Modified Conjugate Gradient Method for ADSL Echo Cancellation

by

Takao Inoue

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Completed June 22, 1998
Commencement June 1999

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF TABLES

MODIFIED CONJUGATE GRADIENT METHOD FOR ADSL ECHO CANCELLATION

1. INTRODUCTION

In the past decade, the rapid growth in use of multimedia capable personal computers to access Internet and the World Wide Web for transfer of image, audio, and video information has resulted in a clamor of high data rate transmission to support the more data rate intensive applications [1]. This has stimulated the market for cheap access techniques with increased data rate. Asymmetric Digital Subscriber Line (ADSL) emerged as an successor to High data rate Digital Subscriber Line (HDSL) access technique to support 6Mbps from central office to the customer premise and 640Kbps from customer premise to the central office. One of the mainstream application is an interactive Video-On-Demand (VOD) where the customer requests particular video (small amount of data) while the central office providing the video stream to the customer premise (large amount of data).

The popularity and the technical difficulty of ADSL system comes from the unique operating environment : twisted pair cable. Twisted pair cables are laid out to almost every home in the U.S. from local telephone companies. Therefore there is no need to lay extra lines for new services. Moreover, ADSL is designed to interoperate with the telephone. Unlike today's computer modems, the customer may use the telephone while the computer is connected to the Internet Service Provider. The achievable data rate is significantly higher in ADSL by almost 100 times compared to todays fastest 56K modems. However, all the above benefits do not come free. Twisted pair cable attenuates signal voltage exponentially with respect to the cable length. A component called hybrid is needed to interface the twisted pair cable which introduces unwanted reflection signal called echo due to impedance mismatch. The presence of the telephone causes rapid changes in the line impedance when the telephone becomes on/off hook. As a result, echo

signal characteristics may also change significantly.

These problems are solved by use of Multicarrier Modulation scheme called Discrete Multitone Modulation, sophisticated channel equalization to combat the ill effects of the twisted pair cables, and echo canceller to reduce the effects of unwanted signals caused by the hybrid. Echo cancellation has been of interest in many areas such as acoustics [17] [18], telephone [19], and especially in the last few decades for data transmission applications [20]. In this thesis, improvement to previously proposed ADSL echo canceller is proposed using Modified Conjugate Gradient for adaptation scheme. It is shown to give better performance in recovering from changes in the twisted pair cable with reasonable computational complexity trade-off.

This thesis is organized as follows. Chapter 2 gives the system overview starting from single carrier systems to multicarrier systems in order to cover the bases needed to understand the echo cancellation problem in ADSL. Chapter 3 gives the outline of the echo canceller architecture that is optimized for use in ADSL. Chapter 4 gives the overview of Conjugate Gradient method and the proposed modified Conjugate Gradient method. Chapter 5 gives the experimental results using the proposed method in a realistic settings. Finally, concluding remarks and possible future research is addressed in Chapter 6.

2. DISCRETE MULTITONE SYSTEM OVERVIEW

To facilitate the understanding of multicarrier systems, passband single carrier communication system is first presented. Single carrier analysis is then extended to multicarrier analysis. Finally, details pertinent to Asymmetric Digital Subscriber Line (ADSL) are presented to support the understanding of ADSL echo canceller.

2.1. Single Carrier Analysis

A general block diagram of single carrier QAM system is shown in Fig.(2.1), where

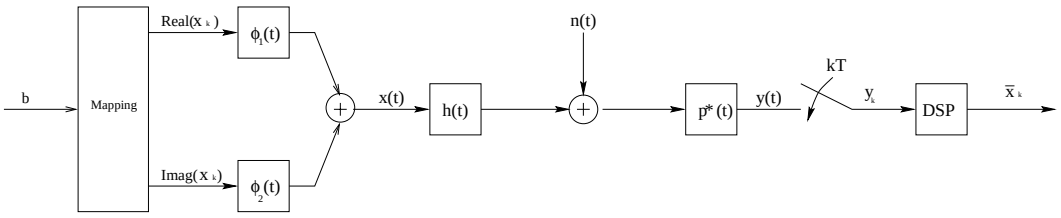

FIGURE 2.1: Block Diagram of General Digtal Communication System

b is the bitstream that is to be transmitted. These bits are mapped to complex symbols $x_k = \Re[x_k] + j\Im[x_k]$. $\Re$ denotes the real part and $\Im$ denotes the imaginary part. The symbol period is T, $\phi_1(t)$ and $\phi_2(t)$ are the transmit filters, $h(t)$ is the channel impulse response, $n(t)$ is the Additive White Gaussian Noise (AWGN), and $p^*(t)$ is the receiver matched filter. The output of the matched filter is sampled and processed to obtain the receive signal estimate $\bar{x}_k$.

