國 立 交 通 大 學

電 機 與 控 制 工 程 學 系

碩 士 論 文

應用於高速數位用戶迴路之
時域等化器設計

Design of Time Domain Equalizer (TEQ) for VDSL System

研 究 生 ： 梁 禮 涵

指 導 教 授 ： 林 源 倍 博士

中華民國九十三年六月

應用於高速數位用戶迴路之時域等化器設計

Design of Time domain Equalizer (TEQ) for VDSL System

研 究 生：梁 禮 涵　　Student：Li-han, Liang

指導教授：林 源 倍 教授　　Advisor：Yuan-Pei, Lin

國 立 交 通 大 學

電 機 與 控 制 工 程 學 系

碩 士 論 文

A Thesis

Submitted to Department of Electrical and Control Engineering

College of Electrical Engineering and Information Science

National Chiao Tung University

in partial Fulfillment of the Requirements

For the Degree of

Master

In

Electrical and Control Engineering

June 2004

Hsinchu, Taiwan, Republic of China

中華民國九十三年六月

應用於高速數位用戶迴路之
時域等化器設計

學生：梁禮涵　　　　　　　　　指導教授：林源倍博士

國立交通大學電機與控制工程學系（研究所）碩士班

摘　　要

本論文中，我們提出一個時域等化器之設計應在高速用戶迴路通訊系統。此傳輸系統利用分頻多工傳輸來分開上游與下游之通信訊號。在下游（上游）的傳輸中上游（下游）之頻道不會被利用於傳輸在此稱做空頻道。如通道之次方數大於循環前置保護塊之長度，那麼在空頻道將會收到前後交錯訊號雜訊與通道雜訊。我們提出的時域等化器之設計方法運用在空頻道的能量來使等效通道之能量集中。本設計方法不需要知道原通道之脈衝響應。設計例子將會顯示所提出之方法能夠設計良好能使能通道量集中的時域等化器。

Design of time domain equalizer(TEQ) for VDSL system

Student:Li-Han,Liang Advisors:Dr. Yuan-Pei Lin

Department of Electrical and Control Engineering
National Chiao Tung University

June 23, 2004

Abstract

In this paper, we propose a semi-blind TEQ design method for VDSL system. In the VDSL system FDD(Frequency Division Duplex) is used to separate upstream and downstream signals. In downstream(upstream) transmission the upstream(downstream) tones are not used and are referred as null tones. If the channel order is larger than the length of cyclic prefix, the null tone will contains the noise and ISI. The proposed TEQ design method exploit the null tone energy to shorten the channel. The design does not require the channel impulse response. Examples will be given to show that the method can design TEQ with good shortening effect.

誌　　謝

感謝指導教授 林源倍 老師，讓我求學階段中能夠不斷地成長。在老師的指導之下，除了學科的知識，老師合藹的待人態度與嚴謹的治學方法，更是我最好的身教。謝謝老師諄諄地教誨，並在我怠惰時能及時拉我一把，謝謝老師。

同時，我也要感謝口試委員： 林清安 教授以及 鄭木火 教授，能夠在百忙之中抽空指教並且給我相當寶貴的建議，讓這篇論文更加完整。

感謝陪我走過這段時間的伙伴。與我同屆的 孟良、 諺棋，一同渡過了許多在研究室為了課業及研究而奮鬥的日子；立人 學長總是不斷地給我鼓勵與指點方向； 已畢業的重慶及昶誠學長不厭其煩地回答我的問題；還有下一屆的 沛如、建樟及 裕彬，在做論文為我打氣，以及給我口試準備上的建議。還有口琴社的各位，讓我當學生的日子充滿歡樂。

最後感謝我的家人，父母給予的支持與大哥表現的榜樣都是我前進的動力，是我完成學業最需要重要的支柱，謝謝你們。

Contents

1 Introduction **4**

1.1 Notations 9

2 System model **10**

2.1 Background 10

2.2 Previous TEQ works 14

2.2.1 Optimal shortening, and TEQ optimized for ISI and noise 14

2.2.2 Training TEQ for channel shortening 17

2.2.3 Pole-zero cancellation 20

2.2.4 MERRY algorithm 21

2.2.5 The SAM 23

2.2.6 Blind equalization using TEQ 26

2.2.7 Bit rate maximizing TEQ 28

3 TEQ calculation **32**

3.1 Design Procedure 38

4 Numerical Simulation **41**

4.1 Measure 41

4.2 Environment 42

4.3 The SIR performance 43

4.4 Example 43

5 conclusion **48**

List of Figures

2.1 The DMT system 10
2.2 DMT system represents in matrix form. 11
2.3 DMT system with TEQ added at the receiver 13
2.4 Effective channel representation 13
2.5 illustration of Pole-zero cancellation 20
2.6 illustration of received $r(n)$ sequence 21
2.7 System model proposed by [9]. 26
2.8 System model proposed in [8]. 28

3.1 System model proposed in the thesis. 32
3.2 VDSL band allocation. 33

4.1 Measure of SIR. 42
4.2 (a)Impulse response of original channel.(b)Frequency response of original channel.(c)The DFT transform of averaged received signal $\overline{\mathbf{r}}$. 45
4.3 (a)Designed TEQ coefficients.(b)Frequency response of TEQ.(c)Output signal $\overline{\mathbf{x}}$ of TEQ 46
4.4 (a)Impulse response of original channel and equalized channel.(b)Frequency response of original channel and equalized channel.(c)Compare of $\overline{\mathbf{r}}$ and $\overline{\mathbf{x}}$ after DFT 47

Chapter 1

Introduction

During the last decades, extensive research has been done to provide broadband communication to customer. Today, the discrete multitone modulation(DMT) is used in digital audio/vedeo broadcasting [1], [2], in wireless local area networks such as IEEE 802.11a and HIPERLAN2, and in digital subscriber lines(DSL).

DMT system has an advantage of modulation is easily performed, it could be model as taking a parallel to serial sequence after the FFT process, by appending the CP in front of each block of DMT symbol, repeating the last L sample of DMT symbol, it exploit the circular convolution property and prevents the interference from previous block. This makes the DMT system so popular due to easy equalization using banks of scales. The length of the CP should be determined from the length of the equivalent discrete time channel. The minimal length of CP should equal the order of the equivalent discrete time channel for perfect equalization, but using long CP reduces the data rate of the system, by $\frac{N}{(N+L)}$. At first, the way to avoid using a long CP, is introducing a time domain filter in to the system. If the effective channel(The equivalent discrete time channel cascaded with TEQ) have energy concentrated in an small interval less than CP, it gets better performance due to less interference, we call such time filter a channel shortening time domain equalizer(TEQ). This concept is first proposed on [3], the paper provides an optimal shortening ,and a pole-zero cancellation using least-square(LS) channel modeling approach to design TEQ. For the optimal shortening method, given the original channel impulse, and the CP length, the algorithm generates the coefficients of the optimal shortening TEQ by maximizing the ratio of signal power over interference power. The performance measure SIR is directly related to the eigenvalue corresponds to the eigenvector which generates the coefficients of TEQ. While the algorithm provides the shortest possible effective channel, the computation of eigenvalues and eigenvectors could be difficult to implement due to high complexity, and the accurate time domain impulse response is considered unknown in the transmission system. The concept of pole-zero cancellation is finding a channel model of pole-zero system, assume the system transfer function is $h(z) = \frac{a(z)}{1+b(z)}$ design the TEQ as $1 + b(z)$. Then the effective channel becomes $h_{eff}(z) = a(z)$, if total order of $a(z)$ is smaller than the length of cyclic prefix, we could get perfect reconstruct of original signal, under the assumption of that channel is stationary and causal, and neglect of noise. The Least-square(LS) channel estimation method, finds the channel pole-zero model by criterion of minimizing the output square error of the estimated channel compared to the received channel. The error is defined by the difference of estimated pole-zero system output and the real channel output that we got. It requires auto-correlation matrix of input signal to output signal, which could be interpreted as the average of samples over a period of time. The estimated equivalent discrete channel model could then be solved in a closed form, it is a great success on channel shortening. At the conclusion authors claim that maybe the channel shortening criterion is not the best measure, since noise is no considered in the equalization and the SNR of each tone in the DMT system directly related to the performance of the transmission system. Therefore, TEQ designed for eliminating only ISI does not give best performance of the system for it lacks the optimization for noise. [4] proposed an optimal equalizer based on criterion of channel shortening with color noise. It points out that minimizing ISI doesn't minimizes the overall mean square error of signal, and noise term in the transmission system should take part in optimizing the TEQ, and a cost function of signal power over interference power and noise power is proposed. If the channel and auto-correlation matrix of noise is given, a Rayleigh ratio problem is formed and TEQ could be generated. There is another adaptive training algorithm that converges to Minimizing the Mean Square Error proposed, it is modified from a TEQ training algorithm designed for channel shortening. When noise is introduced in the system, there is only minor part of adaptive algorithm need to be modified. The error in the algorithm adds a noise term estimated from the receiving signal, and the result converges to minimizing the ISI and noise, same as the criterion described above. The overall noise and interference power is minimized in this kind of criterion, and great improvement is done. But in the DMT system, bit rate and probability of error on each tone which is related to tone SNR is not considered in the criterion.

On the other hand, blind equalization algorithm uses the property of transmitted signal to design TEQ. [9] is the first paper to study about the blind equalization of multicarrier system, the main purpose of the paper is proposing a system without cyclic prefix to increase the throughput, the channel effect equalization is done by a time domain filter. At the transmitter side, information symbol and null symbol are encoded in the Frequency domain, after the DFT block, the parallel vector is convert in to serial sequence, thus transmitted to the receiver. The received sample is first equalized by a time domain filter and convert into vector transformed in Frequency domain. With no guard interval, there will be interference in the receiver side, and value at the Null tone will not be zero. The TEQ is designed to minimized the interference shown in the Null tone, since there is no guard interval, adaptive equalization continues until the effective channel is an impulse. The concept of "minimizing ISI using the information in the Frequency domain" is used in this thesis, thought we are not minimizing the total ISI, but only the ISI in the Null tone. We could still get good performance on channel shortening.

Blind Channel shortening method that uses properties of cyclic prefix is proposed in [6] known as "Multicarrier Equalization by Restoration of RedundancY" (MERRY) algorithm, it claims that if the channel order is shorter than the length of cyclic prefix, and due to the data in cyclic prefix is the repeat of last L symbol in DMT symbol. The receiving sample where location corresponds to last sample of cyclic prefix should match the last of receiving sample of the location of DMT symbol. The solution converges to the Optimal shortening method of [3] in an adaptive way, updating of the coefficient of TEQ is done every DMT symbol received. Modification of the cost function could be done to make more update possible in one iteration, the authors proposed a cost function that matches a pair of samples in one iteration, the modification could done by matching "I" pairs of symbols at each iteration, thus update could be done "I" times each iteration. The MERRY algorithm performs blind adaptive channel shortening for DMT system, but the algorithm often fails when noise is significant, the gradient could easily be distored to the opposite way of channel shortening.

The "Sum-square Auto-correlation Minimization"(SAM) algorithm focus on the nature of channel proposed in [7]. First, the authors give a definition of channel auto-correlation function, and claims that if the channel have order L, the value auto-correlation function is zero outside the window of length $(2L + 1)$, which the center of the window is placed at index 0. The cost function defined as the summation of square-auto-correlation value outside the window, optimization could be done by minimizing the value of cost function. We could get the auto-correlation sequence by estimation of auto-correlation of the received signal, and the TEQ coefficient updates until the criterion is fulfilled, thus we could a effective channel with a smaller order. It is considered totally blind since there is no need of knowing the symbol synchronization parameter, the only restriction is the receiving samples should be random, this will easily fulfilled since in most conditions data is considered to be random variable. Both algorithm above give great performance on channel shortening, but poor in noisy condition. To overcome the noise effect, we must have a tradeoff of adaptive method between the robustness of noise. Due to the properties of VDSL training symbol, averaging the received symbol could be done to get rid of most part of noise and having the ISI part left.

By now,various of TEQ design have been proposed and studied, the goal of the design is application dependent, in wireless scenario, there is no bit allocation action taken in the transmission, so bit-error rate minimization and fast adaptation to non-stationary environment are desired; in DSL, bit-rate maximization in a stationary environment is needed. TEQ design of DSL application is a fascinating topic to many researchers, A filter bank representation of generalized TEQ is proposed as a solution of the problem [8]. The proposed design represents the DMT in filter banks, each tone have it's individual TEQ as the basic structure of the receiver, given the channel, and noise statistical parameters, the ISI and noise term at the output of each tone can be obtained, thus the SNR of each tone related to TEQ is revealed. It could be seen as a generalized form of TEQ, if single TEQ design is needed, we just need to let all the TEQ have the same coefficients. Since tone SNR could be found, optimization could be done in all kinds of criterion. The tone SNR is only related to its TEQ design, optimizing SNR of each tone, we could get the best performance in the system, in the sense of bit rate of each tone is directly related to it's tone SNR under a fixed probability of error. An optimal multiple TEQs design is presented, it could be given in a close form and can serve of a bench mark of TEQ design based on bit rate maximizing.

