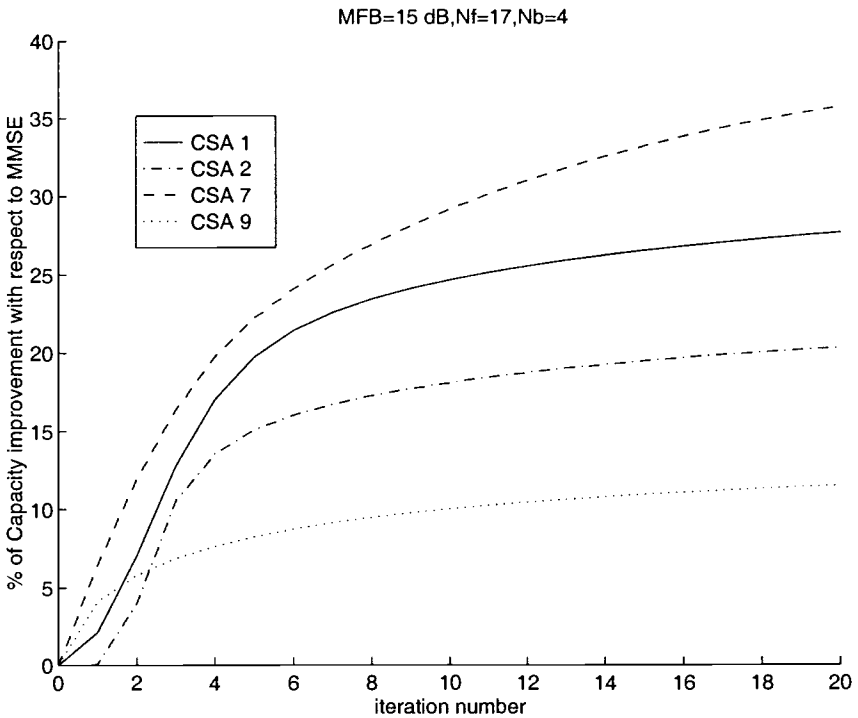


Figure 4.3: Performance Of The Proposed Algorithm For Various CSA Lines

4.7 Simulations and Performance Evaluation of the Algorithm

In this section, we explore the potential performance achievable through the use of the proposed algorithm for equalization of DMT systems. We ran a series of simulations on CSA loops sampled at 276 kHz. The number of subchannels considered is $N = 64$. The TEQ and TIR are assumed to have lengths of $N_f = 17$ and $N_b = 4$, respectively. Receiver and thermal noise is modeled as AWGN noise with power of -30 dBm across the two-sided bandwidth. Near end cross-talk (NEXT) noise is modeled by exciting coupling filter with spectrum of ($|H_x(f)|^2 = 10^{-13} f^{3/2}$) by a white Gaussian noise with power of 10 mW. Unless specified, signal power is set such that the matched filter bound (MFB $\triangleq$ $||h||^2 \sigma_x^2/\sigma_n^2$) of 15 dB is achieved at the receiving point. Furthermore, it is assumed that

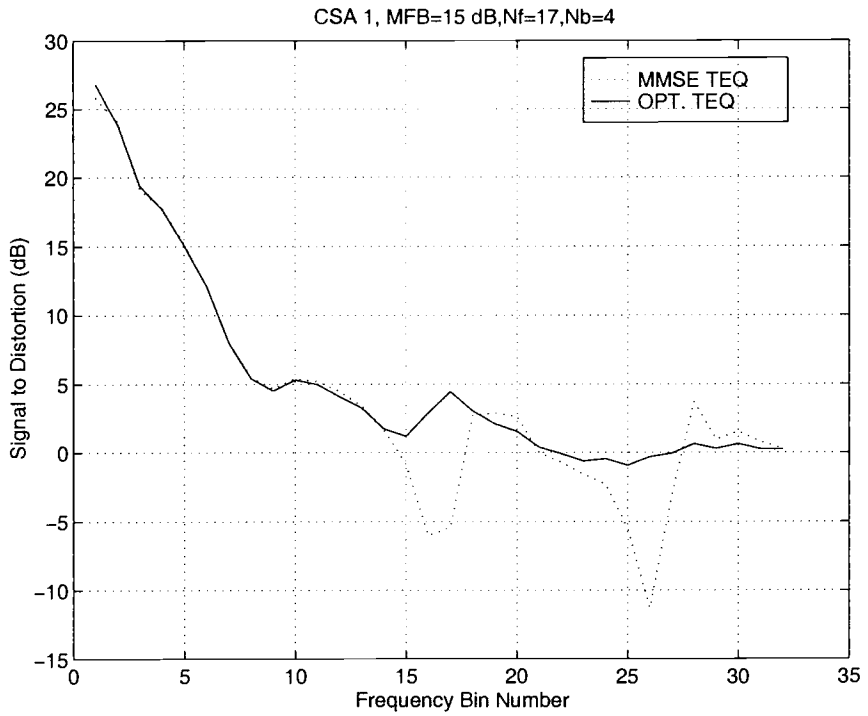


Figure 4.4: Signal To Distortion Profile For MMSE-TEQ and Max Capacity-TEQ

the power is equally divided amongst the entire subchannels. In estimating the capacity of the DMT system, the entire bandwidth is used and no limitation is imposed on the number of bits allocated for each subchannel. Also the coding gain of 0 dB is assumed over entire subchannels.

4.7.1 Effect of Channel Impulse Response

In order to evaluate the performance of the proposed algorithm, computer simulations have been performed on a series of CSA loops. Figure 4.3 shows the percentage of improvement in capacity with respect to the capacity obtained from MMSE-UEC approach. Decision delay, initial condition and QIC are set to the settings furnished by the MMSE-UEC approach. Simulation results indicate that the algorithm exhibits robust

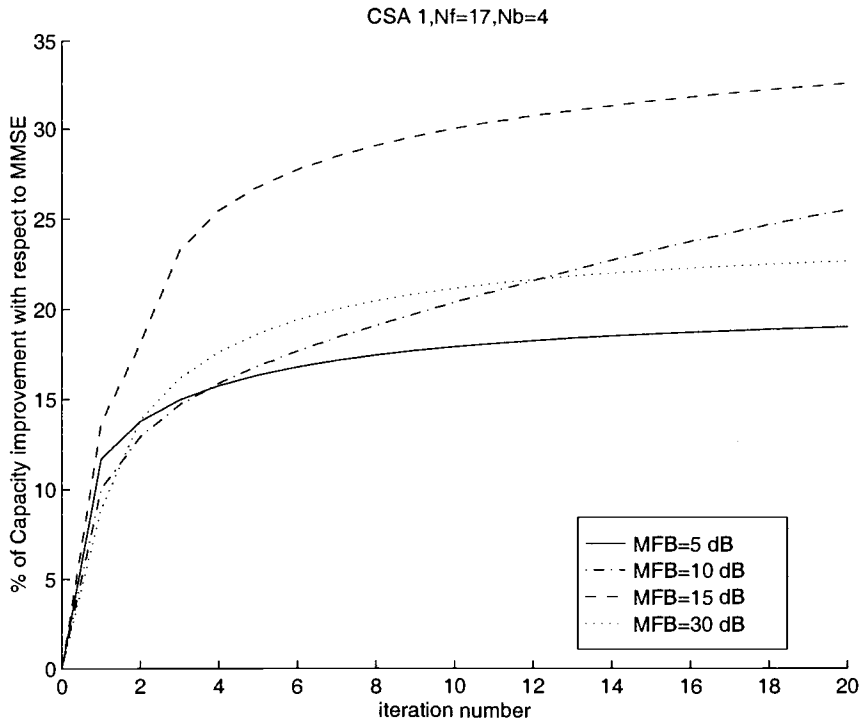


Figure 4.5: Performance Of The Proposed Algorithm For Different MFBs

convergence for all CSA loops used in the study. As shown in the figure, the capacity of the proposed method exceeds that of MMSE-UEC approach in the range of 10% to 35 %. Figure 4.4 compares the signal to distortion ratio of maximum capacity equalization against MMSE-UEC approach for CSA-1 loop. As shown in the figure, the MMSE approach exhibits considerable performance degradation over half of the subchannels. This degradation can be viewed as sharp notches in the signal to distortion profile. Equalization of DMT based on maximum capacity outperforms the MMES-UEC approach by removing these nulls from the signal to distortion profile.

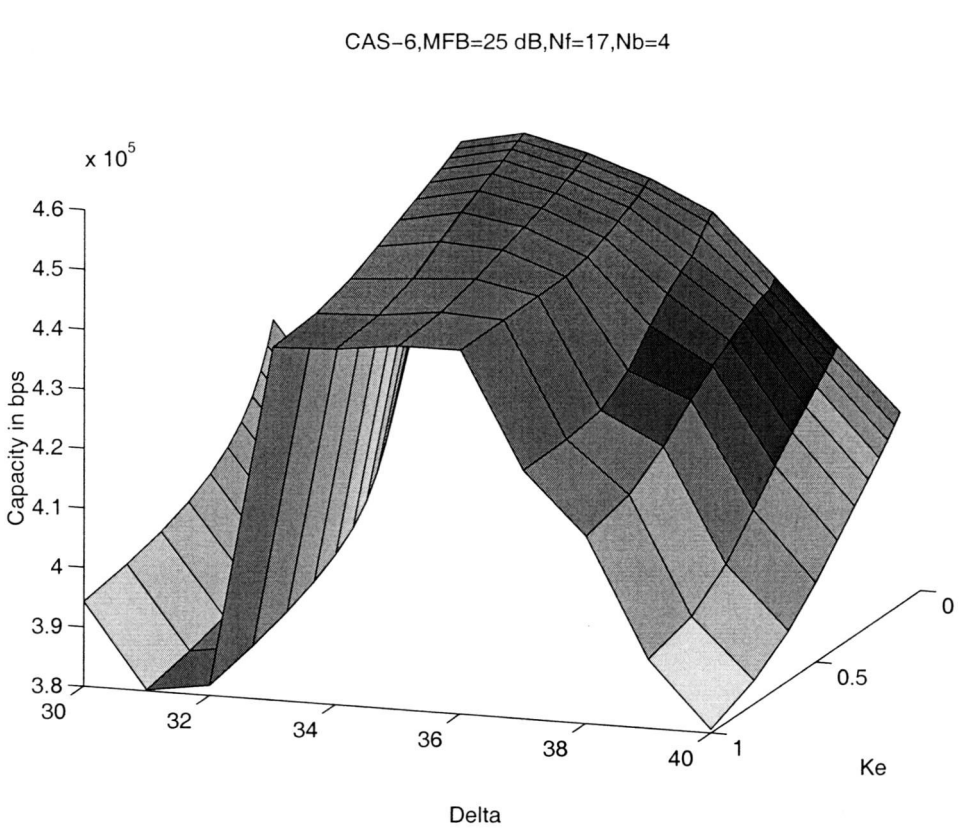


Figure 4.6: Capacity Profile Versus Decision Delay and QIC

4.7.2 Effect of Transmit Power

We applied the maximum capacity equalization on a typical CSA loop, namely CSA 1, under various transmit power. Figure 4.5 shows the relative improvement in capacity with respect to the capacity obtained from MMSE-UEC approach. The transmit power is set so the matched filter bound at the receiving end achieves values of 5, 10, 15, and 30 dB. Likewise previous experiment, initial condition, optimum delay and QIC ϵ^2 are set to the values furnished by MMSE-UEC approach. Simulation results show that a considerable improvement in system performance can be achieved through the use of proposed algorithm. Also as inferred from the figure, robustness of the algorithm is insensitive to the transmit power.