2.1.1. Transmitter

The transmit filters for passband Quadrature Amplitude Modulation (QAM) system are given by :

$$\phi_1(t) = g(t)\cos(2\pi f_c t) \tag{2.1}$$

$$\phi_2(t) = g(t)\sin(2\pi f_c t) \tag{2.2}$$

where $\phi_1(t)$ and $\phi_2(t)$ are called in-phase and quadrature-phase components respectively, $g(t)$ is the pulse shaping filter, and f_c is the carrier frequency. Real and imaginary parts of the complex symbol x_k are passed through the transmit filters $\phi_1(t)$ and $\phi_2(t)$ and summed together to obtain the transmit signal $x(t)$. The transmit QAM signal can be written as :

$$\begin{aligned} x(t) &= \Re[x_k g(t) e^{j2\pi f_c t}] \\ &= \Re[x_k] g(t)\cos(2\pi f_c t) - \Im[x_k] g(t)\sin(2\pi f_c t) \end{aligned} \tag{2.3}$$

We can visualize QAM signal as two dimensional signal which lies in the complex plane (constellation) spanned by two bases functions $\phi_1(t)$ and $\phi_2(t)$. We may see from Eq.(2.3) that the transmit signal is complex symbol modulated by complex sinusoid, or equivalently, sum of real part of symbol modulated by $\phi_1(t)$ and imaginary part of the symbol modulated by $\phi_2(t)$.

For the purpose of understanding multicarrier systems, we will only focus on square QAM constellations. A square QAM constellation contains power of four points arranged regularly with a square boundary on the signal constellation. Fig.(2.2) shows a 16-QAM and 4-QAM contained within 16-QAM. The smallest distance between the points is given by d, and the constellation is centered at the origin. Each point on the constellation is assumed to occur equally likely and the average energy can be approximated by [2],

$$\mathcal{E} = \frac{(K-1)d^2}{6} \tag{2.4}$$

where K is the number of points in the constellation.

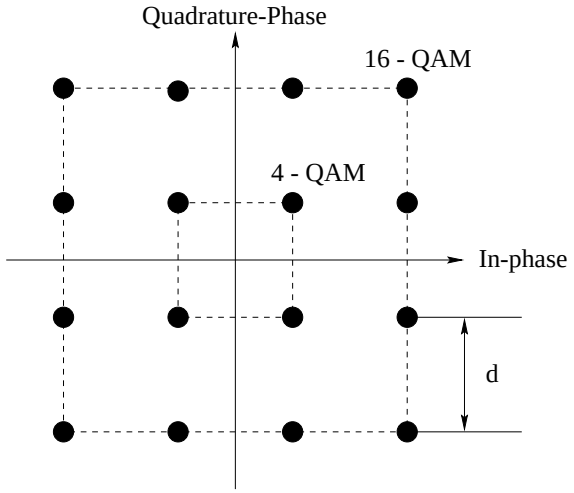

FIGURE 2.2: 4 and 16 point QAM Constellations

2.1.2. Channel

As shown in Fig.(2.1), modulated signal is transmitted through a *channel*, $h(t)$, which is the combined impulse response of all media seen by the signal after the transmitter to the receiver. For instance, the channel may include amplifiers and cables. First, to analyze the performance of the passband QAM system, we assume that the channel is Inter Symbol Interference (ISI) free, or perfectly equalized. The channel has attenuation of $|H|^2$ where H is the Frequency domain response of $h(t)$. The average transceiver output energy of the system is $\mathcal{E}$, and the noise variance of $n(t)$ per dimension is equal to σ^2. The received signal to noise ratio (SNR) is given by,

$$SNR = \frac{\mathcal{E}|H|^2}{2\sigma^2} \tag{2.5}$$

The received SNR gives us quantitative measure of the received signal quality. Low SNR would indicate that the received signal is highly corrupted and high SNR would indicate that the received signal is close to what was transmitted. We will now investigate the effect of the channel in receiver design.

2.1.3. Receiver

The minimum distance between the QAM points at the receiver, d_{min}, is given by,

$$d_{min}^2 = |H|^2 d^2 \tag{2.6}$$

Eq.(2.6) allows us to visualize the amount of degradation caused by the channel in the constellation diagram. Furthermore, in receiver design, Eq.(2.6) must be larger than predefined value of d_{min} in order to achieve prespecified amount of transmission error as we shall see later. Rewriting Eq.(2.4) with Eq.(2.6) and solving for K gives,

$$K = 1 + \frac{6\mathcal{E}|H|^2}{d_{min}^2} \tag{2.7}$$

Eq.(2.7) shows that given $\mathcal{E}$, H, and d_{min}, we may find the optimal constellation size that can be transmitted for this channel. The probability of symbol error, P_e, in QAM system is closely approximated by [2],

$$P_e \leq 4Q[\frac{d_{min}}{2\sigma}] \tag{2.8}$$

where $Q(x)$ is the Q function defined by [9],

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \tag{2.9}$$

This integral cannot be evaluated in closed form, and it is typically found by table lookup [9] or through computer simulation. Fig.(2.3) shows the theoretical Symbol Error Rate (SER) for various constellation sizes using Eq.(2.8). In designing the system, maximum tolerable error rate must be specified. We see from Eq.(2.8) that d_{min} can be specified for given symbol error rate and the AWGN variance.