The goal of this thesis, is proposing a method of TEQ, that is robust of noise.In practical use, channel is unknown and channel noise always degrades the performance of the adaptive methods. Thought it is not a adaptive method, we assume the channel for DSL channel varies little by time and could be neglected. The TEQ is design on a VDSL transmission system, and uses the training stage symbol to design the TEQ. In the system at training stage, there will be tones with data symbol and training symbol and zero input in Null tones. By the same concept, if there is ISI in the system, signals in Null tone will not be zero.Part of the training symbol is constant, and we could average the received symbol, the result of it is signal and ISI information with minor noise effect, and if we could minimize the ISI in the null tone or maximize the Signal power in addition we might get a filter suitable for channel shortening. Thus, we give two cost function for experiment one minimizing the signal power in the Null tone since the Null tone only contain ISI, the other is signal tone power over Null tone.

Outline

In chapter 2, we introduce the DMT system with TEQ added in the receiver and the previous works of Multicarrier system is introduced, it could be consider as the main idea of the proposed method. In chapter 3, we will have detail description about the proposed method and analysis of the method will be carried on. Numerical simulation is presented in chapter 4. A conclusion will be given in chapter 5.

1.1 Notations

1. Bold face upper case letters represents matrices.Bold face lower case letters represents matrices. $\mathbf{A}^{\top}$ denotes transpose of $\mathbf{A}$, and $\mathbf{A}^{\dagger}$ denotes conjugate transpose of $\mathbf{A}$, works both for vector and matrix.

2. $\|\mathbf{x}\|$ denotes 2-norm of vector $\mathbf{x}$

3. The function $E\left[y\right]$ denotes the average value or expect value of y.

4. $eig(\mathbf{A})$ is the operation of finding eigenvalues of $\mathbf{A}$.

5. $*$ denotes the linear convolution operator.

Chapter 2

System model

2.1 Background

The DMT transmitter side, information bits are encoded the into a conjugate symmetry vector before the DFT block to form a real signal through the DFT process. The output sequence is added with cyclic prefix(CP) of length L and then transmitted to the other end of the receiver. The CP consist of last L samples of the transmitted DMT symbol.At the receiver, the receive signal is then passed through the TEQ, and CP is removed at the next stage. This signal, is then transformed back to the frequency domain by a DFT block. After the symbol detector, each of the frequency tone could be decoded into bit stream. In the practical system, the DFT is implemented digitally. Therefore, there is a digital to analog(DAC) pulse shaping filter, and a analog to digital(ADC) pulse matching filter connecting the physical channel between the transmitter and receiver. By combining the DAC, real channel, and the ADC we could get a equivalent discrete

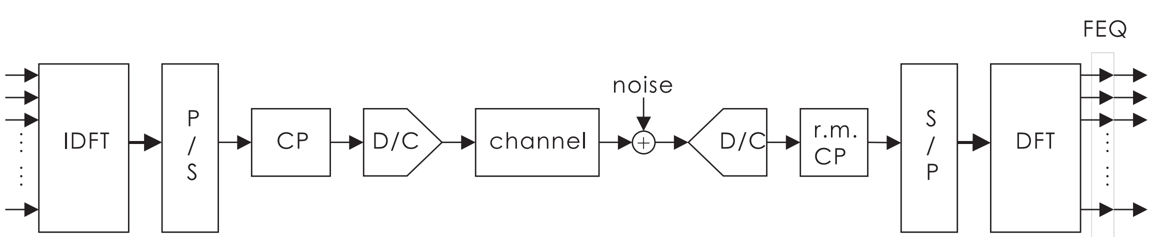

Figure 2.1: The DMT system

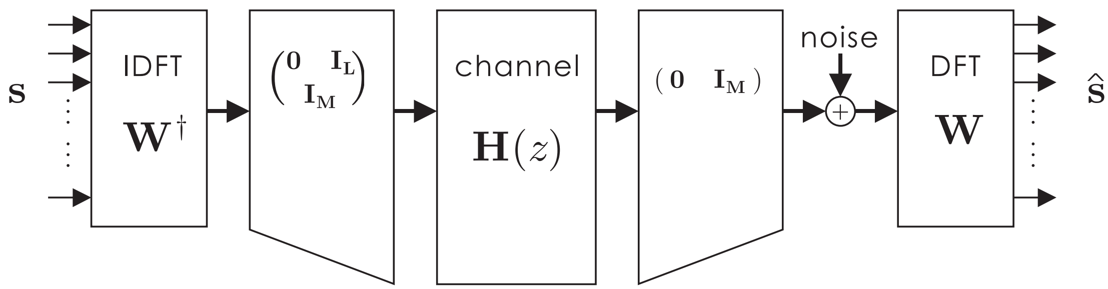

Figure 2.2: DMT system represents in matrix form.

time channel. As shown in Fig. 2.1. The transceiver is a Multiple-input Multiple-output system by considering the inputs at the transmitter and the output at the receiver, we could analyze the signals and the system using matrix and vectors shown in Fig. 2.2. Assuming the order of the channel N_h is smaller than the length of cyclic prefix L and the size of the block length is M. Due to the redundancy in the transmission, interference from the previous symbol will be removed at the cyclic prefix removal block at the receiver, and thanks to the data repeat in the cyclic prefix, the equalization is easily done by a scale on each tone, represent as the FEQ block Λ^{-1} which it's diagonal terms are the M-point DFT of discrete time channel coefficients. There are N samples transmitted to the receiver, where $N = M + L$, and the equivalent discrete channel $h[n]$ could be represent as a $N \times N$ pseudo-circulant matrix $\mathbf{H}(z)$, in the from

$$\mathbf{H}(z) = \begin{pmatrix} h[0] & 0 & \cdots & 0 & z^{-1}h[N_h] & \cdots & z^{-1}h[1] \\ h[1] & h[0] & \ddots & \vdots & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & \vdots & \ddots & z^{-1}h[N_h] \\ h[N_h] & h[N_h-1] & \ddots & h[0] & 0 & \ddots & 0 \\ 0 & h[N_h] & \ddots & \ddots & h[0] & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & h[N_h] & h[N_h-1] & \cdots & h[0] \end{pmatrix}_{N\times N} \tag{2.1}$$

Which the matrix could be represent two sub-matrix,

$$\mathbf{H}(z) = \begin{pmatrix} \mathbf{H_0(z)} \\ \mathbf{H_c} \end{pmatrix}$$

where $\mathbf{H_c}$ is a constant matrix with size $M \times N$ and $\mathbf{H_0}(z)$ is the submatrix that contains z^{-1} which represent that interference that comes from the previous block, size $L \times N$.

$$\mathbf{H_c} = \begin{pmatrix} 0 & \cdots & 0 & h[N_h] & h[N_h-1] & \ddots & h[0] & 0 & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 & h[N_h] & \ddots & \ddots & h[0] & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & h[N_h] & h[N_h-1] & \cdots & h[0] \end{pmatrix}_{M\times N} \tag{2.2}$$

And it is the from of equivalent discrete channel combine with the cyclic prefix removing block, then we could observe that $\mathbf{H_c}\begin{pmatrix} \mathbf{0} & \mathbf{I_L} \\ & \mathbf{I_M} \end{pmatrix}$ is a matrix by adding the first L columns to the last L column of the $M \times M$ right submatrix of $\mathbf{H_c}$, the result is a $M \times M$ circulant matrix given $\mathbf{H_{cyc}}$

$$\mathbf{H_{cyc}} = \begin{pmatrix} h[0] & 0 & \cdots & 0 & h[N_h] & \cdots & h[1] \\ h[1] & h[0] & \ddots & \vdots & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & \vdots & \ddots & h[N_h] \\ h[N_h] & h[N_h-1] & \ddots & h[0] & 0 & \ddots & 0 \\ 0 & h[N_h] & \ddots & \ddots & h[0] & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & h[N_h] & h[N_h-1] & \cdots & h[0] \end{pmatrix}_{M\times M} \tag{2.3}$$

And the circulant of matrix could be diagonalized by $\mathbf{W}$, and the relation could be written as

$$\mathbf{H_{cyc}} = \mathbf{W}^{\dagger}\mathbf{\Lambda}\mathbf{W}$$

, where the Λ is the inverse of Λ^{-1}

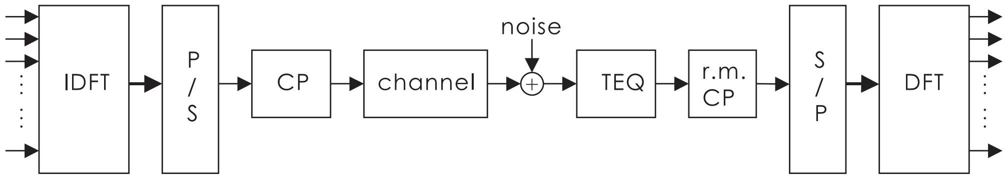

Figure 2.3: DMT system with TEQ added at the receiver

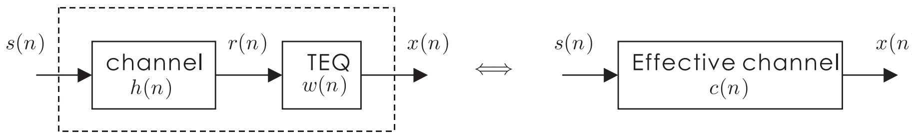

Figure 2.4: Effective channel representation

The relation of signal part could be examined as simplifying the equation

$$\begin{aligned}
\hat{\mathbf{s}} &= \mathbf{\Lambda^{-1}W}\begin{pmatrix} \mathbf{0} & \mathbf{I_M} \end{pmatrix}\mathbf{H}(\mathrm{z})\begin{pmatrix} \mathbf{0} & \mathbf{I_L} \\ \multicolumn{2}{c}{\mathbf{I_M}} \end{pmatrix}\mathbf{W}^{\dagger}\mathbf{s} && (2.4)\\
&= \mathbf{\Lambda^{-1}W}\begin{pmatrix} \mathbf{0} & \mathbf{I_M} \end{pmatrix}\begin{pmatrix} \mathbf{H}_0(z) \\ \mathbf{H}_c \end{pmatrix}\begin{pmatrix} \mathbf{0} & \mathbf{I_L} \\ \multicolumn{2}{c}{\mathbf{I_M}} \end{pmatrix}\mathbf{W}^{\dagger}\mathbf{s} && (2.5)\\
&= \mathbf{\Lambda^{-1}WH_c}\begin{pmatrix} \mathbf{0} & \mathbf{I_L} \\ \multicolumn{2}{c}{\mathbf{I_M}} \end{pmatrix}\mathbf{W}^{\dagger}\mathbf{s} && (2.6)\\
&= \mathbf{\Lambda^{-1}WH_{cyc}W^{\dagger}s} && (2.7)\\
&= \mathbf{\Lambda^{-1}WW^{\dagger}\Lambda WW^{\dagger}s} && (2.8)\\
&= \mathbf{s} && (2.9)
\end{aligned}$$

Therefore, in absence of the channel noise, and the assumption that the length of cyclic prefix is larger than the channel order. We could get a conclusion that it is able to observe the original signal we transferred with perfect reconstruction through the DMT system.

When the channel order is much longer than the length of cyclic prefix, TEQ is added to perform channel shortening to reduce interference that distort the signal. By denoting the equivalent discrete time channel impulse as $h(n)$,and $w(n)$ is the TEQ impulse response. The output of TEQ absence of noise could be expressed as

$$x(n) = (h(n) * w(n)) * s(n) = c(n) * s(n) \tag{2.10}$$

where $c(n)$ is the effective channel. In most TEQ design case, thought we couldn't really shorten the channel into and order of L, but the channel power in time domain is concentrated in a small interval by TEQ designed of channel shortening or minimizing the overall error in the system using other design criterion. By this result, the data rate will increase or the probability of error falls due to less ISI and less noise received, thus give better performance to the transmission system.

2.2 Previous TEQ works

In this chapter, details of previous works will be introduced. They are Optimal shortening, Pole-zero cancellation from [3], and the noise added optimal shortening from [4]. Blind channel shortening of [6] and [7]. the blind equalization of Multicarrier system[9], and the bit rate optimized TEQ [8].