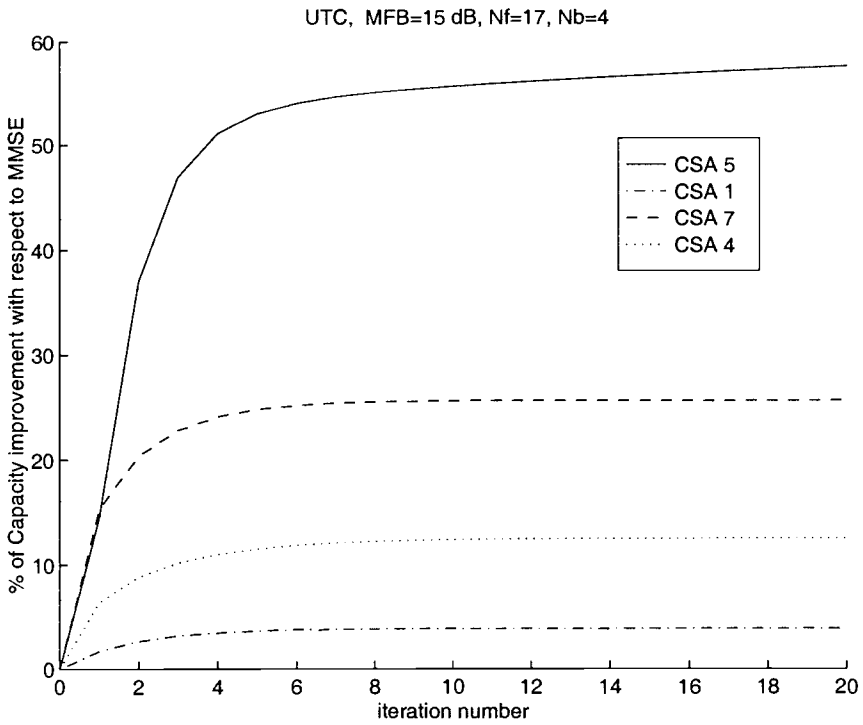


Figure 4.7: Performance of the Proposed Algorithm with UTC

4.7.3 Effect of QIC (ϵ^2) and Decision Delay (Δ)

In order to investigate the effect of QIC, the proposed algorithm is applied for equalization of a typical CSA loop, namely CSA-6. Figure 4.6 depicts the capacity profile as a joint function of QIC and decision delay. The QIC is set to $\epsilon^2 = K_e \epsilon^2_{\text{MMSE}}$, where $K_e \in [0.2 \quad 1]$. As mentioned earlier, increasing QIC would increase the volume of the constraint region which provides more freedom in the search direction. On the other hand, the dual optimization problem would better approximate the capacity maximization problem if QIC is small. Therefore, the admissible range of QIC is bounded from both sides and an exhaustive search should be performed on QIC to obtain its optimum value. As is shown in the figure, the maximum capacity is displaced downward as K_e increases from .2 to 1.

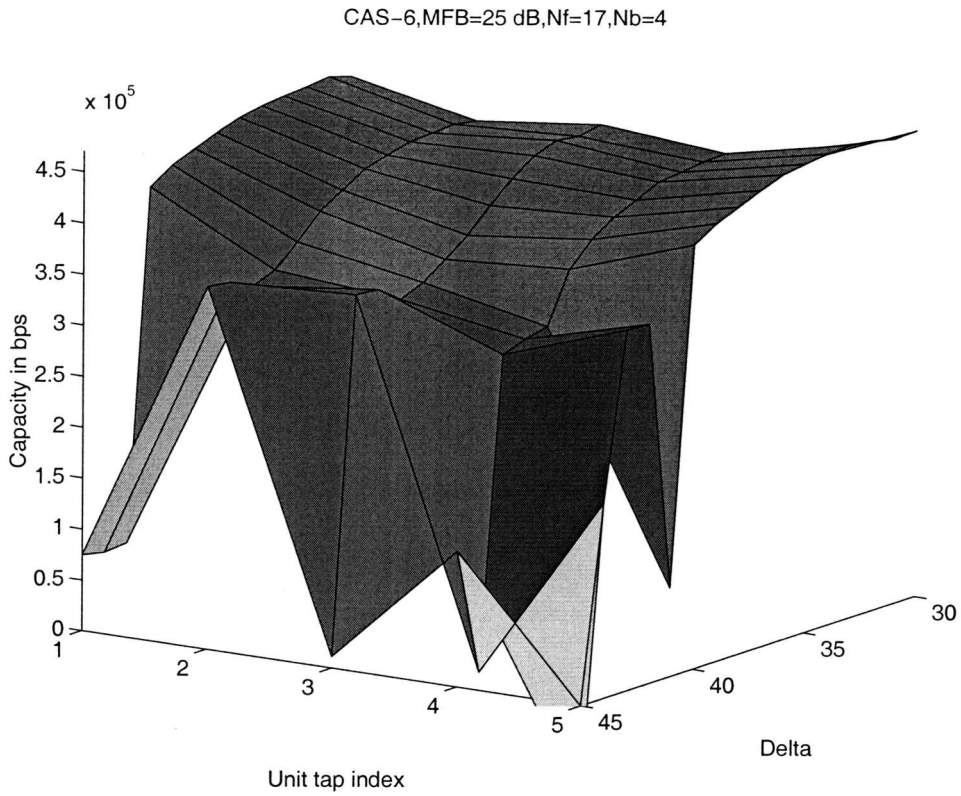


Figure 4.8: Capacity Profile Versus Decision Delay and Noncausality Index

4.7.4 UTC and Effect of Filter Causality

Next we examine the effect of UTC on the optimum equalization of DMT system. Figure 4.7 shows the relative improvement in capacity for maximum capacity equalization under UTC for various CSA lines. In optimizing the performance of the equalizer, QIC, decision delay and unit tap index are set to the optimum values obtained from MMSE-UTC approach. Changing the causality of TIR filter results in a different normal vector for the constraint hyperplane, which shapes the constraint region accordingly. The optimum unit tap index k is found through performing exhaustive search on values k ranging from 1 to $N_b + 1$. In order to show how the causality of TIR can affect the performance of the proposed algorithm, the optimum equalization algorithm is performed on on a typical

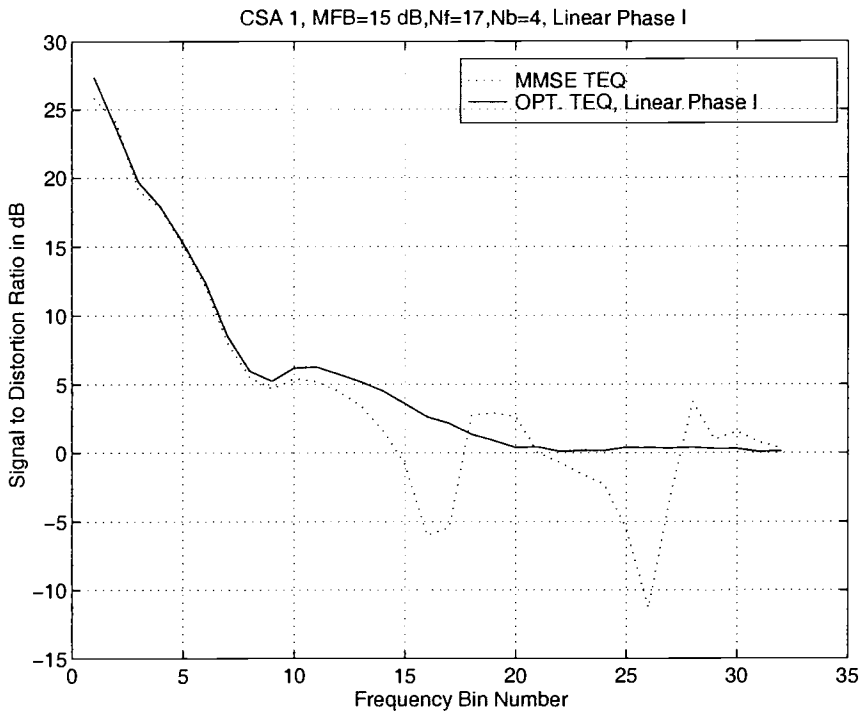


Figure 4.9: Signal to Distortion Ratio for MMSE-TEQ and Max Capacity-TEQ with Linear Phase type (I) Constraint

CSA loop. Figure 4.8 depicts the capacity profile as a joint function of decision delay and unit tap index k for CSA-1 loop.

Unit tap index 1 and 5 correspond to causal and noncausal TIR filters respectively. As shown in the figure, a noncausal TIR would maximize the the performance of the DMT system for this particular case.

4.7.5 Effect of Phase Distortion

In order to investigate the effect of phase nonlinearity on the performance of DMT system we impose the linear phase constraint on the maximum capacity equalization problem. We consider the CSA-1 loop used in section 4.7.1 and impose the linear phase type I constraint on the optimum equalization.

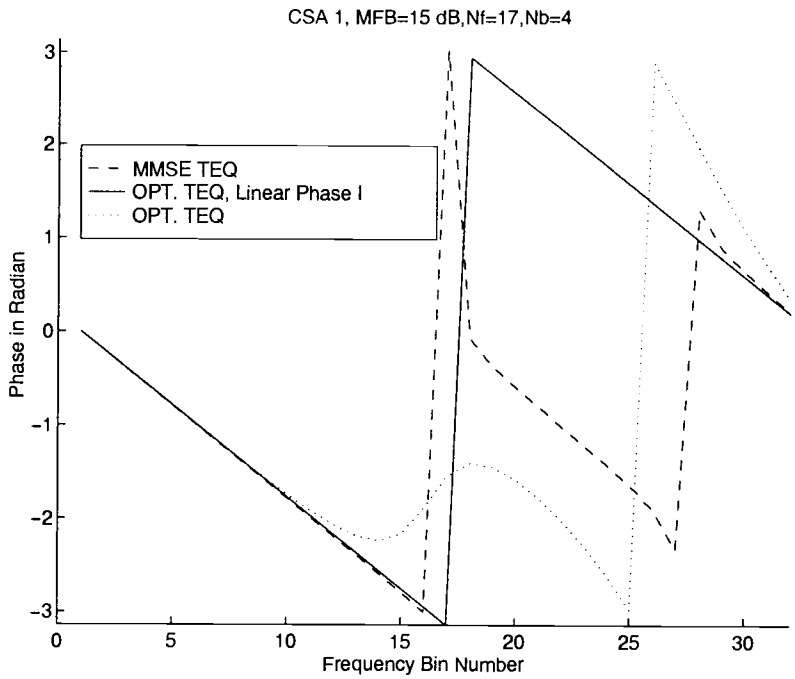


Figure 4.10: Comparison Between Phase Response of TIR Filter Under Various Constraints

4.8 Conclusion

Figure 4.9 shows the signal to distortion ratio over subchannels for this experiment. Figure 4.10 compares the phase response of the TIR for MMSE-UEC, optimum capacity equalization and optimum capacity equalization with linear phase constraint. As expected, the optimum capacity equalization with linear phase outperforms the other schemes through removing the phase distortion from the frequency response. Optimum equalization of multicarrier systems can be viewed as a constrained optimization problem over convex sets. The constraint sets exhibit identifiable geometrical characteristics which make the projection operation significantly efficient. Based on these observations, we have proposed a novel iterative algorithm as a straightforward application of POCS for solving the optimum equalization of multicarrier systems. Further work will also be carried out to demonstrate the impact of the bit loading algorithm on the overall capacity of the system.

5. CLASS OF CYCLIC BASED ESTIMATORS FOR FREQUENCY OFFSET ESTIMATION OF OFDM SYSTEMS

In this chapter we present a *new*, class of non data-aided cyclic based estimators for carrier frequency offset and symbol timing error estimation of orthogonal frequency division modulation (OFDM) systems. Essentially, the proposed approach exploits the properties of cyclic prefix subset to reveal the synchronization parameters in the likelihood function of received vector. This approach is an extension of the previously proposed estimation method given in [37]. However, our research indicates that the previously proposed likelihood metric does not globally characterize the estimation problem. Based on this observation, a new likelihood function for the joint timing and frequency offset estimation is derived. The resulting probabilistic measure is used to develop three class of unbiased estimators namely, maximum likelihood, minimum variance unbiased (MVU), and moment estimators. In comparison to the previously proposed methods, the proposed estimators are computationally and statistically efficient which makes the estimators more attractive for real time applications. Performance of estimators is assessed by computer simulation for OFDM scheme.

5.1 Introduction

OFDM system is a viable modulation scheme for data transmission time varying dynamic channels [51, 52] . However, it is known that performance of such system is highly susceptible to nonideal synchronization parameters [33, 53]. Specifically, symbol timing and carrier frequency offset become an increasingly important concern in using OFDM systems for practical applications [9, 54]. It is known that carrier frequency offset deteriorates performance of OFDM systems by introducing interference among the subchannels [32]. To overcome this imperfection, various compensation methods for estimation and correction of synchronization parameters have recently been proposed [32, 37, 55]. In order

to compare the performance of these estimators, it is required to define a single number representing the goodness of the estimate. Knowing that all the estimators are unbiased, i.e. expectation of the estimate is equal to the parameter, the variance of the estimator is used as a global measure for performance comparison of these estimators.

Cramer-Rao lower bound (CRLB) is a fundamental lower bound on the variance of the estimators and the estimator whose variance equals CRLB is called efficient. When the evaluation of efficient estimator is not possible, it is desirable to obtain an estimator in which its performance becomes as close as possible to the CRLB fundamental bound. The estimator in which its performance is the closest to the CRLB estimator is known as MVU estimator.