2.1.4. Performance Assessment

By defining SNR gap, , [2],

$$, = \frac{d_{min}^2}{12\sigma^2} \tag{2.10}$$

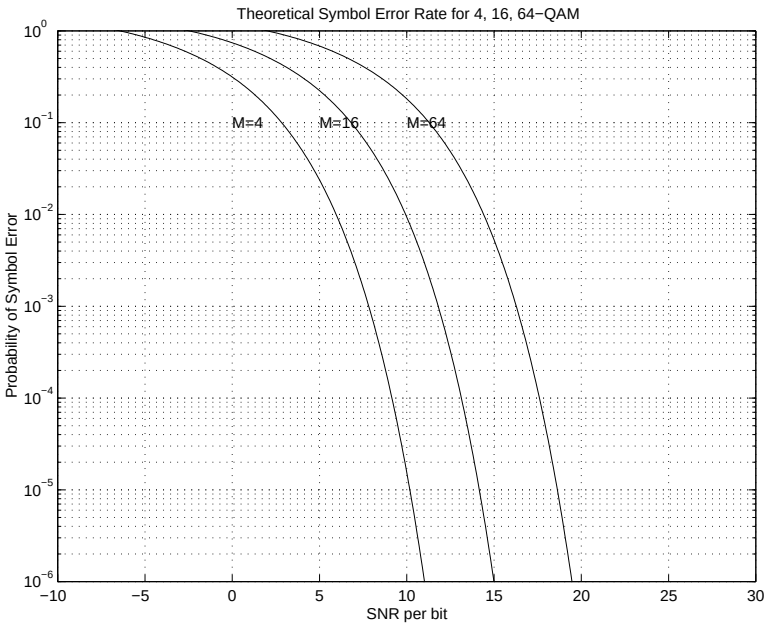

FIGURE 2.3: Theoretical Symbol Error Rate (SER)

and substituting Eq.(2.5) and Eq.(2.10) into Eq.(2.7), the number of points in the constellation can be written as,

$$K = 1 + \frac{SNR}{,} \tag{2.11}$$

The number of bits that can be reliably transmitted is,

$$b = \log_2 K = \log_2(1 + \frac{SNR}{,}) \tag{2.12}$$

with data rate,

$$R = \frac{b}{T} \quad bits/sec \quad (bps) \tag{2.13}$$

which is similar to the well known result derived by Shannon [33]. The interpretation of SNR gap is the loss in performance with respect to the optimum capacity given by $\log_2(1 + SNR)$. As , approaches one, the achievable data rate of the system approaches optimum. Then, given the probability of error that we would like to achieve, the channel characteristics, and the energy allowed for transmission, we may readily compute the achievable data rate across the channel.

2.1.5. Motivation for Multi-carrier Systems

In general, the channel transfer function is a function of frequency, $H(f)$. Then, the received SNR given by Eq.(2.5) is also a function of frequency. In single carrier systems, the bandwidth of a QAM signal must be designed such that satisfactory data rate can be achieved for a given channel. This is illustrated in top figure of Fig.(2.4). In this example, we have selected 650KHz of bandwidth for single QAM signal. The maximum allowed received signal SNR for such system is 10dB with the sample channel SNR as shown. Then using Eq.(2.12), maximum allowed data rate is computed. In other words, *amount of received SNR used* gives direct indication of the amount of the data that can be transmitted through the channel. From this argument, we can see that the top figure of Fig.(2.4) is highly inefficient. However, if we partition the channel into a set

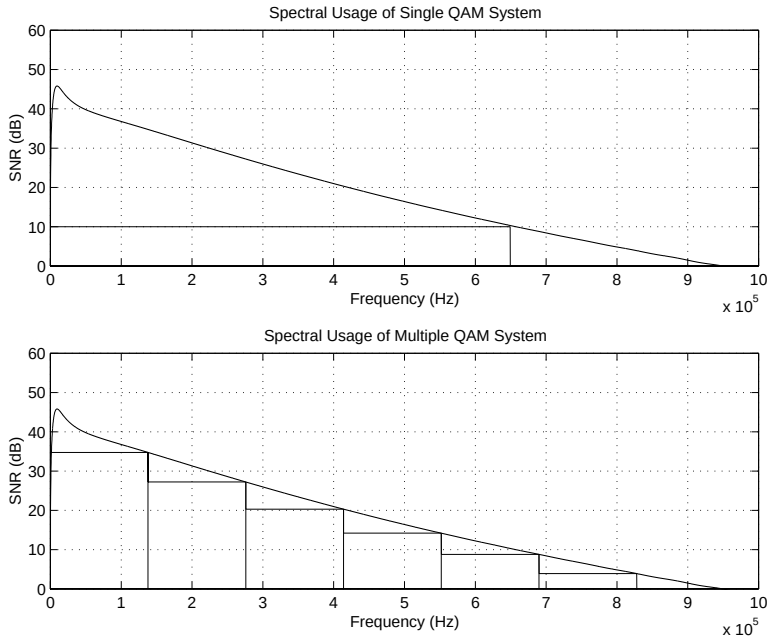

FIGURE 2.4: Channel SNR usage of Single & Multiple QAM Systems

of narrowband subchannels, the bottom figure of Fig.(2.4) is obtained which makes far better use of the channel. In fact, as the number of subchannels are increased, the optimal channel capacity is obtained. The optimally of such channel partitioning is proved in [31]. This is the underlying principle in using multi-carrier modulation in order to maximize the use of available channel resources.