2.2.1 Optimal shortening, and TEQ optimized for ISI and noise

The optimal shortening deals with an imaginary effective channel, choosing a window located at a constant delay on it, and we wish to forces as much power to lie inside the window. Denoting the window length as $N_w + 1$, the sample delay d, the effective channel $\mathbf{c}$ its nth element $c(n)$, and the original channel as $h(n)$ with length N_h, the TEQ coefficient as $\mathbf{t}$, a $(N_t \times 1)$ vector.

we start from representing the effective channel by original equivalent discrete channel and the unknown TEQ value.

$$
\begin{aligned}
\mathbf{c} &= \begin{pmatrix} c(0) \\ c(1) \\ \vdots \\ c(L_h) \\ c(N_h+1) \\ \vdots \\ c(N_h+t-1) \end{pmatrix} \\
&= \begin{pmatrix} h(0) & 0 & \cdots & \cdots & 0 \\ h(1) & h(0) & \ddots & & \vdots \\ \vdots & \ddots & & & \vdots \\ h(N_h-1) & h(N_h-2) & \cdots & h(N_h-t+1) & h(N_h-t) \\ 0 & h(N_h-1) & \cdots & & h(N_h-t+1) \\ \vdots & \ddots & & & \vdots \\ 0 & \cdots & & 0 & h(N_h-1) \end{pmatrix} \begin{pmatrix} t(0) \\ t(1) \\ \vdots \\ t(N_t-1) \end{pmatrix} \\
&= \mathbf{H}\,\mathbf{t} \qquad (2.11)
\end{aligned}
$$

given the delay d and the window size N_w, we define the effective channel inside the window as $\mathbf{c}_{win}$,

$$
\begin{aligned}
\mathbf{c}_{win} &= \begin{pmatrix} c(d) \\ c(d+1) \\ \vdots \\ c(d+N_w) \end{pmatrix} \\
&= \begin{pmatrix} h(d) & h(d-1) & \cdots & h(d-t+1) \\ h(d+1) & h(d) & \cdots & h(d-t+2) \\ \vdots & & \ddots & \vdots \\ h(d+N_w) & h(d+N_w-1) & \cdots & h(d+N_w-t+1) \end{pmatrix} \begin{pmatrix} t(0) \\ t(1) \\ \vdots \\ t(N_t-1) \end{pmatrix} \\
&= \mathbf{H}_{win}\,\mathbf{t} \qquad (2.12)
\end{aligned}
$$

and the effective channel outside the window defined as $\mathbf{c}_{wall}$,

$$\begin{aligned}
\mathbf{c}_{wall} &= \begin{pmatrix} c(0) \\ \vdots \\ c(d-1) \\ c(d+N_w+1) \\ \vdots \\ c(N_h+t-1) \end{pmatrix} \\
&= \begin{pmatrix} h(0) & 0 & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ h(d-1) & h(d-2) & \cdots & h(d-t) \\ h(d+N_w+1) & h(d+N_w) & \cdots & h(d+N_w-t+2) \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & 0 & h(N_h-1) \end{pmatrix} \begin{pmatrix} t(0) \\ t(1) \\ \vdots \\ t(N_t-1) \end{pmatrix} \\
&= \mathbf{H}_{wall}\,\mathbf{t} \qquad (2.13)
\end{aligned}$$

Then the optimal shortening can be expressed as choosing the TEQ $\mathbf{t}$ to maximize the ratio of power inside the window over power outside the window $\frac{\mathbf{c}_{win}^{\dagger}\mathbf{c}_{win}}{\mathbf{c}_{wall}^{\dagger}\mathbf{c}_{wall}}$ subject to, $\mathbf{t}^{\dagger}\mathbf{t} = 1$.To represent the power by TEQ coefficient and original channel.

$$\mathbf{c}_{win}^{\dagger}\mathbf{c}_{win} = \mathbf{t}^{\dagger}\mathbf{H}_{win}^{\dagger}\mathbf{H}_{win}\,\mathbf{t} = \mathbf{t}^{\dagger}\mathbf{A}\,\mathbf{t} \qquad (2.14)$$

$$\mathbf{c}_{wall}^{\dagger}\mathbf{c}_{wall} = \mathbf{t}^{\dagger}\mathbf{H}_{wall}^{\dagger}\mathbf{H}_{\mathbf{wall}}\,\mathbf{t} = \mathbf{t}^{\dagger}\mathbf{B}\,\mathbf{t} \qquad (2.15)$$

where $\mathbf{A}$, $\mathbf{B}$ are symmetric and positive semidefinite matrixes. Optimal shortening could be done by choosing the TEQ $\mathbf{t}$ to maximize the ratio, $\frac{\mathbf{t}^{\dagger}\mathbf{A}\mathbf{t}}{\mathbf{t}^{\dagger}\mathbf{B}\mathbf{t}}$,we could turn the ratio into a Rayleigh ratio problem by Cholosky decomposition, $\mathbf{B} = \mathbf{U}^{\dagger}\mathbf{U}$

$$\frac{\mathbf{t}^{\dagger}\mathbf{A}\mathbf{t}}{\mathbf{t}^{\dagger}\mathbf{B}\mathbf{t}} = \frac{\mathbf{t}^{\dagger}\mathbf{A}\mathbf{t}}{\mathbf{t}^{\dagger}\mathbf{U}^{\dagger}\mathbf{U}\mathbf{t}} \qquad (2.16)$$

and let $\mathbf{l} = \mathbf{U}\mathbf{t}$

$$\frac{\mathbf{t}^{\dagger}\mathbf{A}\mathbf{t}}{\mathbf{t}^{\dagger}\mathbf{U}^{\dagger}\mathbf{U}\mathbf{t}} = \frac{\mathbf{l}^{\dagger}\mathbf{U}^{-\dagger}\mathbf{A}\mathbf{U}^{-\mathbf{1}}\mathbf{l}}{\mathbf{l}^{\dagger}\mathbf{l}} = \frac{\mathbf{l}^{\dagger}\mathbf{Q}\mathbf{l}}{\mathbf{l}^{\dagger}\mathbf{l}} \qquad (2.17)$$

and the maximum value of the ratio is given by the largest eigenvalue of $\mathbf{Q}$ denote as λ_{max} , the corresponding eigenvector $\mathbf{l}_{max}$ could generate the TEQ coefficient by

$$\mathbf{t} = \mathbf{U}^{-\mathbf{1}}\mathbf{l}_{max} \qquad (2.18)$$

also the shortening SIR of the equalized channel could be found as

$$SIR_{opt} = 10\log\left(\frac{\mathbf{t}^{\dagger}\mathbf{A}\mathbf{t}}{\mathbf{t}^{\dagger}\mathbf{B}\mathbf{t}}\right) = 10\log(\lambda_{max}) \tag{2.19}$$

We will follow a similar approach to the previous one, but Total error consist of ISI and noise. We could modify the optimization of ISI into optimization of both noise and ISI by adding the noise factor into the cost function. Let's derive the expression for error due to noise, If $n(k)$ is the noise samples at the input, the noise variance at the equalizer output is

$$\sigma_n^2 = E\left[z^*(m)z(m)\right] = E\left[\left(\sum_k t(m-k)n(k)\right)^* \left(\sum_l t(m-l)n(l)\right)\right] \tag{2.20}$$

where $t(n)$ is the equalizer coefficients, and we need to write this result into a form similar to $\mathbf{t}^{\dagger}\mathbf{A}\,\mathbf{t}$ in order to match the previous result and the same optimization technique.

$$\sigma_n^2 = \mathbf{t}^{\dagger}E\left[\mathbf{n}\mathbf{n}^{\dagger}\right]\mathbf{t} = \mathbf{t}^{\dagger}\mathbf{R}_{\mathbf{x}}\mathbf{t} \tag{2.21}$$

where $\mathbf{Nt}$ corresponds to the linear convolutions process and $\mathbf{n}$ is denoted as,

$$\mathbf{n} = \begin{pmatrix} n(k) & n(k-1) & \cdots & n(k-N_t) \end{pmatrix}^{\top}$$

and $\mathbf{R}_{\mathbf{x}}$ is the noise auto-correlation matrix. This expression is in the same form of the optimization problem in the previous section, and we could easily add the noise term in the ratio.

$$SINR = \frac{\mathbf{t}^{\dagger}\mathbf{A}\mathbf{t}}{\mathbf{t}^{\dagger}\left(\mathbf{B}+\mathbf{R}_{\mathbf{x}}\right)\mathbf{t}} \tag{2.22}$$

and the result is done the same way to solve the optimization problem.

2.2.2 Training TEQ for channel shortening

In absence of noise in the DMT, the transmitted signal is $X(e^{jw})$ and the received signal is $Y(e^{jw})$ the channel frequency response has the relation of

$$H(e^{jw}) = \frac{Y(e^{jw})}{X(e^{jw})} \tag{2.23}$$

We want to model the channel into two FIR filters

$$H(e^{jw}) = \frac{F(e^{jw})}{T(e^{jw})} \tag{2.24}$$

If $T(e^{jw})$ is the TEQ, then $F(e^{jw})$ is the equalized channel. If the order of equalized channel is shorter than the cyclic prefix, we get the relation of

$$F(e^{jw})X(e^{jw}) = T(e^{jw})Y(e^{jw}) \tag{2.25}$$

since the relation is distorted by is the channel have order larger than the length of cyclic prefix, so we wish to design a TEQ to force this relation. By setting $T(e^{jw})$ as the TEQ, and $F(e^{jw})$ as a ideal shortened channel, we evaluate the error by

$$\mathbf{E} = (\mathbf{FX}) - (\mathbf{TY}) \tag{2.26}$$

where E, F, X, T, Y denotes the sample of $E(e^{jw}), F(e^{jw}), X(e^{jw}), T(e^{jw}), Y(e^{jw})$ and the operation $(\cdot)$ denotes a component-wise production, we wish to minimize this error power. Since the input is complex, and it is adaptive calculation cost tone by tone, we could derive the complex gradient function one tone at a time, Let E, F, X, T, Y denote the same tone index entries.

$$J = |E|^2 = E \cdot E^* = E_R^2 + E_I^2 \tag{2.27}$$

where $Y = Y_R + jY_I$, $T = T_R + jT_I$ and

$$\begin{aligned} E &= E_R + jE_I \\ &= (FX) - (TY) \\ &= (FX)_R + j(FX)_I - (T_R + jT_I)(Y_R + jY_I) \\ &= [(FX)_R - T_RY_R + T_IY_I]_{E_R} + j\,[(FX)_I - T_RY_I - T_IY_R]_{E_I} \end{aligned} \tag{2.28}$$

and the real part and imaginary part of the gradient could be expressed as,

$$\frac{\partial J}{\partial T_R} = -2E_RY_R - 2E_IY_I \tag{2.29}$$

$$\frac{\partial J}{\partial T_I} = 2E_RY_I - 2E_IY_R \tag{2.30}$$

and the gradient of T becomes

$$
\begin{aligned}
\frac{\partial J}{\partial T} &= \frac{\partial J}{\partial T_R} + j\frac{\partial J}{\partial T_I} \\
&= -2(E_R Y_R + E_I Y_I) + 2j(E_R Y_I - E_I Y_R) \\
&= -2(E_R + jE_I)(Y_R - jY_I) \\
&= -2E \cdot Y^* \qquad (2.31)
\end{aligned}
$$

and last we summarize the adaptive algorithm.

1. For a given T, compute $\mathbf{F} = H\mathbf{T}$

2. $\mathbf{f}$=IFFT($\mathbf{F}$) has to be windowed. This process searches the location where the window captures the largest energy and zero out the component outside the window.

3. After the windowing, Let

$$\mathbf{E} = (\mathbf{FX}) - (\mathbf{TY})$$

   to be the error and complete the update by

$$\mathbf{T}' = \mathbf{T} + \mu(\mathbf{EY}^*)$$

4. $\mathbf{t}$= IFFT($\mathbf{T}$) is longer than desire had has to be window, keeping the first L_w entries and zero out the rest.

Next we will modify the training algorithm to include the effects of color noise, the adaptive algorithm could converge to a Minimum Mean Square Error(MMSE) solution by adding a term that corresponds to noise into the training algorithm. Note that the noise vector at the equalizer's output could is ($\mathbf{NT}$) and now the error in the algorithm becomes

$$\mathbf{E} = (\mathbf{FX}) - (\mathbf{TY}) + (\mathbf{TN}) \qquad (2.32)$$

the receiving symbol are $\mathbf{P} = \mathbf{Y} + \mathbf{N}$. After averaging is done, $\mathbf{Y}$ is known and we could obtain the noise part by subtracting $\mathbf{Y}$ from the received symbol $\mathbf{P}$

$$\mathbf{E} = (\mathbf{FX}) - 2(\mathbf{TY}) + (\mathbf{TP}) \qquad (2.33)$$

and rest of the algorithm are the same.

2.2.3 Pole-zero cancellation

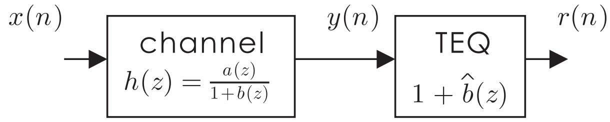

Figure 2.5: illustration of Pole-zero cancellation

Assume that the original channel impulse response of the channel is represented as a transfer function $h(z) = \frac{a(z)}{1+b(z)}$, shown in Fig. 2.5. The poles of the transfer function represent as tail after the peak, so cancellation of pole leaves the zero of the transfer function, and hopefully we will get shortened channel. This approach estimates a model of the equivalent channel, in this section, we will give a Least-Square approach for example of finding the pole-zero model of equivalent channel. We set the estimation transfer function as $\hat{h}(z) = \frac{\hat{a}(z)}{1+\hat{b}(z)}$ and we define the parameter as a vector,

$$\theta = \begin{pmatrix} \hat{a}_0 & \hat{a}_1 & \cdots & \hat{a}_\nu & -\hat{b}_0 & -\hat{b}_1 & \cdots & -\hat{b}_\eta \end{pmatrix}^\top \tag{2.34}$$

and a regress vector of input and output sample of the system.

$$\Phi(n) = \begin{pmatrix} x(n) & x(n-1) & \cdots & x(n-\nu) & y(n-1) & y(n-2) & \cdots & y(n-\eta) \end{pmatrix}^\top \tag{2.35}$$

By using the two equation above, we could get an expression of the estimated $y(n)$

$$\hat{y}(n) = \theta^\top \Phi(n) \tag{2.36}$$

and the error $c(n)$ as

$$c(n) = y(n) - \hat{y}(n) \tag{2.37}$$

and the name least-square (LS) comes from that the desire of minimize the square-error. The result of Least-Square is then

$$\theta_{LS} = \mathbf{R^{-1}r} \tag{2.38}$$

where

$$\begin{aligned}
\mathbf{R} &= E\left[\Phi(n)\Phi^{\top}(n)\right] && (2.39)\\
\mathbf{r} &= E\left[y(n)\Phi(n)\right] && (2.40)
\end{aligned}$$

The coefficient is then the least-square model parameters, and take the poles of the estimate model $1+\hat{b}(z)$ to cascade with the original channel, and if the estimation is good enough we could get

$$\left(\frac{a(z)}{1+b(z)}\right)(1+\hat{b}(z)) \approx a(z) \qquad (2.41)$$

The LS approach developed above calculation and inversion of he autocorrelation matrix $\mathbf{R}$. The inversion makes the algorithm so complicated for practical use. Other system modeling such as Autoregressive modeling, could be taken in practice when using the concept of pole-zero cancellation TEQ.