Categorically, the previously proposed methods for synchronization of OFDM systems can be classified into two main subclasses, namely minimum mean square error (MMSE) and ML estimators. In MMSE approach, estimator uses the side information provided by the reference signal (pilot tones) in order to minimize a cost function associated with the synchronization parameters [34],[35]. A salient feature of this approach is that no probabilistic assumptions are made in regard to the data. Although MMSE estimators usually result in a tractable (globally stable) and easy to implement realization, no optimality criteria (probabilistic or statistical) is associated with these estimators. Also, since part of the transmitted information is allocated to the reference pilots, the bandwidth efficiency of these methods is lower in comparison to the nonpilot schemes.

On the other hand, ML estimators provide the estimate of the unknown parameter subject to minimum probability of error criteria [36], [37], [32]. Although not exactly efficient, ML estimators are asymptotically MVU, i.e. their variance attains that of MVU estimator as length of data record goes to infinity. However, due to the physical constraints, systems with infinitely long data records are not feasible for implementation purposes.

In [32], authors use retransmission technique in order to reveal the frequency offset parameter in the likelihood function of the received signal. Due to the redundancy introduced by repeating the data block, the spectral efficiency is dropped by a factor of

two. To avoid this imperfection, a new ML estimator based on cyclic prefix (CP) was introduced in [37]. In this approach, the side information provided by the CP is used to obtain the likelihood function for joint estimation of symbol timing error and frequency offset in OFDM systems.

However, our research reveals that the likelihood function proposed in [37] does not globally characterize the observation vector over the entire range of the timing offset. Consequently, the ML estimator proposed based on this likelihood function would result in considerable performance loss over a finite range of timing offset interval.

Motivated by the suboptimum performance of this estimator, a new likelihood function for joint estimation of carrier frequency offset and symbol timing error of OFDM systems is introduced in this paper. Based on this result, a new optimum ML estimator for the joint estimation problem is also presented. In attempt to reduce the variance of ML estimator, we also investigate a new class of MVU estimators for frequency offset estimation of OFDM systems. It is shown that there exists but one function of sufficient statistic which provides the MVU estimate of the frequency offset. The proposed estimator provides a closed form expression for the estimator as a function of data statistic. Consequently, it does not suffer from converging to multiple local minima, the problem which arises in ML technique with nonconvex loglikelihood function [36].

The advantage of the proposed MVU estimator over the class of previously proposed estimators is two fold. First, it is MVU , therefore its variance is minimum among the entire class of estimators which use the same probabilistic measure. Secondly, it provides a closed form expression for mapping the statistics into the estimation domain. The former property assures optimality of the estimator, while the later facilitate the closed loop analysis of the system.

The rest of this paper is organized as follows. Section 5.2 introduces the timing and frequency offset estimation problem and addresses the previously proposed estimation approach. In section 5.3 a new global likelihood function for joint estimation of timing and frequency offset is presented. Once the likelihood is found, a new ML estimator for joint estimation problem is proposed in section 5.4. In section 5.5 the powerful Neyman-

Fisher factorization theorem and Rao-Blackwell-Lehmann-Scheffe (RBLS) theorem are applied to obtain the MVU estimator. Lower bound and closed loop performance of this estimator is also investigated in this section. Section 5.6 addresses a moment estimator for estimation of frequency offset under uncertain timing offset knowledge. In section we propose a unified structure which encounters the class of proposed cyclic based estimators. Performance assessments of the estimators for OFDM system are presented in section .

5.2 Preliminaries

This section presents an overview of cyclic based estimation method for synchronization of OFDM system [37]. In the OFDM, an N-point FFT is used to divide the channel spectrum into a set of N parallel subchannels. Due to the intersymbol interference introduced by non-flat spectrum channel, the OFDM symbols are subject to the interblock interference (IBI) among consecutive transmitted blocks [56]. This would result in considerable performance degradation in OFDM system. To mitigate this effect, the last L input samples in each input block of length N are repeated at the beginning of the block. This makes the input sequence look periodic and clears the channel memory at the end of each input block making the successive OFDM symbols independent. Figure 5.1 depicts the transmission blocks in OFDM systems. We denote the OFDM received signal by

$$y[n] = s[n-\vartheta]\exp^{j\frac{2\pi n\varepsilon}{N}} + w[n] \tag{5.1}$$

where $s[.]$ is the transmitted sequence and w is the additive white Gaussian noise (AWGN). Both signal and noise sequences are assumed to be uncorrelated independent identically distributed (iid) random variables with power of σ_s^2 and σ_n^2 respectively. Also, ε and ϑ are the frequency offset and symbol timing error introduced by the synchronization mismatch in the carrier frequency and symbol timing respectively. Let $\mathbf{x} = [x[0]\cdots x[2N+L-1]]$ be a vector of $2N+L$ previously received samples at time n, known as observation vector (OV). With the above notation, the kth entry of this vector can be represented as $x[k] =$

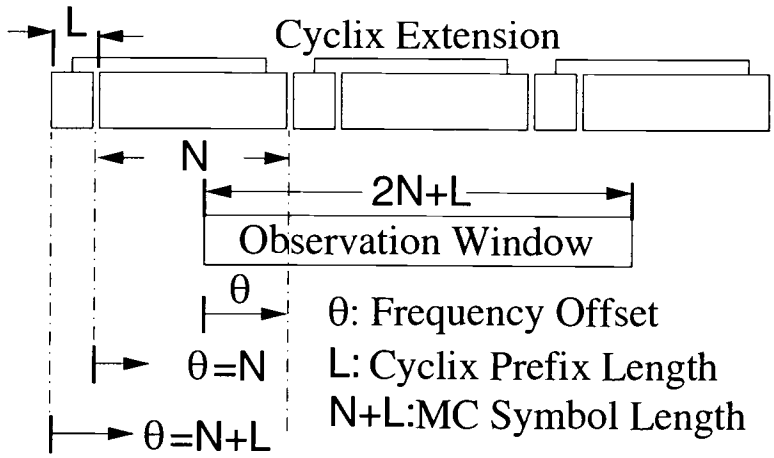


Figure 5.1: Observation Window

$y[n-2N-L+1+k]$. Length of this vector is selected such that there are at least L correlated symbols associated in the OV regardless of the synchronization parameters values. Due to the timing offset error ϑ at the receiver, the starting point of this vector is shifted by ϑ samples with respect to the beginning of the OFDM symbol block.

Under the iid assumption for both signal and noise, the autocorrelation function for the OV $\mathbf{x}$ can be expressed by

$$E\{x[k]x^*[k+m]\} = \begin{cases} \sigma_s^2 + \sigma_n^2 & m = 0 \\ \sigma_s^2 e^{-j2\pi\varepsilon} & m = N, k \in \Omega \\ 0 & \text{otherwise} \end{cases} \tag{5.2}$$

where $\Omega = \{k \in N : \vartheta \leq k \leq \vartheta + L - 1\}$ is the cyclic associated with the OV. Assuming AWGN scenario, the probability density function (PDF) of the observation vector is in the form of

$$p(\mathbf{x}, \vartheta, \varepsilon) = \frac{1}{(2\pi)^{2N+L} det(\mathbf{R})} \exp[\mathbf{x}^* \mathbf{R}^{-1} \mathbf{x}] \tag{5.3}$$

where $\mathbf{R}$ is the autocorrelation matrix of the observation vector. In [37] authors use the Bayes conditional theorem in order to obtain the PDF of the observation vector. However, as we show next, the resulting likelihood measure does not fully characterize the random observation vector $\mathbf{x}$ over the entire range of the timing offset parameter. In the next section we propose a new likelihood function which removes this imperfection.

5.3 Likelihood Measure

The question of the optimal choice of likelihood function for joint estimation of frequency and timing offset in OFDM systems is considered next. The analysis developed here for computing the PDF is based on the standard matrix inversion approach. Depending on the value of timing offset ϑ, the autocorrelation matrix ($\mathbf{R}$) can be cast into one of the following forms

5.3.1 Case I $(1 \leq \vartheta \leq N)$

In this case there are two cyclic sets associated with the observation vector. This would partition the autocorrelation matrix and its inverse into the following forms

$$\mathbf{R} = \begin{bmatrix} r_{xx}[0]\mathbf{I}_{(\vartheta)} & \mathbf{0} & \mathbf{0} \\ 0 & \mathbf{T}_{(N+L)} & 0 \\ 0 & 0 & r_{xx}[0]\mathbf{I}_{(N-\vartheta)} \end{bmatrix} \tag{5.4}$$

$$\mathbf{R}^{-1} = \begin{bmatrix} \frac{\mathbf{I}_{(\vartheta)}}{r_{xx}[0]} & \mathbf{0} & \mathbf{0} \\ 0 & \mathbf{T}^{-1}_{(N+L)} & 0 \\ 0 & 0 & \frac{\mathbf{I}_{(\mathbf{N}-\vartheta)}}{r_{xx}[0]} \end{bmatrix} \tag{5.5}$$

(5.6)

where $\mathbf{T}_{(M)}$ is a tridiagonal toepliz matrix of size M as expressed by

$$\mathbf{T}_M \triangleq Toeplitz([\ r_{xx}[0] \quad r_{xx}[1] \quad \cdots \quad r_{xx}[M-1]\ ]) \tag{5.7}$$

The argument inside the *Toeplitz* operator is the first row of the toeplitz matrix $\mathbf{R}_{(\vartheta)}$ with

$$r_{xx}[m] \triangleq (\sigma_s^2 + \sigma_n^2)\delta[m] + \sigma_s^2 e^{-j2\pi\varepsilon}\delta[m-N] \tag{5.8}$$

Due to the tridiagonal toeplitz property of the matrix $T_{(M)}$, the $[i,j]$ entry of its inverse can be obtained from

$$T_{(M)}^{-1}[i,j] = \begin{cases} \frac{1}{r_{xx}[0](1-|a|^2)} & i=j \ \ \& \ \ (1 \le i \le M-N \ \text{ or } \ N+1 \le i \le M) \\ \frac{-a^*}{r_{xx}[0](1-|a|^2)} & i-j=N \\ \frac{-a}{r_{xx}[0](1-|a|^2)} & j-i=N \\ \frac{1}{r_{xx}[0]} & i=j \ \ \& \ \ M-N+1 \le i \le N \\ 0 & \text{otherwise} \end{cases} \tag{5.9}$$

where

$$a \triangleq \frac{e^{-j2\pi\varepsilon}\sigma_s^2}{\sigma_s^2 + \sigma_n^2} \tag{5.10}$$

Substituting equations (5.9) and (5.5) into (5.3), after some algebraic manipulation, the conditional PDF for this case can be expressed as

$$\begin{aligned} p_1(\mathbf{x},\varepsilon,\vartheta) &= \frac{1}{(2\pi)^{2N+L}det(\mathbf{R})}\exp[\frac{-1}{2(\sigma_s^2+\sigma_n^2)}(\sum_{k=0}^{2N+L-1}|x[k]|^2 \\ &+ \sum_{k=\vartheta}^{L+\vartheta-1}(|x[k]|^2+|x[k+N]|^2)\frac{|a|^2}{1-|a|^2} - 2\Re\{\sum_{k=\vartheta}^{\vartheta+L-1}x^*[k]x[k+N]\frac{a^*}{1-|a|^2}\})] \end{aligned} \tag{5.11}$$

5.3.2 Case II $(N+1 \le \vartheta \le N+L)$

In this case there are three cyclic sets associated with the observation vector and the autocorrelation matrix is partitioned into two tridiagonal toeplitz matrices as expressed by

$$\mathbf{R} = \begin{bmatrix} \mathbf{T}_{(\vartheta)} & 0 \\ 0 & \mathbf{T}_{(2N+L-\vartheta)} \end{bmatrix} \tag{5.12}$$

Substituting equation (5.12) into (5.3), and using (5.9) the conditional PDF for this case can be written as

$$\begin{aligned} p_2(\mathbf{x},\varepsilon,\vartheta) &= \frac{1}{(2\pi)^{2N+L}det(\mathbf{R})}\exp[\frac{-1}{2(\sigma_s^2+\sigma_n^2)}(\sum_{k=0}^{2N+L-1}|x[k]|^2 \\ &+ (\sum_{k=0}^{\vartheta-N-1}|x[k]|^2+|x[k+N]|^2+\sum_{k=\vartheta}^{N+L-1}|x[k]|^2+|x[k+N]|^2)\frac{|a|^2}{1-|a|^2} \\ &- 2\Re\{\sum_{k=0}^{\vartheta-N-1}x^*[k]x[k+N]\frac{a^*}{1-|a|^2}\}-2\Re\{\sum_{k=\vartheta}^{N+L-1}x^*[k]x[k+N]\frac{a^*}{1-|a|^2}\})] \end{aligned} \tag{5.13}$$

Combining this with (5.11), we conclude that the exact PDF which globally characterize the observation vector can be expressed as

$$\begin{aligned} p(\mathbf{x},\varepsilon,\vartheta) &= p_1(\mathbf{x},\varepsilon,\vartheta)(U[\vartheta-1]-U[\vartheta-N-1]) \\ &+ p_2(\mathbf{x},\varepsilon,\vartheta)(U[\vartheta-N-1]-U[\vartheta-N-L+1]) \end{aligned} \tag{5.14}$$

where $U[n]$ is the discrete time unit step function. We conclude this section by noting that the PDF given in [37] is only the first term of the PDF given in (5.14).