2.2. Multi-Carrier Analysis

2.2.1. General Architecture

Multicarrier modulation method uses an optimized frequency division allocation of energy and bits into multiple channels to maximize achievable data rates over bandlimited communication channels [10]. These systems can be easily described as multiplexing of single carrier signals over multiple channels. The general structure of multicarrier system is shown in Fig.(2.5). An input bit stream with data rate R bps is buffered over some time

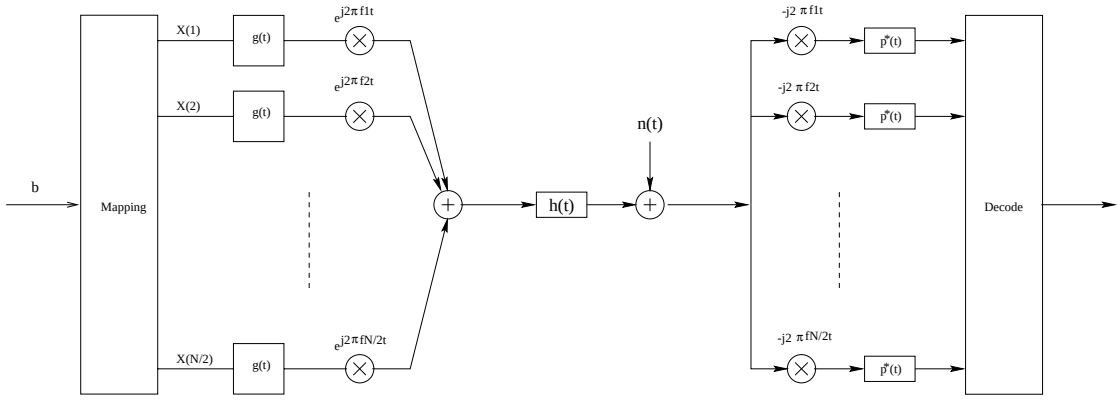

FIGURE 2.5: General Multicarrier System Diagram

T, where T is called the frame period in seconds and $1/T$ is called the frame rate. The number of bits in the buffer is $b = RT$. These b bits are partitioned to $N/2$ subchannels so that b_i (for $i = 1, 2, ..., N/2$) bits are transmitted through i-th subchannel assuming N even. That is,

$$b = \sum_{i=1}^{N/2} b_i \tag{2.14}$$

These b_i bits for each subchannel 1 through N/2 are mapped to complex symbols $X(i)$ ($i = 1, 2, ..., N/2$). Each of these symbols are then modulated by frequencies $f_1, f_2, \cdots, f_{N/2}$ as shown in Fig.(2.5).

This method has been called *Orthogonal Frequency Division Multiplexing* (OFDM) or *Orthogonally Multiplexed QAM* [34] [35] because each of the subchannels are QAM modulated signals with different carrier frequency and each subchannels are orthogonal to each other. However, for implementation purposes, having so many modulators and demodulators are prohibitively expensive. It was noticed by [36] that equivalent of multiple modulators can be realized by using the Discrete Frequency Transform and this method has been called *Discrete MultiTone* (DMT) modulation. With recent developments in hardware and software algorithms, DFT/Inverse DFT (IDFT) is efficiently realized by Fast Fourier Transform (FFT) and Inverse FFT (IFFT) which is very attractive for Large Scale Integrated (LSI) circuits implementation. Fig.(2.6) shows the block diagram of transmitter to receiver in DMT system. The buffer and encoder performs the same operation as described before. Varying number of bits are allocated to each subchannels. These bits are then mapped to appropriate QAM symbols to obtain $X(k)$ for $k = 0, 1, \cdots, N/2$. Next, $X(k)$ must be mapped before IFFT to $\hat{X}(k)$ such that the output of IFFT is purely real. This is because the transmit signal cannot be complex.

First, let us define a normalized N-point DFT and IDFT as

$$\hat{X}(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x(n) e^{-j2\pi nk/N} \qquad \text{for} \quad 0 \le k \le N-1 \tag{2.15}$$

$$x(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \hat{X}(k) e^{j2\pi nk/N} \qquad \text{for} \quad 0 \le n \le N-1 \tag{2.16}$$