2.2.4 MERRY algorithm

The MERRY algorithm is given high respect to its low complexity, it works on DMT system, data repeat of cyclic prefix play an important role in the equalization. Thought there is only one update every DMT samples, it could be modified and give increase on convergence rate.

We give a illustration of the basic concept of the MERRY algorithm, shown in Fig. 2.6 , consider there is a DMT symbol of 8 sample, and cyclic prefix of length 2 samples, if the channel impulse response have 5 samples, the last sample of cyclic prefix will be $r(2)$ and the last sample of cyclic prefix will be $r(10)$

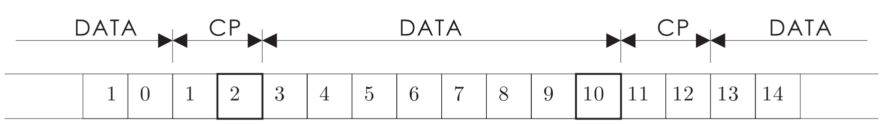

Figure 2.6: illustration of received $r(n)$ sequence

$$\begin{aligned}
r(2) &= x(2)h(0)+x(1)h(1)+[x(0)h(2)+x(-1)h(3)+x(-2)h(4)]\\
r(10) &= x(10)h(0)+x(9)h(1)+x(8)h(2)+x(7)h(3)+x(6)h(4)\\
&= x(2)h(0)+x(1)h(1)+[x(8)h(2)+x(7)h(3)+x(6)h(4)]
\end{aligned}$$

If we add a TEQ $w(n)$ to force the last three sample of channel $h(2) = h(3) = h(4) = 0$, then $r(2) = r(10)$ due to the date repeat in the cyclic prefix $x(2) = x(10), x(1) = x(9)$, and the channel is then shortened to order 1 in the case. The author propose an cost function to express this concept

$$J_{ori,\triangle} = E\left[|y(L+\triangle) - y(L+N+\triangle)|^2\right] \tag{2.42}$$

where $\triangle$ is the date synchronization parameter. and $y(n)$ denotes the nth sample of receiving DMT symbol at TEQ output. We could make the MERRY update more times per iteration by modifying the cost function into

$$J_{mod,\triangle} = E\left[\sum_i |y(L+\triangle-i) - y(L+N+\triangle-i)|^2\right] \tag{2.43}$$

since we could get the expected value, we introduce another parameter to estimate the instant cost instead:

$$J_{inst,\triangle} = \sum_i |y(L+\triangle-i) - y(L+N+\triangle-i)|^2 \tag{2.44}$$

and the gradient could be calculated as,

$$\frac{\partial J_{inst,\triangle}}{\partial t(l)} = \sum_i \left[y(L+\triangle-i) - y(L+N+\triangle-i)\right]\left[r(L+\triangle-i-l) - r(L+N+\triangle-i-l)\right] \tag{2.45}$$

where $r(n)$ denotes the nth sample of receiving DMT symbol with cyclic prefix at TEQ output. And we define $\tilde{e}_{i,\triangle} = [y(L+\triangle-i) - y(L+N+\triangle-i)]$ as the instant error caused by ISI. and $\tilde{r}_{i,\triangle} = [r(L+\triangle-i-l) - r(L+N+\triangle-i-l)]$ for convenience. Thus, update of single TEQ parameter could be done by

$$t'(l) = t(l) - \mu\frac{\partial J_{inst,\triangle}}{\partial t(l)} \tag{2.46}$$

and a normalization is done at the last step of iteration.

$$\mathbf{t}' = \frac{\mathbf{t}}{\parallel \mathbf{t} \parallel} \tag{2.47}$$

forcing the $\mathbf{t}^\dagger\mathbf{t} = 1$ and prevents the trivial solution of the TEQ. And Let's summarize these steps to give a clear view of how simple is the MERRY algorithm at each iteration.

1. Calculate $\tilde{e}_{i,\triangle}$, $\tilde{r}_{i,\triangle}$ the gradient is then to be calculated as,

$$\frac{\partial J_{inst,\triangle}}{\partial t(l)} = \sum_i \tilde{e}_{i,\triangle}\tilde{r}_{i,\triangle}$$

2. update the TEQ coefficients

$$t'(l) = t(l) - \mu \sum_i \tilde{e}_{i,\triangle}\tilde{r}_{i,\triangle}$$

3. normalize the coefficients

$$\mathbf{t}' = \frac{\mathbf{t}}{\parallel \mathbf{t} \parallel}$$

2.2.5 The SAM

We begin with the analysis of SAM cost function, and show how to measure the channel auto-correlation from the received data. We define the auto-correlation sequence of the combined channel-equalizer impulse response,

$$R_{cc}(l) = \sum_{k=0}^{N_c} c(k)c(k-l) \tag{2.48}$$

the parameter N_c denotes the order of combined channel impulse response, if the combined channel have a length of $L+1$, it is necessary for the auto-correlation value $R_{cc}(l)$ to be zero outside the window of length $(2L+1)$

$$R_{cc}(l) = 0, \forall |l| > L. \tag{2.49}$$

and come up of a possible way of channel shortening , forcing the power outside the window to be zero. Hence, we define a cost function

$$J = \sum_{L+1}^{N_c} |R_{cc}(l)|^2 \tag{2.50}$$

and the optimization problem could be stated as,

$$\mathbf{t}^{opt} = arg_{\mathbf{t}} \min_{\mathbf{t}^{\dagger}\mathbf{t}=1} J \tag{2.51}$$

By consider the auto-correlation function of sequence $y(n)$, with the assumptions of

1. source sequence $x(n)$ is a white, zero-mean, and wide sense stationary.

2. The combined channel order N_c holds the relation $2N_c < M$for the DMT system, where N is the DFT size.

3. The noise sequence $\nu(n)$ is zero-mean i.i.d., uncompleted to the source sequence and has a variance σ_ν^2, and the source sequence $x(n)$ is real and has a unit variance.

$$\begin{aligned} R_{yy}(l) &= E\left[y(n)y(n-l)\right] \\ &= E\left[(\mathbf{c}^\top \mathbf{x}_n + \mathbf{t}^\top \mathbf{v}_n)(\mathbf{c}^\top \mathbf{x}_{n-l} + \mathbf{t}^\top \mathbf{v}_{n-l})\right] \\ &= \sum_{k=0}^{N_c} c(k)c(k-l) + \sigma_\nu^2 \sum_{k=0}^{N_t} t(k)t(k-l) \\ &= R_{cc}(l) + \sigma_v^2 R_{tt}(l) \end{aligned} \tag{2.52}$$

There is something to argue about the input source $x(n)$, the authors claims SAM could work on a DMT system, and is simulated in a ADSL environment, how come the source $x(n)$ is considered W.S.S ? since the cyclic prefix is added and there should be correlation between the $x(n)$. It could be explained with the assumption of the channel order relation $2N_c < M$, , since $x_n = \begin{bmatrix} x(n) & x(n-1) & \cdots & x(n-N_c) \end{bmatrix}^\top$ and the biggest $R_{yy}(l)$ that we are dealing with is $R_{yy}(N_c)$ there will be no elements correlated with the samples in cyclic prefix since the auto-correlation function of $x(n)$ could be represent as $R_{xx}(n) = \sigma_x^2\left(\delta(n) + \delta(n-M) + \delta(n+M)\right)$, and $R_{yy}(n) = \sigma_x^2\left(\delta(n) + \delta(n-M) + \delta(n+M)\right) * R_{cc}(l)$, and since $2N_c < M$ there will be no aliasing between the three copies of channel effect and what we are interested is just the interval of $L+1 < |l| < N_c$, and then the input source of the DMT system could be seen as no correlation. Thus, under noiseless scenario, $\sigma_\nu^2 = 0$

$$J = \sum_{L+1}^{N_c} |R_{yy}(l)|^2 = \sum_{L+1}^{N_c} |R_{cc}(l)|^2 \tag{2.53}$$

and in the presence of noise,

$$
\begin{aligned}
\hat{J} &= \sum_{L+1}^{N_c} |R_{yy}(l)|^2 \\
&= \sum_{L+1}^{N_c} |R_{cc}(l)|^2 + 2\sigma_\nu^2 \sum_{L+1}^{N_c} R_{cc}(l)R_{ww}(l) + \sigma_\nu^4 \sum_{L+1}^{N_c} |R_{ww}(l)|^2 \\
&\approx \sum_{L+1}^{N_c} |R_{cc}(l)|^2
\end{aligned}
\tag{2.54}
$$

the approximation could seen as that most TEQ $\mathbf{w}$ have smaller order compare to the length of cyclic prefix and the last term is zero, and the noise term σ_ν^4 is small and the middle term is considered to be neglect.

Now, we begin the part of adaptive algorithm derivation, we could get the true value of $E[\cdot]$, and we use the instant value of it instead, so the cost function turns into

$$
J^{inst}(k) = \sum_{L+1}^{N_c} \left(\sum_{n=kN}^{(k+1)N-1} \frac{y(n)y(n-l)}{N} \right)^2 \tag{2.55}
$$

and the gradient of the algorithm could be expressed as,

$$
\frac{\partial J^{inst}(k)}{\partial w(j)} = 2 \sum_{L+1}^{L_c} \left(\sum_{n=kN}^{(k+1)N-1} \frac{y(n)y(n-l)}{N} \right) \left(\sum_{n=kN}^{(k+1)N-1} \frac{y(n)r(n-l-j) + y(n-l-j)r(n)}{N} \right) \tag{2.56}
$$

k denotes the kth iteration of the algorithm and the gradients could be represent from input and output samples we could summarize the SAM adaptive algorithm as follow,

1. calculate the gradients of $\frac{\partial J^{inst}(k)}{\partial t(j)}$

2. update the TEQ coefficients

$$
t'(j) = t(j) - \mu \frac{\partial J^{inst}(k)}{\partial t(j)}
$$

3. normalize the TEQ coefficients by

$$
\mathbf{t}' = \frac{\mathbf{t}}{\| \mathbf{t} \|}
$$

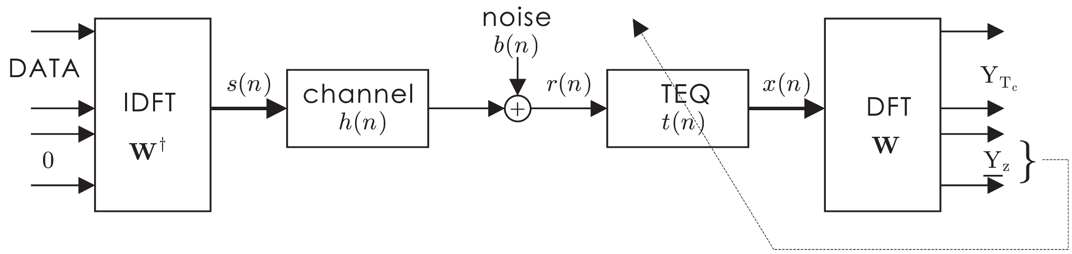

Figure 2.7: System model proposed by [9].