5.4 Maximum Likelihood Estimator

In this section, a ML estimator for joint estimation of carrier frequency and symbol timing error is presented. The ML estimate for the unknown vector $[\varepsilon,\vartheta]$ is defined to be the vector $[\varepsilon_{ML},\vartheta_{ML}]$ that maximizes the $p(\mathbf{x},\varepsilon,\vartheta)$ for fixed realization of the random vector $\mathbf{x}$ as expressed by

$$[\varepsilon_{ML},\vartheta_{ML}] = \arg\max_{\varepsilon,\vartheta} p(\mathbf{x},\varepsilon,\vartheta)$$

The maximization is performed over the entire span of the estimation vector $[\varepsilon,\vartheta]$.

By taking the derivative of the likelihood function given in (5.14), it can be shown that the joint ML estimation of ϑ and ε becomes

$$\vartheta_{ML} = \arg\max_{\vartheta} T_2(\vartheta)|a|+2|T_1(\vartheta)|$$

$$\varepsilon_{ML} = \frac{-1}{2\pi}\angle T_1(\mathbf{x},\vartheta)$$

$$T_1(\mathbf{x},\vartheta) = \begin{cases} \sum_{k=\vartheta}^{L+\vartheta-1} x[k]x^*[k+N] & 1 \le \vartheta \le N \\ \sum_{k=0}^{\vartheta-N-1} x[k]x^*[k+N] + \sum_{k=\vartheta}^{N+L-1} x[k]x^*[k+N] & N+1 \le \vartheta \le N+L \end{cases} \tag{5.15}$$

$$T_2(\mathbf{x},\vartheta) = \begin{cases} \sum_{k=\vartheta}^{L+\vartheta-1} |x[k]|^2 + |x[k+N]|^2 & 1 \le \vartheta \le N \\ \sum_{k=0}^{\vartheta-N-1} |x[k]|^2 + |x[k+N]|^2 + & \\ \sum_{k=\vartheta}^{N+L-1} |x[k]|^2 + |x[k+N]|^2 & N+1 \le \vartheta \le N+L \end{cases} \tag{5.16}$$

The estimator proposed in [37] provides the Likelihood function of the observation vector over a finite range of timing offset parameter. Thus, the resultant ML estimator is suboptimal and obtains the ML estimate by maximizing the conditional likelihood function $p(\mathbf{x},\varepsilon,\vartheta|_1 \le \vartheta \le N)$.

To better visualize this difference, a typical realization of log-likehood measures for both suboptimum and proposed functions are plotted in Figure (5.2). In this experiment, the DFT block size (N) and CP (L) are assumed to be 64 and 8. Signal to noise ratio, frequency offset (ε) and timing offset (ϑ) are set to 25 dB, .01 and 70 samples respectively. The upper plot indicates the log-likelihood function for the suboptimum metric proposed in [37] and the lower plot provides the metric given in equation (5.14). By investigating these plots, it is inferred that the suboptimum metric achieves its maximum at $\vartheta = 48$. However, for the metric given in (5.14), the maximum is 70 which is exactly the unknown symbol timing error. Knowing the fact that during the startup and initialization of the receiver, symbol timing error is uniformly distributed between $\vartheta \in [1 \quad N+L]$, the ML estimator proposed in [37] results in considerable estimation error with probability of $L/(N+L)$.

ML estimators are asymptotically MVU, i.e. the variance of the estimator achieves that of MVU estimator as length of the observation vector goes to infinity. However, for

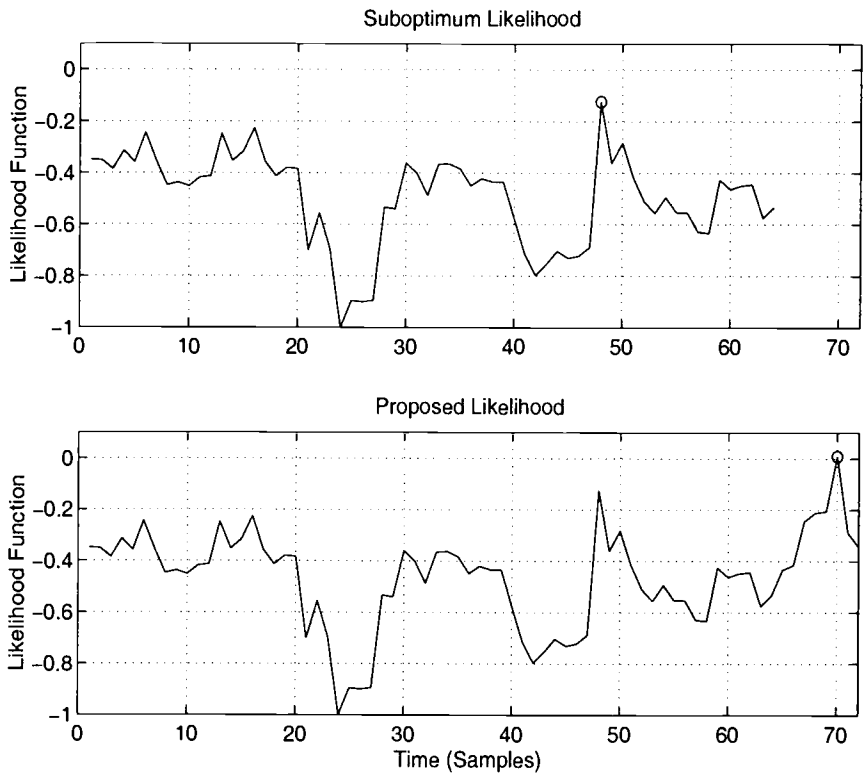


Figure 5.2: Comparison Between Realization of Log-likelihood Function for suboptimum and Proposed Method

many practical purposes, length of observation vector is constrained by physical limitations and can not be chosen arbitrarily long. In the subsequent section we investigate the MVU estimator which has the minimum variance among the set of cyclic based estimators.

5.5 MVU Estimator

In this section we try to find the MVU estimator by resorting to the theory of sufficient statistics [57]. The first step in deriving the MVU estimator is to obtain the sufficient statistic for the PDF given in (5.14). The sufficient statistic is known to be a

function of the OV , namely $T(\mathbf{x})$, such that the conditional PDF of the OV given $T(\mathbf{x})$ does not depend on the unknown estimation parameters $[\varepsilon, \vartheta]$. Evaluating the sufficient statistic is a formidable task for the broad class of PDFs, however the Neyman-Fisher (NF) Factorization theorem can be used for identifying the potential sufficient statistics. According to this theorem if the PDF can be factored in the form $g(T(\mathbf{x}), \varepsilon, \vartheta)h(\mathbf{x})$ where g is a function depending on $\mathbf{x}$ only through $T(\mathbf{x})$ and $h(\mathbf{x})$ is a function depending only on $\mathbf{x}$, then $T(\mathbf{x})$ is a sufficient statistic for estimation of the parameters ε and ϑ. By reformulating the PDF given in 5.14 to

$$p(\mathbf{x}, \varepsilon, \vartheta) = e^{\frac{2\Re[e^{j2\pi\varepsilon}T_1(\mathbf{x},\vartheta)|a|] - |a|^2 T_2(\mathbf{x},\vartheta)}{2(1-|a|^2)(\sigma_s^2+\sigma_n^2)}} \; h_1(\mathbf{x})$$

it is straightforward to verify that there is a direct dependency between the parameter ϑ and the statistics $T_1(\mathbf{x}, \vartheta)$ and $T_2(\mathbf{x}, \vartheta)$. Based on this observation the NF theorem fails to provide a sufficient statistic for estimation of ϑ. However, for a deterministic realization of the parameter ϑ, we can factor the PDF into

$$p(\mathbf{x}, \varepsilon|_\vartheta) = e^{\frac{\Re[e^{j2\pi\varepsilon}T_1(\mathbf{x},\vartheta)|a|]}{(1-|a|^2)(\sigma_s^2+\sigma_n^2)}} \; h_2(\mathbf{x}) \tag{5.17}$$

Clearly then, $T_1(\mathbf{x}, \vartheta)$ forms a sufficient statistic for estimation of the parameter ε.

Next, we apply Rao-Blackwell-Lehmann-Scheffe (RBLS) theorem to find a MVU estimator. According to this theorem, If $\check{\varepsilon}$ is an unbiased estimator of ε and $T(\mathbf{x})$ is a sufficient complete statistic for ε then $\hat{\varepsilon} = E(\check{\varepsilon}|_{T(\mathbf{x})})$ is a valid, unbiased , MVU estimator of ε.

In applying the above theorem, we need to obtain an unbiased estimator of ε, say $\hat{\varepsilon}$, and determine the conditional expectation of this estimator given the statistic $T_1(\mathbf{x}, \vartheta)$. An appropriate candidate for the unbiased estimator of ε can be obtained from the statistical moments of the random vector $\mathbf{x}$. According to equation (5.2) the second moment of the random variable $x[k]$, with $k \in \Omega$, satisfies the following identity

$$E\{x[k]x^*[k+N]\} = \sigma_s^2 e^{-j2\pi\varepsilon} \tag{5.18}$$

Having this observation, we use the second moment estimator as an unbiased estimator for ε as given by

$$\check{\varepsilon} = \frac{\Im}{2\pi} \ln\{\frac{1}{L\sigma_s^2} \sum_{k=\vartheta}^{\vartheta+L-1} x[k]x^*[k+N]\} \tag{5.19}$$

In deriving the above estimator, we replaced $E(x[k]x^*[k+N])$ by its natural estimator $\frac{1}{L}\sum_{k=\vartheta}^{\vartheta+L-1} x[k]x^*[k+N]$. It is straightforward to verify that this estimator is unbiased as it satisfies the condition

$$E\{\check{\varepsilon}\} = \frac{\Im}{2\pi} \ln E\{\frac{1}{L\sigma_x^2} \sum_{k=\vartheta}^{\vartheta+L-1} x[k]x^*[k+N]\} = \varepsilon \tag{5.20}$$

Next, we obtain the conditional expectation of $\hat{\varepsilon}$ given the sufficient statistic $T_1(\mathbf{x}, \vartheta)$ as follows

$$\begin{aligned} \check{\varepsilon}_{\text{MVU}|_\vartheta} &= E(\check{\varepsilon}|_{T_1(\mathbf{x},\vartheta)}) = \frac{\Im}{2\pi} \ln E\{\frac{1}{L\sigma_x^2} \sum_{k=\vartheta}^{\vartheta+L-1} x[k]x^*[k+N]|_{T_1(\mathbf{x},\vartheta)}\} \\ &= \frac{1}{2\pi}\Im\{\ln \frac{T_1(\mathbf{x},\vartheta)}{L\sigma_x^2}\} \end{aligned} \tag{5.21}$$

It is important to emphasize that since the underlying PDF given in (5.14) belongs to the exponential family of PDFs, then the sufficient statistics $T_1(\mathbf{x}, \vartheta)$ forms a **complete** statistic for estimation of the parameter ε. Therefore, the mapping function obtain from applying RBLS theorem, namely $\ln T(x, \vartheta)$, is but *one* function of the statistic $T_1(\mathbf{x}, \vartheta)$ and no other estimator with the same statistic can result in a lower variance with respect to MVU estimator.