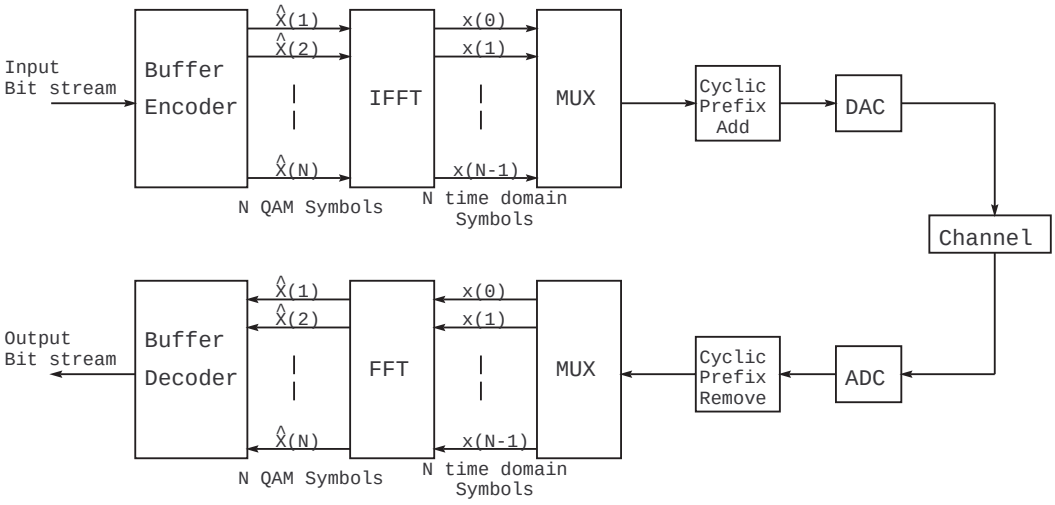

FIGURE 2.6: Discrete Multitone System Diagram

DFT/IDFT in general yields complex numbers. We may make the output of the IDFT to be purely real by ordering $X(k)$ into Hermitian Symmetry.

$$\hat{X}(0) = \Re\{X(1)\}, \qquad \hat{X}(k) = X(k) \qquad \text{for} \quad 2 \leq k \leq N/2 \tag{2.17}$$

$$\hat{X}(N/2) = \Im\{X(1)\}, \quad \hat{X}(N-k) = X^*(k+1) \qquad \text{for} \quad 1 \leq k \leq N/2 - 1 \tag{2.18}$$

where X^* denotes complex conjugate of X. Then, taking IDFT of expression (2.17) and (2.18) yields,

$$x(n) = \frac{1}{\sqrt{N}}\{\hat{X}(0) + \hat{X}(N/2)\cos(\pi n) + 2\sum_{k=1}^{N/2-1} |\hat{X}(k)|\cos(2\pi nk/N + \angle\hat{X}(k)) \tag{2.19}$$

which is purely real time domain signal [32]. We can see from Eq.(2.19) that each of the complex symbols $\hat{X}(k)$ has been modulated by frequencies $0, 1/N, 2/N, ..., (N/2 - 1)/N$. We may visualize the IDFT operation as set of parallel transmit filters, or a set of orthonormal basis functions which modulates $N/2$ complex frequency domain symbols to $N/2$ frequency division multiplexed sinusoids.

2.2.2. Cyclic Prefix

One of the underlying assumptions in using FFT/IFFT is that the signal to be transformed is a periodic signal [30]. In general, this is not the case in DMT systems. Aperiodic signal is transmitted and this is linearly convolved with the channel impulse response. To validate the use of FFT/IFFT, it has been suggested by [14] to modify the transmit signal in such a way that the channel output appears as if it has been circularly convolved with the channel. This method is called the cyclic prefix.

Cyclic prefix is performed by copying L samples from the end of the time domain frame and appending it prior to the original frame. Thus, if $\mathbf{x}$ is the original time domain frame,

$$\mathbf{x} = [x(0) \quad x(1) \quad \cdots \quad x(N-1)]^T \tag{2.20}$$

the cyclic prefixed frame with L cyclic prefix becomes,

$$\mathbf{x}_{cp} = [\underbrace{x(N-L) \quad x(N-L+1) \quad \cdots \quad x(N-1)}_{Cyclic \quad Prefix} \quad \underbrace{x(0) \quad x(1) \quad \cdots \quad x(N-1)}_{Original \quad Data}]^T \tag{2.21}$$

where T denotes the transpose of the vector. At the receiver side, samples corresponding to cyclic prefix are removed before demodulation. The use of cyclic prefix results in effectively no Interframe Interference at the receiver [15] assuming that the channel has impulse response length that is shorter than the prescribed cyclic prefix length. For example, let N=5, L=2, and channel impulse response length M=2. The cyclic prefixed samples become,

$$\mathbf{x}_{cp} = [\underbrace{x(3) \quad x(4)}_{Cyclic \quad Prefix} \quad \underbrace{x(0) \quad x(1) \quad x(2) \quad x(3) \quad x(4)}_{Original \quad Data}]^T \tag{2.22}$$

Table.(2.1) illustrates the linear convolution of cyclic prefixed data and the channel impulse response. Effect of circular convolution can be imitated by repeating the sequence $x(0), x(1), \cdots, x(4)$ and linearly convolving with channel impulse response $h(0)$ and $h(1)$.

$t=-2$	$*$	$x(3)$	$x(4)$	$x(0)$	$x(1)$	$x(2)$	$x(3)$	$x(4)$
	$h(1)$	$h(0)$						
	$y(-2) = h(0)x(3) + h(1)*$							
$t=-1$	$*$	$x(3)$	$x(4)$	$x(0)$	$x(1)$	$x(2)$	$x(3)$	$x(4)$
		$h(1)$	$h(0)$					
	$y(-1) = h(0)x(4) + h(1)x(3)$							
$t=0$	$*$	$x(3)$	$x(4)$	$x(0)$	$x(1)$	$x(2)$	$x(3)$	$x(4)$
		$h(1)$	$h(0)$					
	$y(0) = h(0)x(0) + h(1)x(4)$							
$t=1$	$*$	$x(3)$	$x(4)$	$x(0)$	$x(1)$	$x(2)$	$x(3)$	$x(4)$
				$h(1)$	$h(0)$			
	$y(1) = h(0)x(1) + h(1)x(0)$							
$t=2$	$*$	$x(3)$	$x(4)$	$x(0)$	$x(1)$	$x(2)$	$x(3)$	$x(4)$
					$h(1)$	$h(0)$		
	$y(2) = h(0)x(2) + h(1)x(1)$							
$t=3$	$*$	$x(3)$	$x(4)$	$x(0)$	$x(1)$	$x(2)$	$x(3)$	$x(4)$
						$h(1)$	$h(0)$	
	$y(3) = h(0)x(3) + h(1)x(2)$							
$t=4$	$*$	$x(3)$	$x(4)$	$x(0)$	$x(1)$	$x(2)$	$x(3)$	$x(4)$
							$h(1)$	$h(0)$
	$y(4) = h(0)x(4) + h(1)x(3)$							