2.2.6 Blind equalization using TEQ

The system showed in Fig. 2.7 is proposed in [9]. Its main concept is designing a adaptive method to equalize the channel, by transmitting some zeros in certain carrier, The update continues whenever the value of $\mathbf{Y}_z$,which corresponds to the null which encode at the transmitter, is not zero. The system transmits random input from tone 1 to N_{u-1} represented in Fig.2.7 as $\mathbf{D}$ and rest of the tones from N_u to N is the null tone inputs are zero. After modulated by FFT and turned into serial sequence, the samples are transmitted into the channel, lacking of guard interval in this system, the zero value input at null tones of the transmitter could not be recovered after the DFT process and ISI occurs and zero in Null tones will not recover. We want to represent the system into matrix form in order to proceed optimization. By defining

$$
\begin{aligned}
\mathbf{s}_n &= (s(n), s(n-1), \cdots, s(n-M+1))^\top \\
\mathbf{S}_n &= (\mathbf{s}_n, \mathbf{s}_{n-1}, \cdots, \mathbf{s}_{n-M+1})^\top \\
\mathbf{S}^2 &= (\mathbf{S}_n, \mathbf{S}_{n-M})
\end{aligned}
$$

to represent the convolution process of the signal, and the channel $\mathcal{H}$ is size $2M \times M$ lower triangular matrix Toeplitz matrices whose first column is given by

$$\mathbf{H} = (h(0), h(1), \cdots, h(N_c), 0, \cdots, 0)^\top$$

Defining

$$\begin{aligned}\mathbf{b}_n &= (b(n), b(n-1), \cdots, b(n-M+1))^\top \\ \mathcal{B}_n &= (\mathbf{b}_n, \mathbf{b}_{n-1}, \cdots, \mathbf{b}_{n-M+1})^\top\end{aligned}$$

to represent the convolution process of noise, the received signal at the TEQ output could be expressed as

$$\mathbf{x} = \underline{\mathbf{S}}^2 \mathcal{H}\mathbf{t} + \mathcal{B}_n \mathbf{t} \tag{2.57}$$

$$\mathbf{Y} = \mathbf{F}\mathbf{y} = \begin{pmatrix} \mathbf{F}_{T_c} \\ \mathbf{F}_z \end{pmatrix} \mathbf{y} = \begin{pmatrix} \mathbf{Y}_{T_c} \\ \mathbf{Y}_z \end{pmatrix} \tag{2.58}$$

where $\mathbf{F}_{T_c}$ denotes the first N_u rows of the DFT matrix corresponds to the signal part and the rest of the rows constructs $\mathbf{F}_z$, due to the data inputs are random we could express the power at null tones using the variance of samples. By minimizing the total power of $\mathbf{Y}_z$ setting the sum of total variance of $\mathbf{Y}_z$ as cost function

$$J_z = E\left[\|\mathbf{Y}_z\|^2\right] = E\left[\mathbf{Y}_z^\dagger \mathbf{Y}_z\right] \tag{2.59}$$

Such criterion could be rewritten as,

$$J_z = \mathbf{t}^\dagger (\mathbf{P} + \mathbf{Q})\mathbf{t} \tag{2.60}$$

where,

$$\mathbf{P} = \mathcal{H}^\dagger E\left[\underline{\mathbf{S}}^{2\dagger} \mathbf{T} \underline{\mathbf{S}}^2\right] \mathcal{H} \tag{2.61}$$

$$\mathbf{Q} = E\left[\mathcal{B}^\dagger \mathbf{T} \mathcal{B}\right] \tag{2.62}$$

we simplify $\mathbf{F}_z^\dagger \mathbf{F}_z$ as $\mathbf{T}$. To optimize the cost function subject to $\mathbf{t}$ is constrained to unit norm ($\mathbf{t}^\dagger \mathbf{t} = 1$). The optimization could be viewed as a Rayleigh ratio problem , thus solve the minimum eigenvalue of $(\mathbf{P} + \mathbf{Q})$ give the optimal solution of the cost function J_z.

The cost function could be modified into an adaptive implementation of minimizing J_z, by letting $J_z^{inst} = \mathbf{Y}_z^\dagger \mathbf{Y}_z$ for instantaneous estimation of the cost function, we could get the gradient of the cost function respect to the TEQ coefficients by,

$$\frac{\partial J_z^{inst}}{\partial \mathbf{t}^*} = \mathbf{R}^\dagger \mathbf{F}_\mathbf{z}^\dagger \mathbf{Y}_z \tag{2.63}$$

Then, we could summarize the adaptive equalization algorithm by:

1. get the error of output Y_T and the gradient could be calculated as $\frac{\partial J_z^{inst}}{\partial \mathbf{t}^*} = \mathbf{R}^\dagger \mathbf{F}_\mathbf{z}^\dagger \mathbf{Y}_z$

2. update the equalizer
$$\mathbf{t}' = \mathbf{t} - \mu \frac{\partial J_z^{inst}}{\partial \mathbf{t}^*}$$

3. modify the equalizer into unit norm
$$\mathbf{t}' = \frac{\mathbf{t}}{\| \mathbf{t} \|_2}$$

With no guard interval there must be interference from the previous symbol. By neglecting noise in transmission, the ideal equalization of this method, values of null tone at the receiver is recovered as zero, it means the TEQ is an inverse of the channel thus the effective channel is an impulse. Without the CP the solution leads effective channel into a spike.

2.2.7 Bit rate maximizing TEQ

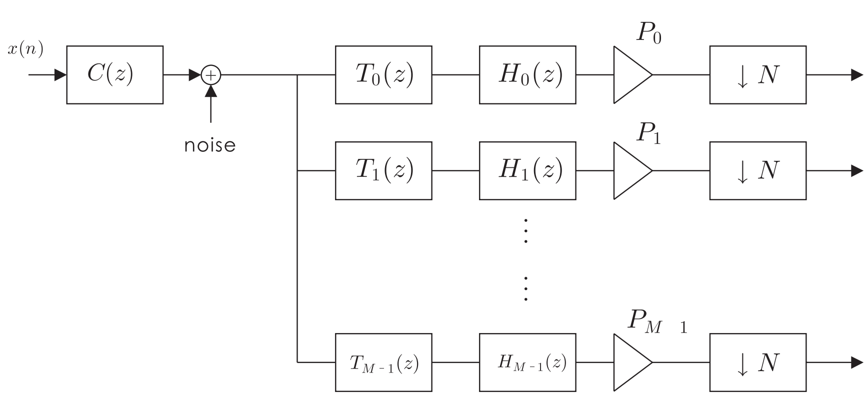

Figure 2.8: System model proposed in [8].

The bit rate maximizing TEQ here, uses a filter bank representation. Using this approach, the ISI and noise error at the FEQ output can be obtained. By the filter bank representation, minimizing the mean square error and the geometric error using one TEQ could be designed. In addition to the criterion propose, the filter bank representation could be done to have a generalized TEQ design, multiple TEQ design is proposed in the paper[8], and given in a close form. We start to introduce the formulation of the ISI error and noise error to obtain the tone SNR at output of FEQ. By choosing a synchronization delay n_s, we denote the cyclic prefix length is L in the system, consecutive L samples after delay n_w forms a window of signal part of the received signal, and the equivalent channel outside is the formation of the ISI part. Define the sequence

$$d(n) = \begin{cases} 0 & \text{for } n_w < n \leq n_w + L \\ 1 & \text{for } 0 < n \leq n_w \text{ or } n_w + L < n \leq N_c + N_t \end{cases}$$

where N_c denotes the order of the original channel and N_t denotes the order of the equalized channel, and we could denote the ISI term of kth tone as

$$p_{isi,k}(n) = d(n)\,(c(n) * t_k(n)) \tag{2.64}$$

thus, the output error at the kth tone is given by $e_k(n) = [e_{isi,k}(n) + e_{\nu,k}(n)]_{\downarrow N}$

$$\begin{aligned} e_{isi,k}(n) &= h_k(n) * p_{isi,k}(n) * x(n)/P_k & (2.65) \\ e_{\nu,k}(n) &= h_k(n) * t_k(n) * \nu(n)/P_k & (2.66) \end{aligned}$$

where P_k denotes the FEQ scale on the kth tone. To obtain the SNR of each tone, we must have the noise power, which is the variance of $e_{isi,k}$, $e_{\nu,k}$ and the total noise power is denote as

$$\sigma^2_{e,k} = \sigma^2_{isi,k} + \sigma^2_{\nu,k}$$

where we have the assumption of signal and noise are uncorrelated. We can express the error variance in matrix form. define the kth TEQ $t(n)$ coefficients and filter bank $w(n)$ in to vector.

$$\begin{aligned} \mathbf{t}_k &= (t_k(0)\ t_k(1)\ \cdots\ t_k(N_t))^\top \\ \mathbf{w}_k &= \left(1\ e^{j2\pi k/M}\ \cdots\ e^{j2\pi k/M}\right)^\top \end{aligned}$$

Let $\mathbf{C}$ and $\mathbf{H}_k$ be respectively $(N_c + N_t + 1) \times (N_t + 1)$ and $(M + N_c + N_t + 1) \times (N_c + N_t + 1)$ lower triangular Toeplitz matrices, which represents the linear convolution process, their first column is given by

$$(c(0)\ c(1)\ \cdots\ c(N_c)\ 0\ \cdots\ 0)^\top$$
$$\left(e^{j2\pi k(M-1)/M}\ \cdots\ e^{j2\pi k/M}\ 1\ 0\ \cdots\ 0\right)^\top$$

and $\mathbf{D}$ is the window matrix size $(N_c + N_t + 1) \times (N_t + N_c + 1)$ with entries $\mathbf{D}_{ii} = d(i)$ using the definition above the error variance could be represent as,

$$\sigma^2_{isi,k} = \frac{\sigma_x^2 \mathbf{t}_k^\dagger \mathbf{C}^\dagger \mathbf{D}^\dagger \mathbf{H}_\mathbf{k}^\dagger \mathbf{H}_\mathbf{k} \mathbf{D} \mathbf{C} \mathbf{t}_k^\dagger}{|C(e^{j2\pi k/M})|^2 \mathbf{t}^\dagger \mathbf{w}_k^\dagger \mathbf{w}_k \mathbf{t}} \tag{2.67}$$

$$\sigma^2_{\nu,k} = \frac{\mathbf{t}_k^\dagger \tilde{\mathbf{H}}_\mathbf{k}^\dagger \mathbf{R}_\nu \tilde{\mathbf{H}}_\mathbf{k} \mathbf{t}_k^\dagger}{|C(e^{j2\pi k/M})|^2 \mathbf{t}^\dagger \mathbf{w}_k^\dagger \mathbf{w}_k \mathbf{t}} \tag{2.68}$$

where $\tilde{\mathbf{H}}_\mathbf{k}$ is the smaller version of $\mathbf{H}_\mathbf{k}$ of size $(M + N_t) \times (N_t + 1)$. Defining two matrix for convenient

$$\mathbf{Q}_{isi,k} = \frac{\sigma_x^2 \mathbf{C}^\dagger \mathbf{D}^\dagger \mathbf{H}_\mathbf{k}^\dagger \mathbf{H}_\mathbf{k} \mathbf{D} \mathbf{C} \mathbf{t}_k^\dagger}{|C(e^{j2\pi k/M})|^2}, \qquad \mathbf{Q}_{\nu,k} = \frac{\tilde{\mathbf{H}}_\mathbf{k}^\dagger \mathbf{R}_\nu \tilde{\mathbf{H}}_\mathbf{k}}{|C(e^{j2\pi k/M})|^2}$$

the optimization of Minimization of Mean Square Error at FEQ output using single TEQ could be expressed as

$$arg \min_t \sum_{k=0}^{M-1} \frac{\mathbf{t}^\dagger (\mathbf{Q}_{isi,k} + \mathbf{Q}_{\nu,k}) \mathbf{t}}{\mathbf{t}^\dagger \mathbf{w}_k^\dagger \mathbf{w}_k \mathbf{t}} \tag{2.69}$$

the average is done by the summation of all error variance of each tone, and the scaler is neglected. but the optimization could be highly non-linear. The bit rate maximizing problem using one TEQ could be interpreted as minimizing the geometric mean of the variance. That the maximum achievable per tone bit rate under $P_e = 10^{-7}$ is given by

$$\begin{aligned} b &= \sum_k \log 2 \left(1 + \frac{\sigma_x^2/\sigma_{e_k}^2}{10}\right) \\ &\approx \sum_k \log 2 \left(\frac{\sigma_x^2/\sigma_{e_k}^2}{10}\right) \\ &= \prod_k \left(\frac{\sigma_x^2/\sigma_{e_k}^2}{10}\right) \end{aligned}$$

we could see from above, maximizing the bit rate using one TEQ is the same as minimizing the geometric mean of one TEQ, and the optimization problem could be expressed as.

$$arg\min_{t} \prod_{k=0}^{M-1} \frac{\mathbf{t}^{\dagger}(\mathbf{Q}_{isi,k} + \mathbf{Q}_{\nu,k})\mathbf{t}}{\mathbf{t}^{\dagger}\mathbf{w}_k^{\dagger}\mathbf{w}_k\mathbf{t}} \tag{2.70}$$

at last, note that the e_k^2 is only affected by the kth TEQ, so the overall optimization could be done by optimizing each TEQ. The optimization could be expressed as

$$arg\min_{t} \frac{\mathbf{t}^{\dagger}(\mathbf{Q}_{isi,k} + \mathbf{Q}_{\nu,k})\mathbf{t}}{\mathbf{t}^{\dagger}\mathbf{w}_k^{\dagger}\mathbf{w}_k\mathbf{t}} \tag{2.71}$$

we could decomposite the matrix inside $(\mathbf{Q}_{isi,k} + \mathbf{Q}_{\nu,k})$ as $Q_h^{\dagger}Q_h$ since $\mathbf{Q}_{isi,k}, \mathbf{Q}_{\nu,k}$ is semi-positive, and letting $\mathbf{u}_k = \mathbf{Q_h}\mathbf{t}_k$ thus the optimization could be done by solving

$$arg\max_{u_k} \frac{u_k^{\dagger}\mathbf{Q}_h^{-\dagger}\mathbf{w}\mathbf{w}^{\dagger}\mathbf{Q}^{-\dagger}u_k}{u_k^{\dagger}u_k} \tag{2.72}$$

since the matrix $\mathbf{Q}_h^{-\dagger}\mathbf{w}\mathbf{w}^{\dagger}\mathbf{Q}^{-\dagger}$is rank one, the eigenvalue that maximized the above function is $u_{k,opt} = \mathbf{Q}_h^{-\dagger}\mathbf{w}$ and the TEQ become the close form solution of

$$t_{k,opt} = \mathbf{Q}_h^{-1}\mathbf{Q}_h^{-\dagger}\mathbf{w} \tag{2.73}$$

though could not be implemented, the optimization result could serve as a bench mark of TEQ design respect to the system performance.