5.5.1 Cramer-Rao Lower Bound

It is known that under broad conditions, the variance of any unbiased estimator of a nonrandom parameter ε satisfies the CRLB as

$$var(\check{\varepsilon}_{\text{MVU}|_\vartheta}) \geq \frac{1}{I(\varepsilon)} \tag{5.22}$$

where $I(\varepsilon)$ is the Fisher information given by

$$I(\varepsilon) = -E[\frac{\partial^2 \ln pr(\mathbf{x}, \varepsilon|_\vartheta)}{\partial \varepsilon^2}] \tag{5.23}$$

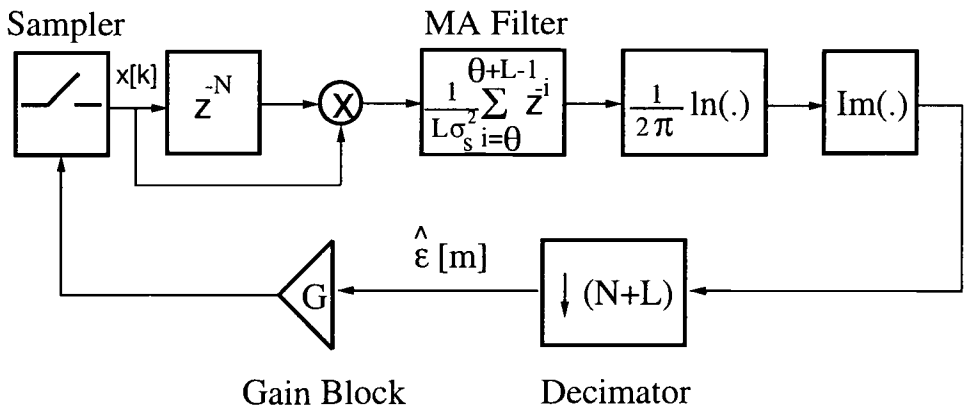


Figure 5.3: Closed Loop Offset Estimator

Substituting (5.17) into (5.23), after some algebraic manipulations, the CRLB of the MVU estimator becomes

$$var(\check{\varepsilon}_{\mathrm{MVU}|_\vartheta}) = \frac{(1-|a|^2)(\sigma_s^2+\sigma_n^2)}{(2\pi)^2|a|E\{T_1(\mathbf{x},\vartheta)\}} = \frac{(1+\frac{1}{\mathrm{SNR}})^2-1}{(2\pi)^2 L} \tag{5.24}$$

where $\mathrm{SNR} \triangleq \frac{\sigma_s^2}{\sigma_n^2}$ is the signal to noise ratio at the receiving end.

5.5.2 Closed Loop Performance

The frequency recovery loop for estimation of frequency offset is depicted in figure 5.3. The closed loop system is obtained by feeding back the information obtained from the estimator into the sampler block (boot strap). The sampler updates its frequency at the beginning of the each observation vector (every $(N+L)$ samples). To match the various sampling frequencies used in the system, a down sampler block is used prior to the sampler. Also, a gain block (G) is used to control the closed loop characteristic of the system (stability, settling time, noise sensitivity). According to figure 5.3, the frequency offset for the mth observation vector can be expressed as

$$\hat{\varepsilon}[m] = \frac{1}{2\pi}\ln\frac{1}{L\sigma_s^2}\sum_{i=0}^{L-1}\mathsf{x}_{(m-1)}[i]e^{-j\frac{2\pi\hat{\varepsilon}[m-1](\vartheta+i)}{N}G}\mathsf{x}^*_{(m-2)}[i]e^{j\frac{2\pi\hat{\varepsilon}[m-2](\vartheta+i)}{N}G}$$

$$= \frac{1}{2\pi}\ln\frac{1}{L\sigma_s^2}e^{j\frac{2\pi\hat{\Delta}\varepsilon[m-1](\vartheta)}{N}}\sum_{i=0}^{L-1}\mathsf{x}_{(m-1)}[i]\mathsf{x}^*_{(m-2)}[i]e^{-j\frac{2\pi\hat{\Delta}\varepsilon[m-1](i)}{N}} \quad (5.25)$$

where $\Delta\hat{\varepsilon}[m-1] \triangleq (\varepsilon[m-1]-\varepsilon[m-2])G$ and $\mathsf{x}_m[i] \triangleq x[m(N+L)+\vartheta+i]$. The term inside the sum is a stochastic quantity and does not have a closed form expression. However, for reasonably high signal to noise ratio it can be well approximated by its expected value ($E[x_{(m-1)}[i]x^*_{(m-2)}[i]] = \sigma_s^2$). Therefore, the expression inside the sum can be written as

$$\sum_{i=0}^{L-1}e^{-j\frac{2\pi\Delta\hat{\varepsilon}[m-1](i)}{N}} = \frac{1-e^{-j\frac{2\pi\hat{\delta}\varepsilon[m-1]L}{N}}}{1-e^{-j\frac{2\pi\Delta\hat{\Delta}\varepsilon[m-1]}{N}}} = e^{-j\frac{2\pi\Delta\hat{\varepsilon}[m-1](L-1)}{2N}}\frac{sin(\frac{\Delta\hat{\varepsilon}[m-1]L}{2N})}{sin(\frac{\Delta\hat{\varepsilon}[m-1]}{2N})}$$

$$\overset{\hat{\varepsilon}[m-1]\to 0}{\longrightarrow} e^{-j\frac{2\pi\Delta\hat{\varepsilon}[m-1](L-1)}{2N}}L \quad (5.26)$$

Substituting (5.26) into (5.25), after some algebraic manipulations, the frequency offset of m'th OV becomes

$$\hat{\varepsilon}[m] = -(\hat{\varepsilon}[m-1]-\hat{\varepsilon}[m-2])\frac{G}{N}(\vartheta+\frac{L-1}{2}) \quad (5.27)$$

The above equation represents a second order finite difference (FD) system in which its dynamic can be obtained from solving the following equation

$$\hat{\varepsilon}[m]+\beta\hat{\varepsilon}[m-1]-\beta\hat{\varepsilon}[m-2]=0$$

where $\beta \triangleq \frac{G}{N}(\vartheta+\frac{L-1}{2})$. Clearly, the solution to the above DE is the form of

$$\hat{\varepsilon}[m] = c_1(z_1)^m + c_2(z_2)^m$$

where $z_{1,2} \triangleq \frac{-\beta\overset{+}{-}\sqrt{\beta^2+4\beta}}{2}$ are two dynamical modes of the system. The smaller root (negative) results in a high frequency oscillation in the frequency offset estimate. However, as we show in the computer simulation, this term is filtered out by the moving average filter . To assure stability, the gain block should be set such that both poles lie inside the unit circle.

$$|\beta_{max}+\sqrt{\beta_{max}^2+4\beta_{max}}| \le 2 \quad (5.28)$$

where

$$\beta_{max} \triangleq \max_{\vartheta} \frac{G}{N}(\vartheta + \frac{L-1}{2}) = \frac{G}{N}(N + L + \frac{L-1}{2}) \tag{5.29}$$

5.6 Moment Estimator

As discussed earlier, when the timing offset parameter is not known to the receiver or if the the noise PDF differs from Gaussian distribution, finding optimum estimator (ML, MVU, CRLB) may not be an easy task. However, as we show next, there exists a moment estimator which provides a consistent estimate for estimation of frequency offset regardless of noise distribution and timing offset values. Although there is optimality criteria associated with this estimator , it can be globally used as an initial estimate for other estimators such as ML estimator.

Consider a sequence of first $N + L$ samples of vector $\mathbf{x}$. Using equation (5.2), autocorrelation of kth entry of this vector satisfies the following identity

$$r_{xx}[N] = \begin{cases} \sigma_x^2 e^{-j2\pi\varepsilon} & k \in \Omega \\ 0 & k \notin \Omega \end{cases} \tag{5.30}$$

Using Base Rule, the expected value of the above function (with respect to parameter k) can be expressed as

$$E_k[r_{xx}[N]] = \sigma_s^2 \exp^{-j2\pi\varepsilon} pr(k \in \Omega) + 0pr(k \notin \Omega) = \frac{L}{L+N}\sigma^2 \exp^{-j2\pi\varepsilon} \tag{5.31}$$

Substituting the Nth autocorrelation lag with its natural estimator , the moment estimator for frequency offset under uncertain timing offset can be found as

$$\hat{\varepsilon}_{mom} = \frac{\Im}{2\pi}\{\ln \mathbf{T}_3(\mathbf{x})\} \tag{5.32}$$

where the statistic $\mathbf{T}_3(\mathbf{x})$ is defined as

$$\mathbf{T}_3(\mathbf{x}) = \frac{1}{L\sigma_s^2} \sum_{k=0}^{N+L-1} x[k]x^*[k+N] \tag{5.33}$$

Statistical assessment of moment estimator is a formidable task for entire span of SNR. However, for relatively high SNRs, the random OV is heavily concentrated about its mean. Using the statistical linearization, we can use a first-order Taylor expression of the estimator about its mean to obtain the variance of estimate. In doing so, we substitute the random variable x in 5.33 with the expression given in (5.1) and obtain

$$\begin{aligned}\mathbf{T}_3([\mathbf{x}]) &= f([\mathbf{s},\mathbf{w}]) = \frac{\Im}{\ln}\{L\sigma_s^2 \sum_{k=0}^{N+L-1} (s[k-\vartheta]\exp^{j\frac{2\pi k\varepsilon}{N}} \\ &+ w[k])(s^*[k-\vartheta+N]\exp^{-j\frac{2\pi(k+N)\varepsilon}{N}} + w^*[k+N])\} \end{aligned} \tag{5.34}$$

where the signal ($\mathbf{s}$) and noise ($\mathbf{w}$) vector are defined as

$$\mathbf{s} \triangleq [s[0]\cdots s[N+L-1]] \tag{5.35}$$

$$\mathbf{w} \triangleq [w[0]\cdots w[N+L-1]] \tag{5.36}$$

By virtue of the above equation, the expected value of OV for a fixed realization of signal vector $\mathbf{s}$ would be $E[\varepsilon_{mom}] = f(\mathbf{s})$. We then perform a first order Taylor expansion of $f([\mathbf{s},\mathbf{w}])$ about the point $\mathbf{s} = E_w[\mathbf{x}]$ to yield

$$\hat{\varepsilon}_{mom} = f([\mathbf{s},0]) + \nabla_{\mathbf{w}} f([\mathbf{s},\mathbf{0}])^*\mathbf{w} = f([\mathbf{s},\mathbf{0}]) + \sum_{n=0}^{L+N-1} \frac{\partial f([\mathbf{s},\mathbf{w}])}{\partial w[n]}|_{\mathbf{w}=0} w[n] \tag{5.37}$$

Taking the derivative of (5.34) with respect to $w[n]$ and setting $\mathbf{w} = 0$, we obtain

$$\frac{\partial h}{\partial w[n]}|_{\mathbf{w}=0} = \frac{1}{2\pi}\Big(\sum_{i=0}^{N+L-1} s[i-\vartheta]s^*[i-\vartheta+N]e^{-j2\pi\varepsilon}\Big)^{-1} s^*[n+N-\vartheta]e^{-j\frac{2\pi\varepsilon(n+N)}{N}}$$

The second term in 5.37 represents the contribution of noise in the estimate. Knowing that noise samples are iid with power of σ_w^2, variance of estimate can be obtain from

$$var(\varepsilon) \simeq \frac{\sum_{n=0}^{L+N-1} |s[n+N-\vartheta]|^2\sigma_w^2}{\Big(2\pi \sum_{i=0}^{N+L-1} s[i-\vartheta]s^*[i-\vartheta+N]\Big)^2} \tag{5.38}$$

For sufficiently large block lengths (N), the above term can be well approximated as

$$var(\varepsilon) \simeq \frac{(N+L)\sigma_s^2\sigma_w^2}{(2\pi L\sigma_s^2)^2} = \frac{(N+L)}{(2\pi L)^2\, SNR} \tag{5.39}$$

In the next section we use the resemblance between the estimators to propose a model to cast all the proposed estimators into a unified structure.

5.7 Unified Structure

In this section we shall show how the proposed estimators can be classified into a single unified structure. This provides a unique framework for analysis of the proposed estimators. Moreover, it allows us to investigate the effect of symbol timing error in the estimation of carrier offset for each individual scheme.

5.7.1 MVU and Moment Estimator

Comparing the moment estimator given in (5.21) to the MVU estimator in (5.32) reveals some similarities in the structure of the estimators. Clearly, both moment and MVU estimators use the same mapping function, namely log function, to project the data statistics into the estimation domain. The only difference is in the form of statistics used for each scheme. Figure 5.4 depicts the block diagram of these two estimators. As shown in the figure, both estimators obtain the statistic by correlating the samples with their Nth delayed samples. This operation is performed by using a moving average (MA) filter in the structure of estimators. However, MVU and moment estimators use different upper and lower bound for the MA filter. In moment estimator, the averaging is performed over the first $N+L$ samples of OV. This would remove the requirement of knowing the exact timing offset parameter in estimation of carrier frequency offset. However, the estimate obtained from using this scheme results in less accurate estimate (more variance) in comparison to MVU estimate. On the other hand, MVU estimator requires the knowledge of timing offset parameter which can be obtained from using pilot tones in the OFDM symbol. This would result in a lower bandwidth efficiency: the price which is paid for improvement in the performance.