TABLE 2.1: Graphical Convolution with Cyclic Prefix

In Table.(2.1), we see that the same effect is achieved. More compactly, we may write this in matrix notation,

$$\begin{bmatrix} y(-2) \\ y(-1) \\ y(0) \\ y(1) \\ y(2) \\ y(3) \\ y(4) \end{bmatrix} = \begin{bmatrix} x(3) & * \\ x(4) & x(3) \\ x(0) & x(4) \\ x(1) & x(0) \\ x(2) & x(1) \\ x(3) & x(2) \\ x(4) & x(3) \end{bmatrix} \begin{bmatrix} h(0) \\ h(1) \end{bmatrix} \tag{2.23}$$

where $*$ is an unknown sample from the previous frame. We may see that if $y(-2)$ and $y(-1)$ is discarded from the received frame (samples corresponding to the cyclic prefix), the remaining output signal appears as if the input data $\mathbf{x}$ was cyclically convolved with the channel. However, if the channel response is longer than the specified cyclic prefix length, cyclic prefix is not sufficient to remove the tail of the impulse response and we require equalization known as Time Domain Equalizer (TEQ). The objective of TEQ is to shorten long channel impulse response such that the Inter-frame Interference can be removed by cyclic prefix [10]. Moreover, in designing the length of cyclic prefix, one must determine the trade-off between data rate and channel impulse response length since increase in cyclic prefix will result in lower throughput due to the redundancy introduced. Cyclic prefix operation and its concept will play a vital role in the design of the echo canceller described in Chapter 3.

2.2.3. Other System Blocks

The cyclic prefixed $N + L$ time domain samples are then successively applied to Digital-to-Analog (D/A) converter which gives the continuous time modulated signal $x(t)$. In the receive path, received signal $y(t) = x(t) * h(t)$ is sampled by Analog-to-Digital (A/D) converter where $*$ denotes linear convolution. After removal of cyclic prefix and

demodulation by FFT, $N/2$ frequency domain symbols are obtained which are decoded to obtain the output bit sequence.

2.3. ADSL

Asymmetric Digital Subscriber Line (ADSL) uses DMT for multicarrier modulation [16]. Asymmetricity comes from the fact that the data transfer required from central office to the customer premise is different than that from customer premise to the central office. The ADSL Transceiver Unit for Remote Terminal or the customer premise side is called ATU-R. ADSL Transceiver Unit for the Central Office side is called ATU-C. Asymmetric data rate is achieved by taking different bandwidth for FFT/IFFT pair in each directions. The upstream, from ATU-R to ATU-C, uses the Nyquist frequency at 276KHz with 64 point FFT/IFFT and cyclic prefix length of 4. The downstream, from ATU-C to ATU-R, the Nyquist frequency is taken at 1.104MHz with 512 point FFT/IFFT and cyclic prefix length of 32. Due to the Hermitian symmetry as explained in Section 2.2.1, the number of subchannels used are half of the FFT/IFFT sizes. Therefore, downstream uses 256 subchannels, whereas upstream uses 32 subchannels. This is summarized in Table.(2.2). The difference in data rate is determined to be a factor of eight from surveys in use of this technology. The frame rate, $1/T$, is $4312.5Hz$. Fig.(2.7) shows the overall block diagram with the bus width shown. The spectral allocation of downstream can be either using Frequency Division Multiplexing (FDM) or Echo Cancelled (EC) as described in the next section.

2.3.1. Multiplexing in ADSL

Multiplexing in ADSL can be accomplished in one of two ways. One is Frequency Division Multiplexing (FDM) where the upstream and downstream spectrums are sepa-

	Downstream (ATU-C to ATU-R)	Upstream (ATU-R to ATU-C)
A: FFT/IFFT	512 points	64 points
B: Nyquist Freq.	1.104MHz	276KHz
C: CP Length	32	4
D: Subchannel Bandwidth	4312.5Hz	4312.5Hz
E: # of Subchannels	256	32

TABLE 2.2: Summary of ADSL Specification

rated. This is typical of what is found in radio communications where each user is assigned to different frequency bands. Fig.(2.8) shows the spectral usage of the FDM system. The frequency separation is accomplished by frequency selective filtering.

Second approach is to use overlapped spectrum using echo canceller to differentiate the two signals. Fig.(2.9) shows the spectral usage of EC system. In FDM system, a filter is required after coming in from the hybrid as illustrated in Fig.(2.10). In EC system, no filtering is required, but instead digital implementation of echo canceller is required. The use of EC system can be justified by comparing the spectral usage shown in Fig.(2.8) and Fig.(2.9). In EC system, wider downstream bandwidth is used compared to FDM system. Therefore EC system is capable of transmitting at higher data rate in downstream than FDM system.