Chapter 3

TEQ calculation

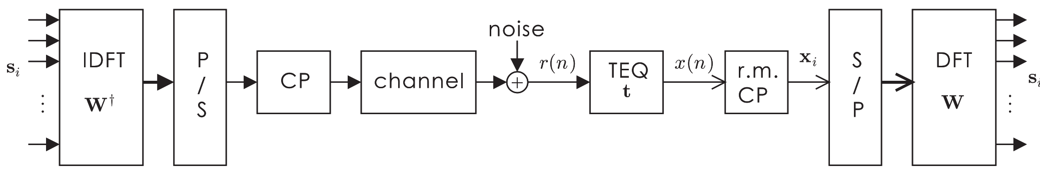

Figure 3.1: System model proposed in the thesis.

Let $\mathbf{x}_n$ be the training symbol of transmission system and $\mathbf{r}_n$ be the received signal at the transmitter. the dash line block represents TEQ will be added into the system after calculation of the method. Since the work is done on the VDSL system at training stage, we will have a introduction about the signal. VDSL transceiver uses a Frequency Division Duplex (FDD) to separate upstream and downstream transmission. In the trial standard [10], the frequency plan consists of two upstream bands are denoted by 1U, 2U and two downstream bands are denoted by 1D, 2D. Two upstream bands and two downstream bands as shown in Fig. 3.2. The values of the splitting frequencies f_i are given in Table. 3.1.

Table 3.1: VDSL band separating frequency

Separating Frequencies	f_0	f_1	f_2	f_3	f_4	f_5
(MHz)	0.25	0.138	3.75	5.2	8.5	12

The optional band between 25 kHz and 138 kHz is to be negotiated during initialization for use as upstream band or downstream band or not used. In the

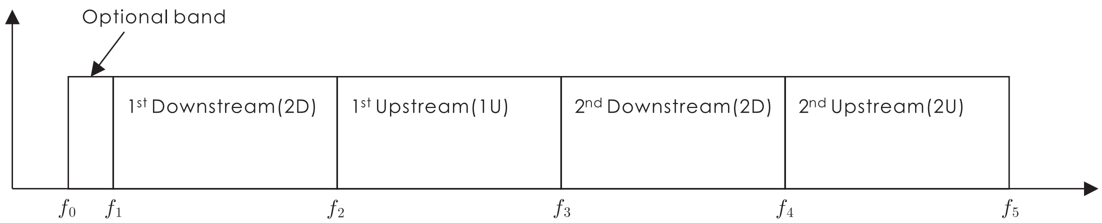

Figure 3.2: VDSL band allocation.

FDD schemes, each tone is used for either downstream or upstream, but not simultaneously. For example, in upstream application, zero are padded in the downstream tones. In this case, the downstream tones are referred to as the null tones. Similarly, in downstream application, the upstream tones are referred to as the null tones. In the training VDSL symbol, some of the data tones are reserved for pilots and the others used for transmitting Special Operation Channel (SOC) message. In the training stage, even tones are reserved and constellation point of 00 is transmitted on even tones. A Special Operation Channel (SOC) message which carries one byte of information is transmitted in every DMT training symbol. The bit mapping of training symbol is given in the Tab. 3.2 .

Table 3.2: Training symbol bit mapping

Tone index	Constellation point
Even	00
1, 11, 21, $\cdots$,10n+1,$\cdots$	SOC message bits 0,1
3, 13, 23, $\cdots$,10n+3,$\cdots$	SOC message bits 2,3
5, 15, 25, $\cdots$,10n+5,$\cdots$	SOC message bits 4,5
7, 17, 27, $\cdots$,10n+7,$\cdots$	SOC message bits 4,5
9, 19, 29, $\cdots$,10n+9,$\cdots$	00

The selected constellation points shall be pseudo-random rotated by $0, \pi/2, \pi$ or $3\pi/2$, and the sequence is reset at every DMT symbol. Once the message bit is generated, it is mapped into 4QAM constellation and the encode the constellation point in to the complex vector $\mathbf{z}$ size $(M/2)$, half of the DFT size$(M/2)$ is considered to be integer since we use FFT process and the input is the power of 2). Before the Modulation by the IDFT, we must map $\mathbf{z}$ into a double sized $\mathbf{s}$ in order to generate a complex-to-real IDFT, thus the vector $\mathbf{s}$ is in the form,

$$s_i = z_i, i = 0, \cdots, (M/2 - 1) \tag{3.1}$$

$$s_i = conj(z_{M-i}), i = M/2, \cdots, (M - 1) \tag{3.2}$$

where $conj(\cdot)$ denotes the conjugate of the complex value, and s_i, z_i denotes the ith element of $\mathbf{s}, \mathbf{z}$. Thus we could have a real vector after the IDFT block. We summarize the properties of VDSL training symbol, there are Null tones in frequency domain, and there are constant part of signal, this enable us to get rid of the noise and knowing where part of the ISI are in frequency domain.

When the channel order is smaller than the length of the cyclic prefix, we know there is no IBI (inter-block interference) after removing guard samples (cyclic removal). In the absence of channel noise, the outputs of the DFT matrix at the receiving end are the scaled versions of the transmitter inputs. The scalars are the M-point DFT of the channel impulse response. In this case, the null tones will be nothing but channel noise. However, if the channel order is larger than the length of cyclic prefix, there will be IBI even after removing guard samples. The output of the null tones now has not only channel noise but also interference from the data tones of the previous block due to IBI (assuming channel order is smaller than N, the length of one block). We observe that the outputs of the null tones will be small if TEQ has effectively shortened the channel.

In this thesis, we propose to a semi-blind TEQ design method for VDSL systems by minimizing the ISI present in the null tones. The design does not require the channel impulse response. To be more specific, suppose the number of data tones is M_d and the number of null tones is M_n, where $M = M_d + M_n$. The numbers M_d, M_n are determined by the spectral plan. Considering the i-th output block, we collect the outputs of the data tones and the outputs of the null tones respectively in vectors $\mathbf{d}_i$ and $\mathbf{n}_i$, for $i = 1, 2, \cdots, B$, where B denotes the number of received output block available for equalizer design. The dimensions of $\mathbf{d}_i$ and $\mathbf{n}_i$ are respectively M_d and M_n. We compute the averaged vectors,

$$\overline{\mathbf{d}} = \frac{1}{B}\sum_{i=1}^{B}\mathbf{d}_i,$$

$$\overline{\mathbf{n}} = \frac{1}{B}\sum_{i=1}^{B}\mathbf{n}_i.$$

We note that the averaged null tone vector $\overline{\mathbf{n}}$ is mostly interference from data tones as averaging remove most noise. Moreover the interference comes mainly from pilot tones because the symbols in message-bearing tones are different from block to block. Two objective functions will used here.

$$\phi_1 = \overline{\mathbf{n}}^{\dagger}\overline{\mathbf{n}}. \tag{3.3}$$

$$\phi_2 = \frac{\overline{\mathbf{d}}^{\dagger}\overline{\mathbf{d}}}{\overline{\mathbf{n}}^{\dagger}\overline{\mathbf{n}}}. \tag{3.4}$$

In the first case, we will optimize the TEQ to minimize interference in null tones, characterized by ϕ_1. In the second case, we will find TEQ to maximize the ratio of data tone energy over the null tone energy. In both cases, the TEQ is constrained to have unit energy, i.e., $\sum_{i=0} |t(i)|^2 = 1$.

Remarks. Notice that our method is semi-blind. Namely, the receiver knows the training symbol contains pilots tones but it knows neither the channel impulse response nor the spectrum of channel noise. We can compare our method to the blind channel equalization method in [9]. The close form equalizer solution requires the channel impulse response and the second order statistics of channel noise in [9].

In what follows, we will see that the objective functions in (3.3) and (3.4) can be formulated as quadratic terms of the TEQ coefficients and the problem can be solved elegantly by computing eigen vectors of appropriately defined positive definite matrices. Suppose the TEQ has order T,

$$T(z) = \sum_{i=0}^{T} t(i)z^{-i}.$$

The output of the TEQ can be written as

$$x(n) = \sum_{\ell=0}^{T} t(\ell) r(n-\ell).$$

Let the i-th intput vector of the DFT matrix be

$$\mathbf{x}_i = \begin{pmatrix} x_{iN+\Delta} \\ x_{iN+\Delta+1} \\ \vdots \\ x_{iN+\Delta+M_1} \end{pmatrix}.$$

Then $\mathbf{x}_k$ can be written in terms of TEQ coefficients as

$$\mathbf{x}_i = \underbrace{\begin{pmatrix} r(iN+\Delta) & r(iN+\Delta-1) & \cdots & r_(iN+\Delta-T) \\ r(iN+\Delta+1) & r(iN+\Delta) & \cdots & r_(iN+\Delta+1-T) \\ \vdots & \vdots & \ddots & \vdots \\ r(iN+\Delta+M-1) & r(iN+\Delta+M-2) & \cdots & r(iN+\Delta+M-1-T) \end{pmatrix}}_{\mathbf{R}_i} \underbrace{\begin{pmatrix} t_0 \\ t_1 \\ \vdots \\ t_T \end{pmatrix}}_{\mathbf{t}}, \tag{3.5}$$

where $\mathbf{R}_i$ is an $M \times (T+1)$ matrix and $\mathbf{t}$ is a column vector of size $(T+1)$. The i-th data tone vector $\mathbf{d}_i$ can be expressed as

$$\mathbf{d}_i = \mathbf{W}_1 \mathbf{x}_i,$$

where $\mathbf{W}_1$ is an $M_d \times M$ submatrix of the $M \times M$ DFT matrix $\mathbf{W}$, obtained by removing the rows that correspond to the null tones. Similarly, we can express the null tone vector $\mathbf{n}_i$ as

$$\mathbf{n}_i = \mathbf{W}_2 \mathbf{x}_i,$$

where where $\mathbf{W}_2$ is an $M_n \times M$ submatrix of the $M \times M$ DFT matrix $\mathbf{W}$, obtained by removing the rows that correspond to the data tones. Using (3.5), we can write $\mathbf{d}_i$ and $\mathbf{n}_i$ respectively as

$$\mathbf{d}_i = \mathbf{W}_1 \mathbf{R}_i \mathbf{t}, \quad \mathbf{n}_i = \mathbf{W}_2 \mathbf{R}_i \mathbf{t}.$$

Using these expressions, we have

$$\overline{\mathbf{d}} = \mathbf{W}_1 \overline{\mathbf{R}} \mathbf{t}, \quad \mathbf{n}_i = \mathbf{W}_2 \overline{\mathbf{R}} \mathbf{t},$$

where

$$\overline{\mathbf{R}} = \frac{1}{B}\sum_{i=1}^{B}\mathbf{R}_i.$$

Therefore, we have

$$\overline{\mathbf{d}}^{\dagger}\overline{\mathbf{d}} = \mathbf{t}^{\dagger}\underbrace{\overline{\mathbf{R}}\mathbf{W}_1^{\dagger}\mathbf{W}_1\overline{\mathbf{R}}}_{\mathbf{A}}\mathbf{t} = \mathbf{t}^{\dagger}\mathbf{A}\mathbf{t},$$

$$\overline{\mathbf{n}}^{\dagger}\overline{\mathbf{n}} = \mathbf{t}^{\dagger}\underbrace{\overline{\mathbf{R}}\mathbf{W}_2^{\dagger}\mathbf{W}_2\overline{\mathbf{R}}}_{\mathbf{B}}\mathbf{t} = \mathbf{t}^{\dagger}\mathbf{B}\mathbf{t},$$

where $\mathbf{A}$ and $\mathbf{B}$ are square matrices of size $(T+1)$. Also, both matrices are positive definite. The function functions given in (3.3) and (3.4) become

$$\phi_1 = \mathbf{t}^{\dagger}\mathbf{B}\mathbf{t},$$

$$\phi_2 = \frac{\mathbf{t}^{\dagger}\mathbf{A}\mathbf{t}}{\mathbf{t}^{\dagger}\mathbf{B}\mathbf{t}}.$$

Now both objective functions are written as quadratic forms of the TEQ coefficients. The energy constraint $\sum_{i=0}|t(i)|^2 = 1$ becomes $\mathbf{t}^{\dagger}\mathbf{t} = 1$.

Optimal solutions

- *Objection function ϕ_1.* The problem of We can use Rayleigh's principle to minimize ϕ_1 subject to the constraint $\mathbf{t}^{\dagger}\mathbf{t} = 1$. The optimal $\mathbf{t}$ is the eigen vector corresponding to the smallest eigen value of $\mathbf{B}$.