5.7.2 MVU and ML estimator

Although the resemblance between MVU and ML estimators may not be as evident as that of MVU and moment estimator, it can be shown that ML estimator can also be classified into the same family. Knowing the fact that $\Im\{logT_1(\mathbf{x},\vartheta)\} = \angle T_1(\mathbf{x},\vartheta)$ the ML estimator can be expressed as

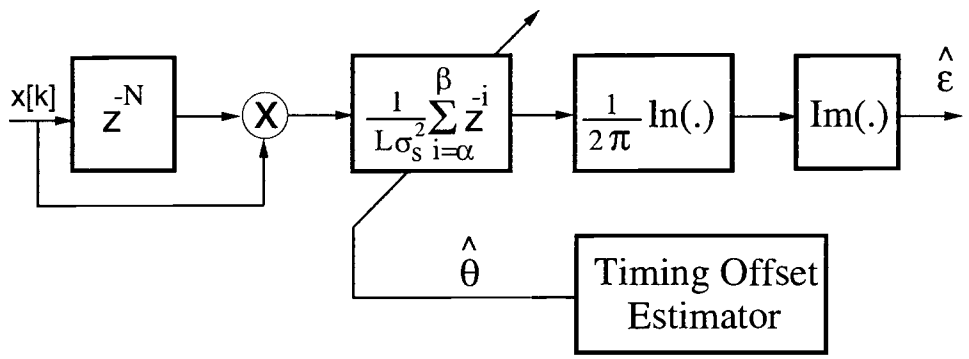
$$\varepsilon_{ML} = \frac{-1}{2\pi}\Im\{\sum_{k=\alpha}^{\beta} x^*[k]x[k+N]\} \tag{5.40}$$

where the parameters α and β are functions of ϑ_{ML} and can be obtained from 5.15. Thus, the ML estimator falls into the same family of estimators. It is noticeable that ML estimator provides the upper and lower bound of the moving average filter by extracting the timing parameter from the likelihood function. Although ML has the advantage of exploiting the entire bandwidth by removing the requirement for having pilot tones, the symbol timing estimate obtained from ML scheme has larger confident interval. This may result in a considerable performance degradation in comparison to the pilot based schemes.

5.8 Discussion and Simulation

In this section we shall perform computer simulation to assess the performance of proposed estimator for synchronization of OFDM systems. As it is usually the case, we shall choose the variance of estimator as a performance measure through our study. The simulation parameters used are typical of the environments. More specifically, the chosen FFT size (N) for OFDM is 64. Unless specified, the length of cyclic prefix is $L = 8$. Unless specified we choose the signal to noise ratio level to be 20 dB. Also, in exception of closed loop system, the carrier frequency offset is set to $\varepsilon = .01$. We carry out a Monte Carlo simulation to evaluate the performance of proposed estimators.

In doing so, we start with a comparison between performance of proposed ML estimator with the suboptimum ML estimator given in [37] over the range of timing



Estimator	α	β
ML	θ_{ML}	θ_{ML}+L-1
MVU	θ	θ+L-1
Moment	0	N+L-1

Figure 5.4: Unified Structure for Class of Cyclic Bases Estimators

offset parameter ($\vartheta \in [1 \quad N+L]$). Since the estimation parameter is varying itself, we use a normalized variance $E[(\varepsilon - \hat{\varepsilon})^2/(\varepsilon)^2]$ as a performance measure for timing offset estimator. Figures 5.5 and 5.6 compares the variance of the proposed ML estimator against the suboptimum ML estimator. As expected, the suboptimum ML estimator exhibits anomalous statistical behaviour over the range of ($\vartheta \in [N+1N+L]$).

Figure shows the performance of MVU estimator for frequency offset estimation under complete knowledge of timing offset error. A careful examination of the variances reveals that the gap between MVU estimator and CRLB tends to zero as SNR increases. Also as illustrated in the figures, the departure from CRLB happens rapidly as SNR goes bellow a threshold. The threshold also depends on the length of CP and is moved toward

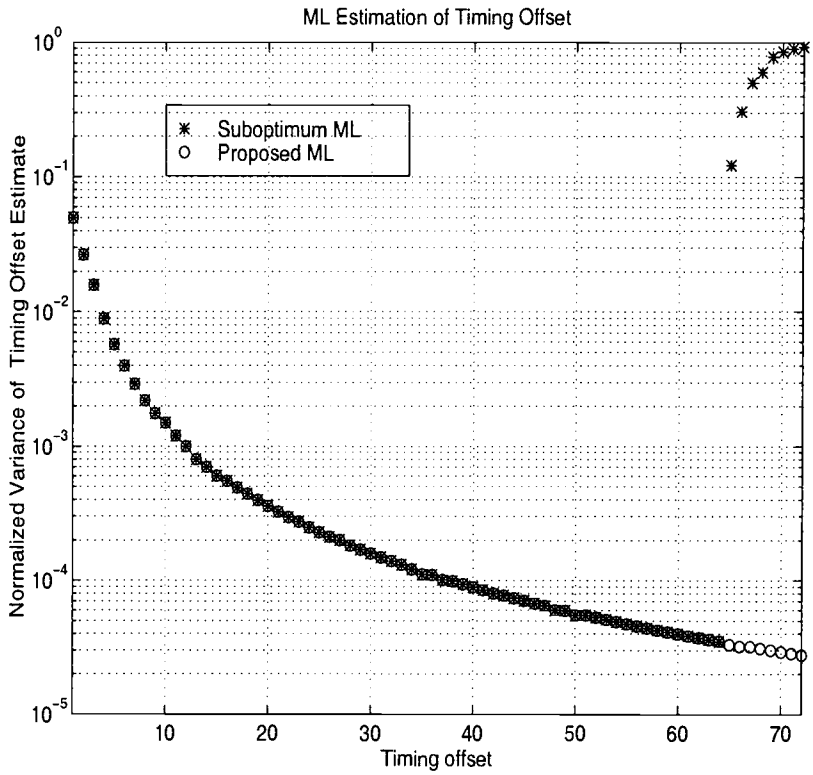


Figure 5.5: Comparison Between Proposed and Suboptimum ML Timing Offset Estimator

lower SNRs as L increases. This can be justified in terms of having more observation samples in estimating the unknown parameter. The choice of cyclic prefix length L represents a tradeoff between data rate reduction and performance (lower variance). Increasing L brings the performance of MVU estimator closer to the CRLB, nevertheless, could result in a considerable data rate reduction due to redundancy introduced by CP.

To illustrate the dynamical behaviour of the closed loop MVU estimator, we have plotted in Figure (5.8), the frequency offset parameter of the closed loop MVU estimator together with the analytical derivation given in equation 5.28. It is clear that the simulation result very closely resembles the analytical model, thus consolidates the approximate model of the closed loop system.

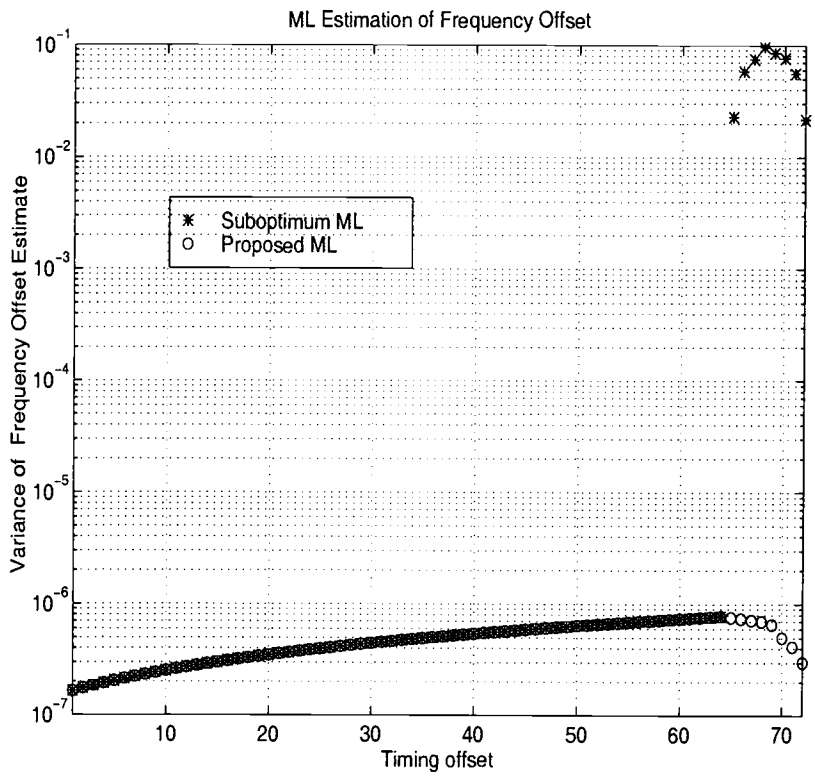


Figure 5.6: Comparison Between Proposed and Suboptimum ML Frequency Offset Estimator

5.9 Conclusion

We have proposed a class of non data-aided cyclic based robust estimators for frequency offset estimation of MC systems. We determined a likelihood function for joint estimation of symbol timing error and carrier frequency offset in MC systems. As a result, it is used for deriving a maximum likelihood estimator for the joint estimation problem. The proposed estimator outperforms the previously proposed suboptimum ML estimator. We also used the concept of sufficient statistics to obtain the minimum variance unbiased estimate of the frequency offset under certain knowledge of timing offset error. In doing so, we apply the factorization and RBLS theorems to identify the sufficient statistic and

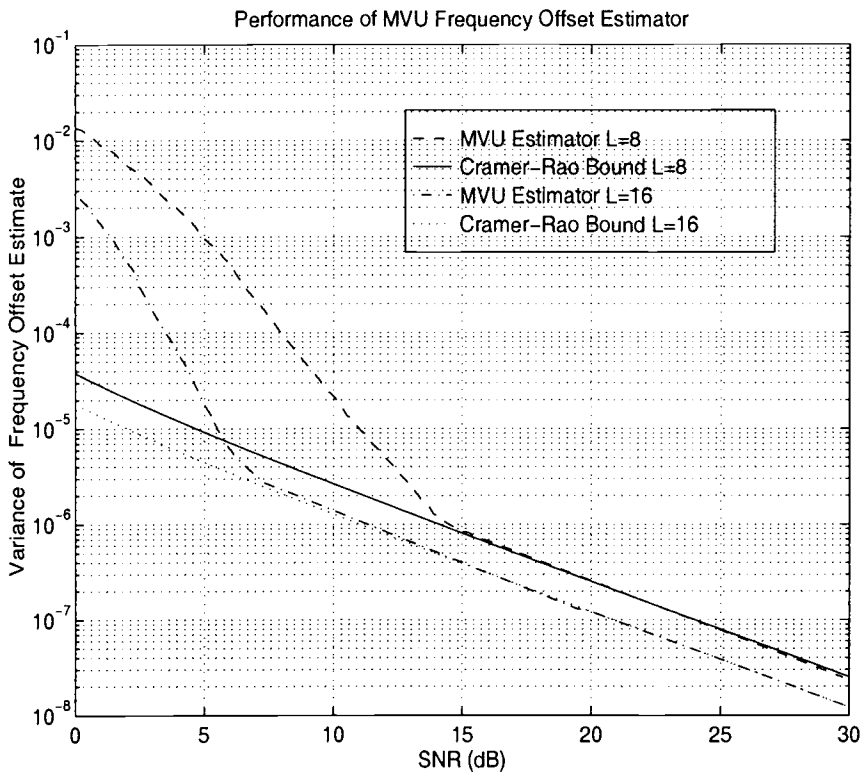


Figure 5.7: Performance of MVU Estimator as a Function of SNR

appropriate mapping functions. It is shown that there is but one function of the sufficient statistics which results in the minimum variance estimate among the possible class of estimators. Also, a moment estimator is proposed to obtain the consistent estimate of carrier offset under uncertain symbol timing error. The moment estimator does not rely on any probabilistic assumptions. Thus, its performance is insensitive to the distribution of the additive noise. Finally a unified structure for modeling all the proposed estimators is proposed.