2.3.2. Frame Synchronization

Frame synchronization (or alignment) in ADSL system plays an important role in the design of the echo canceller. Recall that each frame is given by $N + L$ time domain

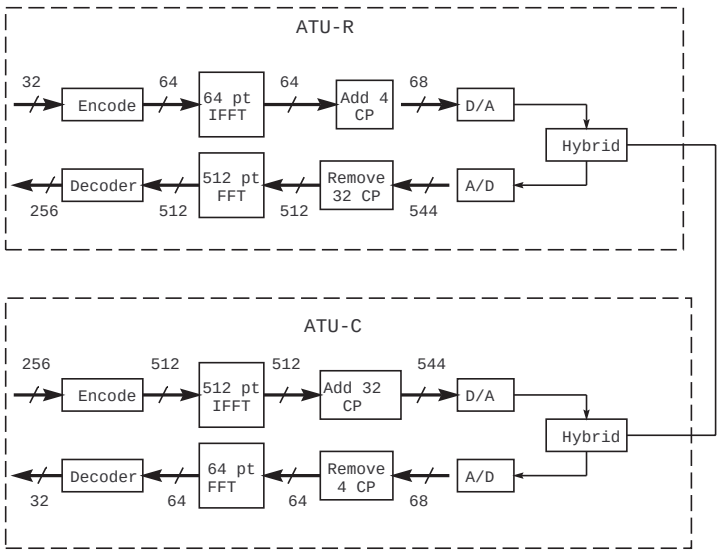

FIGURE 2.7: General ADSL Block Diagram with the Bus widths

samples. When a frame is transmitted and received at the other end, there is a time delay t_p due to propagation delay of the twisted pair cable. If both of the transceivers transmit at the same time instance, the received frames arrive asynchronous with respect to the transmitted frames as is shown in Fig.(2.11). In the receive path, the demodulation FFT is synchronized to the received frame during the initialization process. However, the echo signal couples into the received frame with negligible delay with respect to t_p (i.e. the echo frame arrival is any time in the middle of received frame). The demodulated signal after FFT, therefore, consists of properly demodulated receive signal and "*partial*" information of the echo signal. Hence, the echo canceller must perform all of its operation in the time domain prior to demodulation. This can be simplified by transmitting ATU-C frame with t_p advance time with respect to the ATU-R transmit frame as illustrated in Fig.(2.12). During the initialization, transmitter's frame timing at ATU-C is adjusted such that the transmit frame and receive frame at the remote side are synchronized. The demodulated

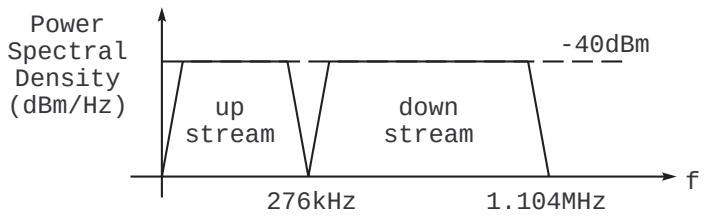

FIGURE 2.8: Spectral Usage of FDM System

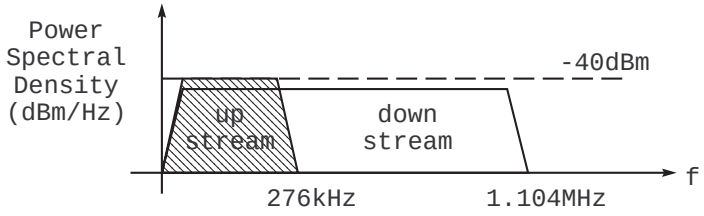

FIGURE 2.9: Spectral Usage of EC System

signal at the RT side contains full information of both received signal and echo signal. Since the ATU-R received frame and echo frame are synchronized, frequency domain echo canceller can be used which is much simpler than time domain echo canceller. Based on this, the technique of synchronous echo canceller is used for ATU-R and asynchronous echo canceller for ATU-C.