- *Objection function ϕ_2.* We can use two methods to find the optimal $\mathbf{t}$ that maximize ϕ_2 subject to the constraint $\mathbf{t}^{\dagger}\mathbf{t} = 1$.

  - Method 1: As $\mathbf{B}$ is positive definite, we can write decompose $\mathbf{B}$ as $\mathbf{B} = \mathbf{C}^{-\dagger}\mathbf{C}^{-1}$. Then ϕ_2 can written as the ratio $\phi_2 = \frac{\mathbf{t}^{\dagger}\mathbf{A}\mathbf{t}}{\mathbf{t}^{\dagger}\mathbf{C}^{\dagger}\mathbf{C}\mathbf{t}}$. Let $\mathbf{u} = \mathbf{C}^{-1}\mathbf{t}$, then $\mathbf{t} = \mathbf{C}\mathbf{u}$ and

    $$\phi_2 = \frac{\mathbf{u}^{\dagger}\mathbf{C}^{\dagger}\mathbf{A}\mathbf{C}\mathbf{u}}{\mathbf{u}^{\dagger}\mathbf{u}}.$$

    Using Rayleigh's principle, ϕ_2 can be maximized by choosing $\mathbf{u}$ to be the eigen vector corresponding to the largest eigen value of $\mathbf{C}^{\dagger}\mathbf{A}\mathbf{C}$.

– Method 2: The design problem can be stated as "maximize $\overline{\mathbf{d}}^{\dagger}\overline{\mathbf{d}}$. subject to the constraint $\overline{\mathbf{n}}^{\dagger}\overline{\mathbf{n}} = 1$.", which reduce to

$$\max_{t} \left(\mathbf{t}^{\dagger}\mathbf{A}\mathbf{t}\right) \text{ subject to } \left(\mathbf{t}^{\dagger}\mathbf{B}\mathbf{t}\right) = 1 \tag{3.6}$$

Solving 3.6 leads to a TEQ that satisfies the generalized eigenvector problem

$$\mathbf{A}\mathbf{t} = \lambda\mathbf{B}\mathbf{t} \tag{3.7}$$

The solution for $\mathbf{t}$ is the eigenvector corresponding to the largest generalized eigenvalue of $(\mathbf{B}^{-1}\mathbf{A})$.

Remarks on method 1 and method 2. The same TEQ solution could be found by both methods, however method 2 has a computation advantage to method 1, since method 2 only needs a matrix inversion to form the appropriate defined positive definite matrix instead of one cholosky decomposition, matrix inverse, and two matrix multiplication of method 1. The computation effort is reduced.

3.1 Design Procedure

In our proposed TEQ design, we need to compute $M \times (T+1)$ matrix $\overline{\mathbf{R}}$ and $(T+1) \times (T+1)$ matrix $\mathbf{A} = \overline{\mathbf{R}}^{\dagger}\mathbf{W}_1^{\dagger}\mathbf{W}_1\overline{\mathbf{R}}$, $\mathbf{B} = \overline{\mathbf{R}}^{\dagger}\mathbf{W}_2^{\dagger}\mathbf{W}_1\overline{\mathbf{R}}$. These matrices can be computed efficiently as detailed below.

Efficient computation of $\overline{\mathbf{R}}$. Observe that the entries of $\overline{\mathbf{R}}$ are drawn from the $(M+T) \times 1$ vector $\overline{\mathbf{r}}$ given by

$$\overline{\mathbf{r}} = \begin{pmatrix} \frac{1}{B}\sum_{i=1}^{B} r(iN + \Delta - T) \\ \frac{1}{B}\sum_{i=1}^{B} r(iN + \Delta + 1 - T) \\ \vdots \\ \frac{1}{B}\sum_{i=1}^{B} r(iN + \Delta + M - 1) \end{pmatrix}. \tag{3.8}$$

Therefore the matrix $\overline{\mathbf{R}}$ can be obtained by simply computing the average received vector $\overline{\mathbf{r}}$.

Efficient computation of A and B. Notice that

$$\mathbf{A} = \mathbf{E}_1^\dagger \mathbf{E}_1, \quad \mathbf{B} = \mathbf{E}_2^\dagger \mathbf{E}_2,$$

where

$$\mathbf{E}_1 = \mathbf{W}_1 \overline{\mathbf{R}}, \quad \mathbf{E}_2 = \mathbf{W}_2 \overline{\mathbf{R}}.$$

and $\mathbf{E}_1, \mathbf{E}_2$ can be formed from collecting the rows of $\mathbf{W}\overline{\mathbf{R}}$, where $\mathbf{E}_1$ corresponds to data tones and $\mathbf{E}_2$ corresponds to the null tones. We could also observe the column of $\overline{\mathbf{R}}$ is just a data shifting with two different samples from the previous column, we gave a example of fast calculating $\mathbf{E}_1$.

$$\mathbf{E}_1 = (\mathbf{e}_0 \, \mathbf{e}_1 \, \cdots \, \mathbf{e}_T)$$

$$\overline{\mathbf{R}} = (\overline{\mathbf{r}}_0 \, \overline{\mathbf{r}}_1 \, \cdots \, \overline{\mathbf{r}}_T)$$

where $\mathbf{e}_i$,$\overline{\mathbf{r}}_i$ represents the ith column of $\mathbf{E}_1$,$\overline{\mathbf{R}}$. the first column $\mathbf{e}_0$ could be calculated as

$$\mathbf{e}_0 = \mathbf{W}_1 \mathbf{r}_0$$

which is the DFT of $\mathbf{r}_0$ and collects the elements corresponds to the data tones. The next column is a data shift and two different samples from the previous one which we could observe from

$$\begin{pmatrix} \mathbf{e}_{0,0} \\ \mathbf{e}_{0,1} \\ \vdots \\ \mathbf{e}_{0,M_d-1} \end{pmatrix} = \mathbf{W}_1 \begin{pmatrix} \overline{r}(\triangle) \\ \overline{r}(\triangle + 1) \\ \vdots \\ \overline{r}(\triangle + M - 1) \end{pmatrix}$$

$$\begin{pmatrix} \mathbf{e}_{1,0} \\ \mathbf{e}_{1,1} \\ \vdots \\ \mathbf{e}_{1,M_d-1} \end{pmatrix} = \mathbf{W}_1 \begin{pmatrix} \overline{r}(\triangle - 1) \\ \overline{r}(\triangle) \\ \vdots \\ \overline{r}(\triangle + M - 2) \end{pmatrix}$$

and

$$e_{1,0} = r(\triangle - 1) + w^{-k_0}(e_{0,0} - r(\triangle + M - 1)w^{-k(M-1)})$$

thus we could have the relation of

$$\mathbf{e}_{l+1} = \begin{pmatrix} r(\triangle - l + 1) \\ r(\triangle - l + 1) \\ \vdots \\ r(\triangle - l + 1) \end{pmatrix} + \begin{pmatrix} w^{-k_0} \\ w^{-k_1} \\ \vdots \\ w^{-k_{M-1}} \end{pmatrix} \otimes (\mathbf{e}_l + r(\triangle + M - l)) \begin{pmatrix} w^{-k_0(M-1)} \\ w^{-k_1(M-1)} \\ \vdots \\ w^{-k_{M-1}(M-1)} \end{pmatrix} \tag{3.9}$$

where $\otimes$ denotes component wise production. and w^{-kl} denotes the (k, l) component of DFT matrix. We can summarize the design procedure for TEQ optimization using ϕ_2 as follows. The design using ϕ_1 is similar.

1. Collect received signal and compute the average received vector $\bar{\mathbf{r}}$ given in (3.8).
2. Obtain $\mathbf{A}$ and $\mathbf{B}$ by first computing $\mathbf{E}_1 = \mathbf{W}_1\overline{\mathbf{R}}$ and $\mathbf{E}_2 = \mathbf{W}_2\overline{\mathbf{R}}$ using DFT, and then computing $\mathbf{A} = \mathbf{E}_1^\dagger\mathbf{E}_1$ and $\mathbf{B} = \mathbf{E}_2^\dagger\mathbf{E}_2$.
3. Compute $\mathbf{B}^{-1}\mathbf{A}$.
4. Obtain the optimal $\mathbf{t}$ by computing the eigen vector corresponding to the largest eigen value of $\mathbf{B}^{-1}\mathbf{A}$.

The choice of $\triangle$. It is the effective channel that we are dealing with, thus we are choosing the best SIR window delay of a equalized channel that is not revealed. However, from the nature of the effective channel it's main impulse is still near the peak of the original channel, and we place the center of SIR window where the effective channel main impulse is and most of the power of effective channel is inside the SIR window. The delay can be choose as

$$\triangle = \triangle_{gp} - \frac{L}{2} \tag{3.10}$$

where $\triangle_{gp}$ is the group delay of original channel and denoted as the estimation of main peak of original channel.

Chapter 4

Numerical Simulation

In this chapter, we will first introduce the measure of the performance, the simulation environment, and the SIR performance of two cost function verses symbol synchronization delay and designed TEQ tap length. The channel is static and it is the model for VDSL loop7.

4.1 Measure

The performance adopt in the thesis is the Signal to interference ratio(SIR), the SIR depends on the length of the cyclic prefix and the symbol synchronization delay of the DMT system. Finding the SIR respect to a DMT with TEQ system is step as follows:

1. construct the effective channel $c(n)$ by cascading the original impulse response $h(n)$ and the TEQ coefficients $w(n)$ using linear convolution.

$$c(n) = h(n) * w(n)$$

   And $c(n)$ have value form 0 to E_d.

2. the effective channel $c(n)$ is shown in 4.1, the SIR at synchronization delay $\triangle$ with window size $\triangle + L$ is defined as:

$$\mathrm{SIR}_{\triangle,L} = \frac{\sum_{n=\triangle+1}^{\triangle+L} c^2(n)}{\sum_{n=0}^{\triangle} c^2(n) + \sum_{n=\triangle+L+1}^{E_d} c^2(n)}$$

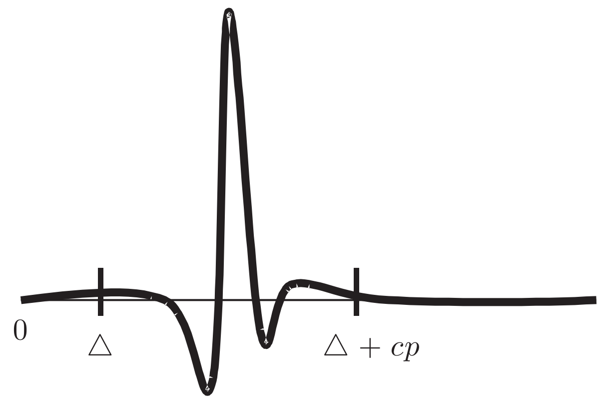

Figure 4.1: Measure of SIR.

which is consider to be the channel power in the window as signal power over channel power out side the window as interference window.

3. proceeds exhaustive search for largest $\mathrm{SIR}_{\triangle,L}$ by changing $\triangle$ one index at a time.

The synchronizing delay $\triangle$ corresponds to the largest $\mathrm{SIR}_{\triangle,L}$ will usually show near the peak of effective channel.

4.2 Environment

The DMT symbol size in the simulation is $N = 8192$ contains 4096 tones with cyclic prefix $L = 640$. A 4-QAM modulation is used for the DMT symbol, and uses downstream bands. Far-end crosstalk(FEXT), near-end crosstalk(NEXT) noise and additive white Gaussian Noise(AWGN) with -140dBm/Hz are considered in the transceiver channel model. The channel is considered to be static and there is no change throughout the simulation the VDSL loop7 is used as channel model.

4.3 The SIR performance

This chapter, we will compare the performance of TEQ, using the measure described in the first section. There will be plots of SIR verses synchronization delay and TEQ tap length. TEQ designed in two objective function will be compared to show the behavior of the TEQ designed by the different objective function.

4.4 Example

We given an example of the design, using a 20 taps filter as TEQ and we average 1000 blocks of received signal $\mathbf{r}_i$. In Fig. 4.2 the channel used in the system is VDSL loop7 and the impulse response and frequency response is shown in (a) and (b). Fig. 4.2(c) is the DFT transform of averaged signal $\bar{\mathbf{r}}$ we could see that the null tone part of the signal is corrupt by noise. The designed TEQ is shown in Fig. 4.3(a) and (b) is the frequency response of the TEQ. Fig. 4.3(c) is the DFT of TEQ output $\bar{\mathbf{x}}$, and we could see that the signal power in the null tone degrades since the cost function of the TEQ is to minimize the energy in null tone. At last, the equalized channel is shown in Fig. 4.4. The comparison of impulse response of equalized and original channel is shown in (a), we could see in the figure the equalized channel have shorter tail than the original channel, and the SIR of the equalized channel is better than the original with a gain of 25.4dB and the frequency response is shown in (b). The table. 4.1 show the semi-blind TEQ performance on different channel, and compared with MERRY and MSSNR method. MERRY is the blind TEQ discussed in previous chapter. MSSNR represents the optimal TEQ designed by optimal shortening, thought it is not a blind or semi-blind method, both method is based on channel shortening and finds the maximum SIR. Therefore, both method have the same criterion, and MSSNR could serve as a benchmark.