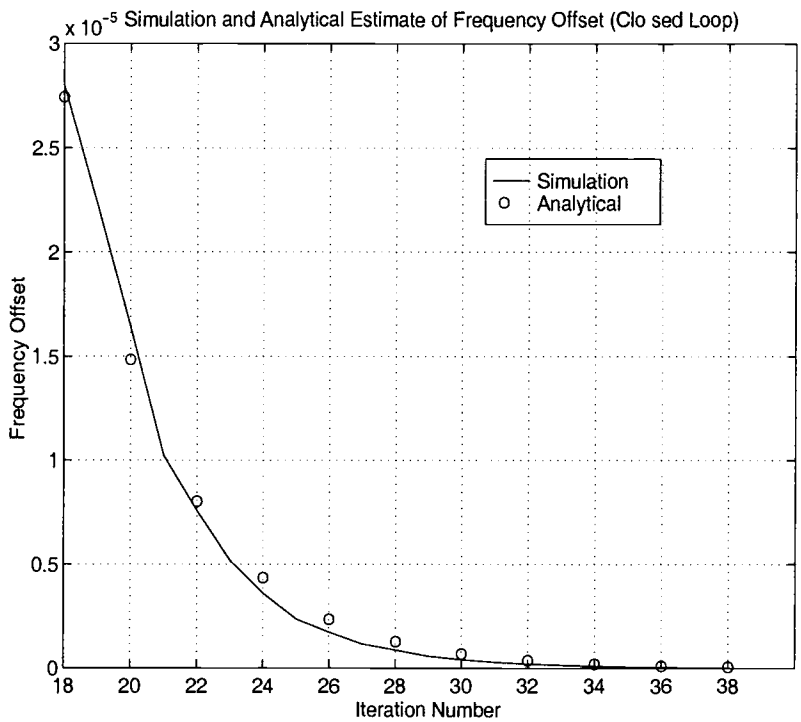


Figure 5.8: Closed Loop Performance of MVU estimator

6. CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

Since the mid 1980's, multicarrier systems have been a viable solution for data transmission over bandlimited channels. Their immunity to nonflat frequency response channels and also the inherent parallelism in their structure, make these systems a competitive candidate in high speed data transmission. Nevertheless, this research area is far from being mature as evidenced by the large number of publications devoted to this subject around the globe.

Due to the high computational complexity involved with equalization of multicarrier modulation, the practical realization of these systems is still a great challenge. Mathematically, the minimum mean square error equalization can be viewed as the problem of computing extreme (minimum) eigenvector of a positive definite matrix. The standard approach for this purpose requires inverting a high dimensional Hessian matrix. Due to the sensitivity of this operation to the truncation and quantization errors and because of the relatively high floating point operations involved with this scheme, the existing digital signal processing blocks are not able to handle this operation in real time. Chapter 3 of this thesis focused on the low complexity implementation of minimum mean square error equalizers for discrete multitone systems. This method has the advantage of being conceptually simple and practical to implement. The robust performance to quantization and truncation effects is inherited from the characteristic of Rayleigh minimization approach. On the other hand, the proposed algorithm is characterized by its FFT based structure, thus the existing infrastructure of multicarrier system can be used to reduce the hardware complexity. The proposed algorithm exploits asymptotic equivalence of Toeplitz and circulant matrices to estimate Hessian matrix of a quadratic form. It was shown that the Hessian matrix exhibits a specific structure. As a result, when combined with the Rayleigh

minimization algorithm, it provides an efficient method to obtain the global minimum of the constrained optimization problem. A salient feature of this algorithm is that extreme eigenvector of the Hessian matrix can be obtained without direct computation of the matrix. The proposed algorithm performs well provided only that the length of equalizer filter is large enough for the asymptotic equivalence of Toeplitz and circulant matrices.

In chapter 4 we addressed the problem of optimum equalization of multicarrier systems. It is known that MMSE is not the optimum criterion in conjunction with multicarrier systems and the optimum setting for the equalizer is obtained at expense of solving a highly complex constrained optimization problem. Motivated by this shortcoming, a new framework for training the time domain equalizer subject to maximum system performance was proposed. The proposed approach exploits an appropriate transformation to shape the constraint region into an identifiable convex boundary, namely a canonical ellipsoid. Due to the canonical property of the constraint set, projection onto this set is obtained efficiently through projection onto convex set (POCS) method. Armed with this, the gradient projection method is applied to obtain the search direction in each iteration of the algorithm. Due to versatility of this method, it is then generalized to two classes of systems namely, unit tap and linear phase filters. Also an upper bound on the performance of equalizer is obtained. As an assessment for the performance of equalizer, it is then applied to equalization of discrete multitone systems for asynchronous digital subscriber line (ADSL) application.

The primary contribution of chapter 5 has been to introduce a new family of estimators for joint symbol timing and carrier frequency synchronization of orthogonal frequency division modulation scheme. Essentially, the approach uses the periodic property of the cyclic sets to reveal the synchronization parameters in the likelihood function of the received vector. The likelihood function is then used as a basis for deriving three classes of efficient estimators namely, maximum likelihood (ML) , minimum variance unbiased (MVU), and moment estimators. The ML estimator obtains the solution for the joint estimation problem by maximizing the likelihood function of the observation vector. On the other hand, MVU estimator exploits the property of sufficient statistic to obtain an

estimate of carrier frequency offset subject to minimum variance criterion. The most appealing aspect of these estimators is that they are are classified in a single unified structure. This would facilitate the analysis of cyclic based estimators within a common framework.

6.2 Recommendations for Future Investigation

It should be stated that although the work presented in this thesis makes some contributions to the area of multicarrier systems [58, 59, 60, 61, 62, 63], there remain plenty of challenges requiring further research work.

In the MMSE equalization, there are two important areas where more work needs to be done. In the context of eigenvalue estimation, few more approaches might be considered. For example one may use Lanczos or subspace methods to provide an estimate of extreme eignevectors. The second area is that various aspects of vector norms beside l_2 norm may also be considered. Although l_2 norm result in a mathematically tractable solution, one may consider l_1 or l_∞ norms in minimizing the residual error of impulse response shortening problem. Thus, rather than minimizing power of noise sequence, sum of absolute values of noise error or even maximum element of noise sequence can be minimized (minimax). The interesting question is " which norm results in the optimum performance in conjunction with multicarrier systems?" There are also other techniques that one can choose for the impulse response shortening problem. For example, instead of using an FIR filter as TEQ one may consider a block-digital filter as an equalizer [64]. Then, one interesting question is that " what is the best way to realize such a filter in terms of reducing the overall complexity while maintaining reasonable system sensitivity?"

The fourth chapter of this thesis, dealing with optimum equalization of multicarrier systems, has provided a new method for training the equalizer. With this approach, one can also combine the optimum loading technique for joint equalization and bit allocation of the multicarrier systems. By including the effect of bit loading on the optimum equalization, a question might be " How to jointly optimize the performance of blocks such that

the data rate becomes maximized?" An effective solution to this problem is necessary to benefit fully from maximum data rate. There are also several interesting issues regarding to the equalization of multicarrier systems over time varying dynamic fading channels [65, 66]. In particular, due to the dynamical variation of channel impulse response , sophisticated channel identification methods are required to compensate this shortcoming [67, 68]. At present most system designers use complicated estimation techniques for this purpose [69]. One possible course of action would be to exploit the temporal coding to transform the time-varying channel into an asymptotic AWGN channel [70]. This would transform the fading channel into highly dispersive static ISI channel. Thus, one can use multicarrier systems to mitigate the effect of ISI in the resulting system.

BIBLIOGRAPHY

1. L. J. Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Transactions on Communications*, vol. 33, pp. 665–675, July 1985.

2. J. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Communication magazine*, pp. 5–14, May 1990.

3. J. W. Lechleider, "High bit rate digital subscriber lines: A review of HDSL progress," *IEEE Journal on Selected Area in Communication*, vol. 9, pp. 769–784, August 1991.

4. K. Sistanizadeh, P. Chow, and J. M. Cioffi, "Multitone transmission for asymetric digital subscriber lines (adsl)," pp. 756–760, International Conference on Communication, May 1993.

5. "Radio broadcasting systems; digital audio broadcasting (DAB) to mobile portable and fixed receivers," ETS 300 401, ETSI-European Telecommunications standards institute, Valbonne,France, February 1995.

6. T. Willink and W. H., "Optimization and performance evaluation of multicarrier systems," *IEEE Transactions on Information Theory*, vol. 43, pp. 426–440, March 1997.

7. N. Al-Dhahir and J. M. Cioffi, "A bandwidth-optimized reduced-complexity equalized multicarrier transceiver," *IEEE Transactions on Communications*, vol. 45, pp. 948–956, May 1997.

8. N. Al-Dhahir and J. M. Cioffi, "Optimum finite-length equalization for multicarrier transceivers," *IEEE Transactions on Communications*, vol. 44, pp. 54–63, Jan. 1996.

9. T. Pollet and M. Peeters, "Synchronization with DMT modulation," *IEEE Communication magazine*, pp. 80–86, April 1999.

10. I. Kalet and S. Shamai, "On the capacity of twisted pair : Gaussian model," *IEEE Transactions on Communications*, vol. 38, pp. 379–383, March 1990.

11. Proakis, *Digital Communications*. Prentice Hall, second ed., 1995.

12. A. Oppenheim, R. R. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*. New Jersy: Prentice Hall, second ed., 1999.

13. J. M. Cioffi, G. Dudevoir, M. Eyuboglu, and G. D. ForneyJr., "Minimum mean-square error decision feedback equalization and coding– parts I and II," *IEEE Transactions on Communications*, vol. 43, pp. 2582–2604, October 1995.

14. I. Kalet, "The multitone channel," *IEEE Transactions on Communications*, vol. 37, pp. 119–124, February 1989.

15. T. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley and Sons, Inc., 1991.

16. Hughes-Hartogs, "Emsemble modem structure for imperfect transmission media," US patent No. 4833706.

17. P. Chow, J. M. Cioffi, and J. A. C. Bingham, "A practical multitone transceiver loading algorithm for data transmission over spectrally shaped channels," *IEEE Transactions on Communications*, vol. 43, pp. 773–775, February, March, April 1995.

18. R. F. H. Fischer and J. B. Huber, "A new loading algorithm for discrete multitone transmission," p. 724, IEEE global telecommunications conference, 1996.

19. A. Leke and J. M. Cioffi, "A maximum rate loading algorithm for discrete multitone modulation systems," p. 1514, IEEE global telecommunications conference, 1997.

20. S. Kasturia, J. Aslanis, and J. M. Cioffi, "Vector coding for partial response channels," *IEEE Transactions on Information Theory*, vol. 36, pp. 741–761, July 1990.

21. P. Melsa, R. C. Younce, and C. E. Rohrs, "Impulse response shortening for discrete multitone transceivers," *IEEE Transactions on Communications*, vol. 12, pp. 1662–1672, December 1996.

22. K. Kammeyer, "Time truncation of channel responses by linear filtering: A method to reduce the complexity of viterbi equalization," *Journal of Signal Processing*, vol. 48, no. 5, pp. 237–243, 1994.

23. S. B. Weinstein and P. M. P. Ebert, "Data transmission by frequency division multiplexing using the discrete fourier transform," *IEEE Transactions on Communications*, vol. 19, pp. 628–634, October 1971.

24. B. Hirosaki, "An analysis of automatic equalizers for orthogonally multiplexed qam systems," *IEEE Transactions on Communications*, vol. 28, pp. 73–83, January 1980.

25. J. S. Chow, J. C. Tu, and J. M. Cioffi, "A discrete multitone transceiver system for HDSL applications," *IEEE Journal on Selected Area in Communication*, vol. 9, pp. 895–907, Aug. 1991.

26. J. Chow and J. M. Cioffi, "A cost-effective maximum likelihood receiver for multicarrier systems," (Chicago), pp. 948–952, International Conference on Communications, June 1992.

27. N. Al-Dhahir and J. M. Cioffi, "Efficiently computed reduced-parameter input-aided mmse equalizers for ML detection: A unified approach," *IEEE Transactions on Information Theory*, vol. 42, pp. 903–915, May 1996.

28. I. Lee and J. M. Cioffi, "A fast computation algorithm for the decision feedback equalizer," *IEEE Transactions on Communications*, vol. 43, pp. 2742–2749, Nov. 1995.

29. I. Lee, J. Chow, and J. M. Cioffi, "Performance evaluation of a fast computation algorithm for the dmt in high-speed subscriber loop," *IEEE Journal on Selected Area in Communication*, vol. 13, pp. 1564–1570, December 1995.

30. D. Pal, G. Iyengar, and C. J. M., "A new method of channel shortening with applications to discrete multitone systems," pp. 763–768, International conference on communications, 1998.

31. P. A. Regalia, "An unbiased equation error identifier and reduced-order approximations," *IEEE Transactions on Signal Processing*, vol. 42, pp. 1397–1412, May 1994.

32. P. H. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Transactions on Communications*, vol. 42, pp. 2908–2913, October 1994.

33. T. Pollet, M. V. Bladel, and M. Moenelaey, "BER sensitivity of OFDM systems to carrier frequency offset and wiener phase noise," *IEEE Transactions on Communications*, vol. 43, pp. 191–193, February/March/April 1995.

34. H. Liu and U. Tureli, "A high-efficiency carrier estimator for OFDM communications," *IEEE communications letters*, pp. 104–106, April 1998.

35. H. Roh, K. Cheun, and J. Park, "An MMSE fine carrier frequency synchronization for OFDM systems," *IEEE Transactions on Consumer Electronics*, vol. 43, p. 761, August 1997.

36. F. Daffara and A. Chouly, "Maximum likelihood frequency detectors for orthogonal multicarrier systems," *IEEE Transactions on Communications*, pp. 766–771, May 1993.

37. J. J. van de Beek, M. Sandell, and P. Borjesson, "ML estimation of timing and frequency offset in OFDM systems," *IEEE Transactions on Communications*, vol. 45, pp. 1800–1805, July 1997.

38. X. Yang, T. Sarkar, and E. Arvas, "A survey of conjugate gradient algorithms for solution of extreme eigen-problems of a symmetric matrix," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 37, pp. 1550–1555, October 1989.

39. Z. Fu and E. Dowling, "Conjugate gradient eigenstructure tracking for adaptive spectral estimation," *IEEE Transactions on Signal Processing*, vol. 43, pp. 1151–1160, May 1995.

40. T. Sarkar and X. Yang, "Application of the conjugate gradient and steepest descent for computing the eigenvalues of an operator," *IEEE Transactions on Signal Processing*, vol. 17, pp. 31–38, 1989.

41. R. M. Gray, "On the asymptotic eigenvalue distribution of toeplitz matrices," *IEEE Transactions on Information Theory*, vol. 18, pp. 725–730, November 1972.

42. R. M. Gray, "Toeplitz and circulant matrices: A review," tech. rep., Information System Labratory, Department of Electrical Engineering Stanford University, Stanford, California, 1998.

43. N. Al-Dhahir and J. M. Cioffi, "Stable pole-zero modeling of long FIR filters with application to mmse-dfe," *IEEE Transactions on Communications*, vol. 45, pp. 508–513, May 1997.

44. J. M. Cioffi, *Course notes for EE379C, Stanford University, Ca.* 1998.

45. A. Peled and A. Ruiz, "Frequency domain data transmission using reduced computational complexity algorithms," *International Conference on Acoustics, Speech and Signal Processing*, pp. 964–967, April 1980.

46. J. Chow, J. M. Cioffi, and J. A. Bingham, "Equalizer training algorithms for the multicarrier modulation systems," (Geneva), pp. 761–765, International Conference on Communications, May 1993.

47. P. Combettes, "The foundation of set theoretic estimation," *Proceedings of the IEEE*, vol. 81, pp. 182–208, February 1993.

48. M. W. Inc., "*MATLAB* high-performance numeric computation and visualization," Natwick, Massachusetts, 1992.

49. D. P. Bertsekas, *Nonlinear Programming.* Belmont, Mass.: Athena Scientific, 1995.

50. H. Stark and Y. Yang, *Vector Space Projection.* New York: John Wiley and Sons, first ed., 1998.

51. E. Casas and C. Leung, "OFDM for data communication over mobile radio channels-part I: Analysis and experimental results," *IEEE Transactions on Communications*, vol. 30, pp. 783–793, May 1991.

52. W. Y. Zou and Y. Wu, "COFDM: An overview," *IEEE transactions on Broadcasting*, vol. 41, pp. 1–8, March 1995.

53. L. Wei and C. Schegel, "Synchronization requirement for multi-user OFDM on satellite mobile and two path rayleigh-fadging channels," *IEEE Transactions on Communications*, vol. 43, pp. 887–895, February/March/April 1995.

54. W. D. Warner and C. Leung, "OFDM/FM frame synchronization fo rmobile radio data communication," *IEEE Transactions on Vehicular Technology*, vol. 42, no. 3, pp. 302–313, 1993.

55. T. M. Schmidl and D. C. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Transactions on Communications*, vol. 45, pp. 1613–1621, December 1997.

56. M. Russell and G. Stuber, "Interchannel interference analysis of OFDM in a mobile environment," pp. 820–824, IEEE Vehicular Technology Conference, July 1995.

57. M. S. Kay, *Fundamentals Of Statistical Signal Processing Estimation Theory.* Prentice Hall, first ed., 1993.

58. N. Lashkarian and S. Kiaei, "Fast algorithm for finite-length mmse equalizers with application to discrete multitone systems," (Phoenix, AZ), International Conference on Acoustics, Speech and Signal Processing, March 1999.

59. N. Lashkarian and S. Kiaei, "Optimum equalization of multicarrier systems via projection onto convex sets," (Vancouver, BC, Canada), International Conference on Communications, June 1999.

60. N. Lashkarian and S. Kiaei, "Globally optimum ML estimation of timing and frequency offsets in OFDM systems," submitted to international conference on communications, 2000.

61. N. Lashkarian and S. Kiaei, "MVU estimation of frequency offset in OFDM system," (Istanbul, Turkey), submitted to International Conference on Acoustics, Speech and Signal Processing, 2000.

62. N. Lashkarian and S. Kiaei, "Optimum equalization of multicarrier systems: A unified geometric approach," *submitted to IEEE Transactions on Communications.*

63. N. Lashkarian and S. Kiaei, "Class of cylic based estimators for frqeuency and timing synchronization of OFDM systems," *manuscript in preparation.*

64. Vaidyanathan, *Multirate Systems and Filter Banks.* Englewood Cliffs, NJ: Prentice-Hall, 1993.

65. S. K. Wilson, Khayate, R. E., and J. M. Cioffi, "16-QAM modulation with orthogonal fequency-division multiplexing in Rayleigh fading environment," *Vehicular Technology Conference*, pp. 1660–1664, June 1994.

66. R. Ziegler and J. M. Cioffi, "Estimation of time-varying digital radio channels," *IEEE transactions on vehicular technology*, pp. 134–151, May 1992.

67. S. Bulumulla, S. Kassam, and S. Venkatesh, "An adaptive reciver for OFDM in fading channels," pp. 1325–1328, International Conference on Communications, 1998.

68. Y. Li, L. Cimini, and N. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Transactions on Communications*, vol. 46, pp. 902–914, July 1998.

69. O. Edfors, M. Sandell, J. J. v. d. Beek, S. L. Wilson, and P. O. Borjesson, "OFDM channel estimation by singular value decomposition," *IEEE Transactions on Communications*, vol. 46, no. 7, p. 931, 1998.

70. G. W. Wornell, "Spread-response precoding for communication over fading channels," *IEEE Transactions on Information Theory*, vol. 42, pp. 488–501, March 1996.

71. A. Edelman and S. T. Smith, "On conjugate gradient-like methods for eigen-like problems," tech. rep., Department of Mathematics, Massachusets Institute of Technology, Cambridge, Ma., 1996.

APPENDICES

A. CIRCULANT MATRICES

A circulant matrix $\mathbf{C}$ is one having the form

$$\mathbf{C} = \begin{bmatrix} c_0 & c_1 & c_2 & \cdots & c_{n-1} \\ c_{n-1} & c_0 & c_1 & \cdots & c_{n-2} \\ \cdots & c_{n-1} & c_0 & \cdots & c_{n-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ c_1 & \cdots & \cdots & c_{n-1} & c_0 \end{bmatrix} \tag{A.1}$$

(A.2)

The matrix $\mathbf{C}$ it itself a special type of Toeplitz matrix. The eigenvalues ψ_k and the eigenvectors y_k of $\mathbf{C}$ are the solutions of

$$\mathbf{C}y_k = \psi_k y_k \tag{A.3}$$

in which the m^{th} eigenvalue is the m^{th} point of DFT of the first row of matrix $\mathbf{C}$. Any circulant matrix $\mathbf{C}$ can be decomposed as

$$\mathbf{C} = \mathbf{U}^*\mathbf{\Psi}\mathbf{U} \tag{A.4}$$

where

$$\mathbf{U} = \{y_0|y_1|\cdots|y_{n-1}\} = n^{-1}\{\exp\frac{-2\pi imk}{n};\ m,k = 0,1,\cdots,n-1\} \tag{A.5}$$

Also if $\psi_k \neq 0; k = 0,1,\cdots,n-1$, then $\mathbf{C}$ is nonsingular and

$$\mathbf{C}^{-1} = \mathbf{U}\Psi^{-1}U^* \tag{A.6}$$

B. RAYLEIGH MINIMIZATION APPROACH

Conjugate gradient method is a versatile algorithm that can be customized to provide balance between performance and computational complexity [71]. Use of the conjugate gradient algorithm in adaptive filtering applications allows flexibility that ranges from LMS-like performance and cost at one extreme, to RLS-like performance and cost at the other. In general, conjugate gradient methods are proposed and analyzed for purely quadratic problems. However, conjugate gradient technique can be successfully applied in many quadratic problems, where it can be argued that near the solution the error surface becomes approximately quadratic. With this assumption, the Rayleigh minimization algorithm exploits conjugate gradient method to obtain the minimum point of the second order approximation of the Rayleigh quotient, i.e.,

$$R(\mathbf{b}) = R(\mathbf{b}_1) + \frac{1}{2}(\mathbf{b} - \mathbf{b}_1)^* H(\mathbf{b}_1)(\mathbf{b} - \mathbf{b}_1) \tag{B.1}$$

where $H(\mathbf{b})$ is the Hessian of the Rayleigh quotient as expressed by

$$H(\mathbf{b}) = \nabla_{\mathbf{b}}^2 \left(\frac{\mathbf{b}^* \mathbf{R}_\Delta \mathbf{b}}{\mathbf{b}^* \mathbf{b}}\right) \tag{B.2}$$

- Initialization

  Compute the minimum eigenvalue estimate, residual error vector and descent direction vector as following;

$$\lambda^0 = \mathbf{b}^{0*} \mathbf{R}_\Delta \mathbf{b}^0 \tag{B.3}$$

$$\mathbf{r}^0 = \lambda^0 \mathbf{b}^0 - \mathbf{R}_\Delta \mathbf{b}^0 \tag{B.4}$$

$$\mathbf{p}^0 = \mathbf{r}^0 \tag{B.5}$$

$$\tag{B.6}$$

- Iteration

  Update the minimum eigenvector solution as

$$\mathbf{b}^{i+1} = \mathbf{b}^i + t^i \mathbf{p}^i \tag{B.7}$$

where the parameters are defined as

$$t^i = \frac{-B + \sqrt{B^2 - 4CD}}{2D} \tag{B.8}$$

$$D = P_b^i P_c^i - P_a^i P_d^i \quad B = P_b^i - \lambda^i P_d^i \quad C = P_a^i - \lambda^i P_c^i \tag{B.9}$$

$$P_a{}^i = \frac{\mathbf{p}^{i*}\mathbf{R}_\Delta \mathbf{b}^i}{\mathbf{b}^{i*}\mathbf{b}^i} \qquad P_b{}^i = \frac{\mathbf{p}^{i*}\mathbf{R}_\Delta \mathbf{p}^i}{\mathbf{b}^{i*}\mathbf{b}^i} \tag{B.10}$$

$$P_c{}^i = \frac{\mathbf{p}^{i*}\mathbf{b}^i}{\mathbf{b}^{i*}\mathbf{b}^i} \qquad P_d{}^i = \frac{\mathbf{p}^{i*}\mathbf{p}^i}{\mathbf{b}^{i*}\mathbf{b}^i} \tag{B.11}$$

$$\tag{B.12}$$

Update the minimum eigenvalue estimate, residual error vector and descent direction vector and normalize the vector as following;

$$\lambda^{i+1} = \frac{\mathbf{b}^{i+1*}\mathbf{R}_\Delta \mathbf{b}^{i+1}}{\mathbf{b}^{i+1*}\mathbf{b}^{i+1}} \qquad \mathbf{r}^{i+1} = \frac{\lambda^{i+1}\mathbf{b}^{i+1} - \mathbf{R}_\Delta \mathbf{b}^{i+1}}{\mathbf{b}^{i+1*}\mathbf{b}^{i+1}} \tag{B.13}$$

$$\mathbf{p}^{i+1} = \mathbf{r}^i + \beta^i \mathbf{p}^i \tag{B.14}$$

$$\beta^i = \frac{\mathbf{r}^{i+1*}\mathbf{R}_\Delta \mathbf{p}^i + (\mathbf{r}^{i+1*}\mathbf{r}^{i+1})(\mathbf{b}^{i+1*}\mathbf{p}^{i+1})}{\mathbf{p}^{i*}\mathbf{R}_\Delta \mathbf{p}^i - \lambda^{i+1}\mathbf{p}^{i*}\mathbf{p}^i} \tag{B.15}$$

$$\tag{B.16}$$

- Normalization

Normalize the minimum eigenvector estimate to provide the unit energy constraint

$$\mathbf{b}^{i+1} = \frac{\mathbf{b}^{i+1}}{||\mathbf{b}^{i+1}||} \tag{B.17}$$