Table 4.1: SIR performance on different VDSL loop in dB

VDSL loop	original	semi-blind		MERRY		MSSNR
		AWGN	ALL	AWGN	ALL	
VDSL-1L	47.6	69.8	69.5	69.5	60.9	96.5
VDSL-2L	49.7	59.4	59.4	57.9	59.2	84.8
VDSL-3L	43.3	69.6	69.1	71.5	62.7	86.8
VDSL-4L	30.8	63.1	61.7	55.9	54.9	73.6
VDSL-5	72.7	90.9	90.8	72.0	90.2	98.8
VDSL-6	52.9	85.1	84.2	52.0	70.8	92.1
VDSL-7	35.7	67.7	67.1	55.1	40.5	79.8
VDSL-1_2km	19.0	57.9	57.3	51.9	45.3	75.0
VDSL-2_2km	28.4	63.4	63.0	42.1	51.8	78.4
VDSL-3_2km	27.1	67.7	67.1	51.6	41.7	79.8
VDSL-4_2km	17.2	43.8	43.8	37.3	28.8	44.1

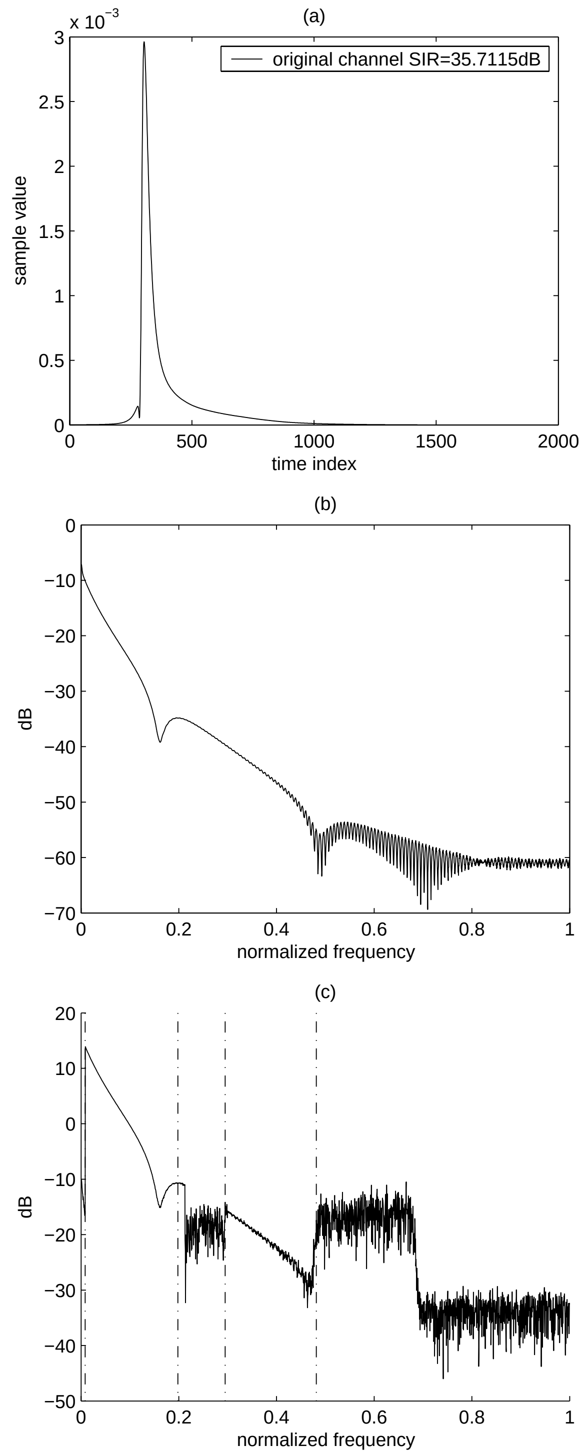

Figure 4.2: (a)Impulse response of original channel.(b)Frequency response of original channel.(c)The DFT transform of averaged received signal $\bar{\mathbf{r}}$.

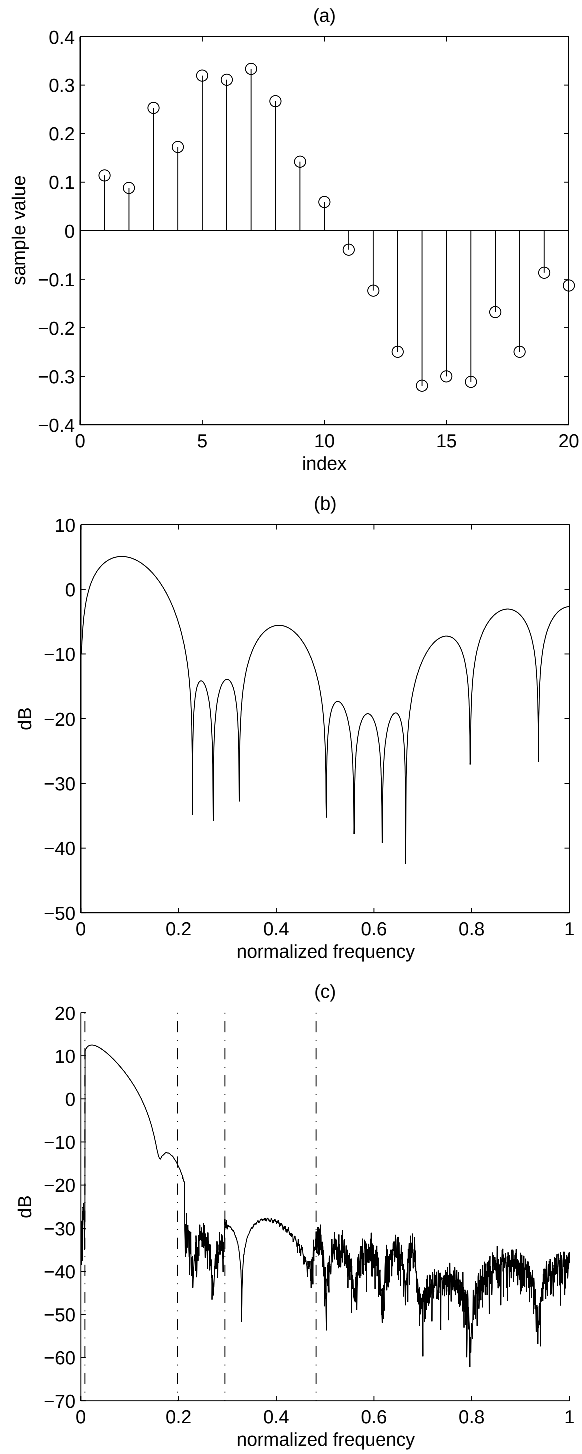

Figure 4.3: (a)Designed TEQ coefficients.(b)Frequency response of TEQ.(c)Output signal $\overline{\mathbf{x}}$ of TEQ

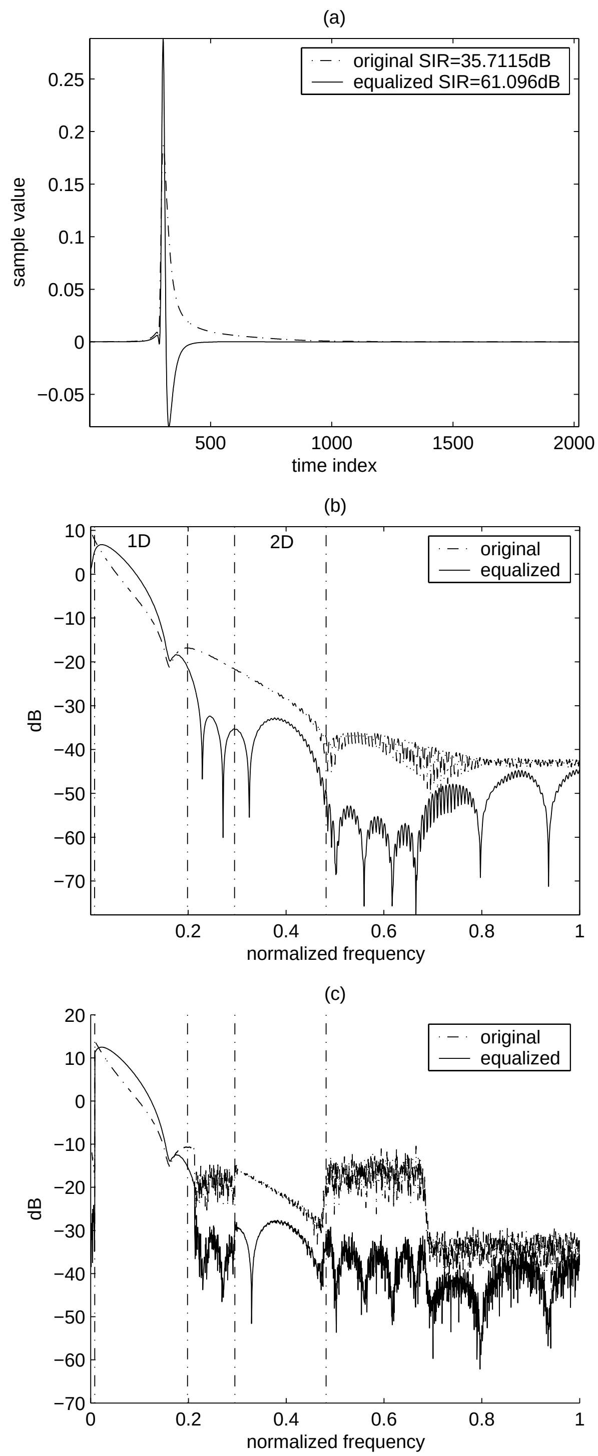

Figure 4.4: (a)Impulse response of original channel and equalized channel.(b)Frequency response of original channel and equalized channel.(c)Compare of $\overline{\mathbf{r}}$ and $\overline{\mathbf{x}}$ after DFT

Chapter 5

conclusion

TEQ design method is studied in the thesis, estimation of peak location of original channel is aided to the design of TEQ. In our simulation experiments thought the cost function is not in sense that it is not minimizing all the ISI, but it would be the trade off between the robust to noise effect and adaptive method implementation. Thought the cost function and the data sample we gather by averaging the incoming DMT symbol has flaws in basic concept, but we could still managed to have a working solution of TEQ. Therefore, the goal of the designing TEQ under noisy environment is accomplished. Further studies of TEQ are the properties of the cost function and if there is any rules of TEQ design that could fit the poles of the channel. Thought noise is not a problem in this method, maximizing bit rate is still a important topic to give raise better performance of the overall DMT transceiver system.

Bibliography

[1] The European Telecomm. Standards Inst, "*Radio Broadcasting System, Digital Audio Broadcasting(DAV) to Mobile, Portible, and Fixed Receivers,*" *ESTI EN 300 401* 1995-1997.

[2] The European Telecomm. Standards Inst, "*Digital Video Broadcasting(DVB); Framing Structure, Channel Coding and Modulation for Digital Terrestrial Television*" *ESTI EN 300 744 V1.4.1* 2001 Edition.

[3] Melsa P.J.W., Younce R.C., Rohrs C.E., "*Impulse response shortening for discrete multitone transceivers* " *Communications, IEEE Transactions* on ,Dec. 1996 pp.1662-1672.

[4] Igor Djokovic, "*MMSE Equalizers for DMT Systems with and without Crosstalk,*" Signals, Systems & Computers, 1997. Conference Record of the Thirty-First Asilomar Conference on Vol.1, Page(s): 545 -549, 1998.

[5] Milos Milosevic, Lucio F. C. Pessoa, Brian L. Evans, and Ross Baldick. "*DMT Bit Rate Maximization WIth Optimal Time Domain Equalizer Filter Bank Architecture*" in *Proc. IEEE Asilomar Conf.,* Pacific Grove, CA, Nov. 2002, vol.1 pp.377-382.

[6] Martin R.K., Balakrishnan J., Sethares W.A., Johnson C.R. Jr. , "*Blind, adaptive channel shortening for multicarrier systems* " *Signals, Systems and Computers, 2002. Conference Record of the Thirty-Sixth Asilomar Conference* on Nov. 2002 pp.372 - 376 vol.1

[7] Balakrishnan J., Martin R.K., Johnson C.R. Jr., "*Blind, adaptive channel shortening by sum-squared auto-correlation minimization (SAM)* " *Signal Processing, IEEE Tran.* on Dec. 2003 pp.3086 - 3093

[8] Chun-Yang Chen, See-May Phoong. "*Bit rate optimized time-domain equalizers for DMT systems* " *Proc. of the 2003 International Symposium* May 2003 pp:IV-37 - IV-40 vol.4

[9] M. de Courville, P. Duhamel, P. Madec, and Palicot, "*Blind Equalization of OFDM System Based on the Minimization of a Quadraic Criterion,*" in *Proceedings of the Int. Conf. on Communications,* Dallas, TX, June, 1996.

[10] T1E1/2002-xxx, Trial-Use standard. Draft Trial-Use Standard For Telecommunications Interface Between Networks and Customer Installation Very-high-bit-rate Digital Subscriber Lines(VDSL) interface, "*Alliance for Telecommunications Industry Solutions*"