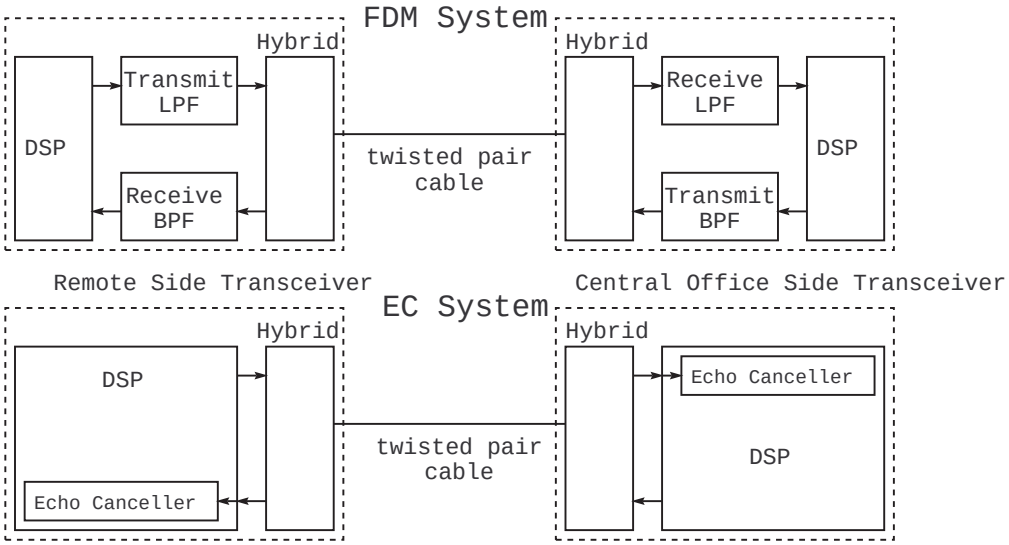

FIGURE 2.10: FDM and EC System Architectural Difference

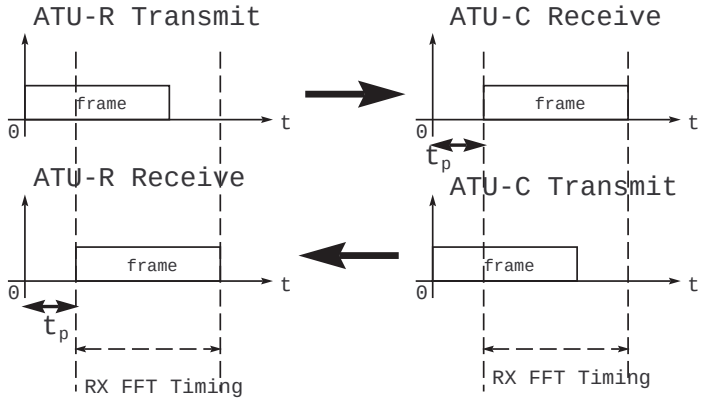

FIGURE 2.11: Asynchronous Timing Diagram

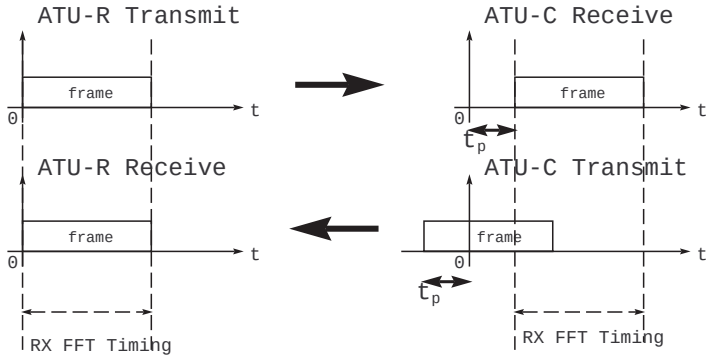

FIGURE 2.12: Synchronous Timing Diagram

3. ADSL ECHO CANCELLATION

3.1. Echo Cancellation : General Overview

Echo is a phenomenon in which a delayed and, perhaps, distorted version of an original signal is reflected back to the source [17]. In ADSL, the echo is caused by the hybrid transformer used for four-wire to two-wire conversion. Figure (3.1) shows the functional block diagram of end to end twisted pair data communication system. We may

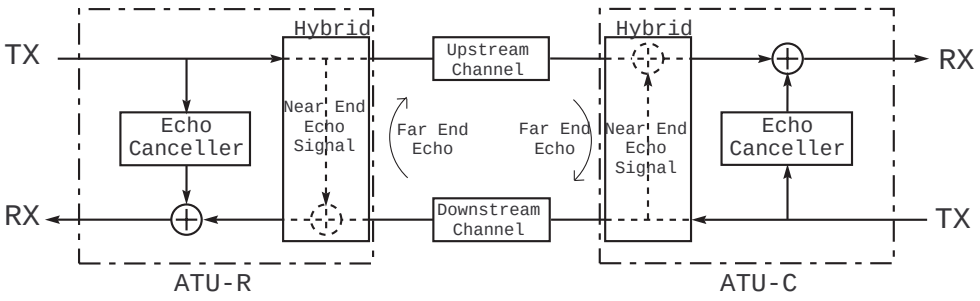

FIGURE 3.1: Twisted Pair End to End Block Diagram

identify two types of echos : near end echo and far end echo. The most dominant and problematic type is the near end echo caused by the hybrid's inability to match the line impedance. The far end echo occurs at the far end modem's hybrid and is reflected back across the channel. Since the far end echo is attenuated by the channel, it is assumed to be negligible.

Fig.(3.2) shows the basic structure of the hybrid. The sum of transmit and receive signal current in the twisted pair cable is I_{loop} and the current inside the matching circuit is I_{hyb}. The functions of the hybrid are impedance matching and isolation between the transceiver and the twisted pair cable. Ideally, all transformers are matched and the

balancing network is matched to the twisted pair cable impedance. Under such ideal conditions, signal originating on the "TX" side (I_{TX}) of the four-wire circuit gets coupled to the two-wire circuit, but produces no response in the "RX" side of the four-wire circuit since $I_{hyb} = I_{TX}$. A signal originating in the two wire circuit couples to both paths of the four-wire circuit, however, it has no effect on the "TX" side because the amplifier in the TX path points in the opposite direction. If the bridge is not perfectly balanced, the

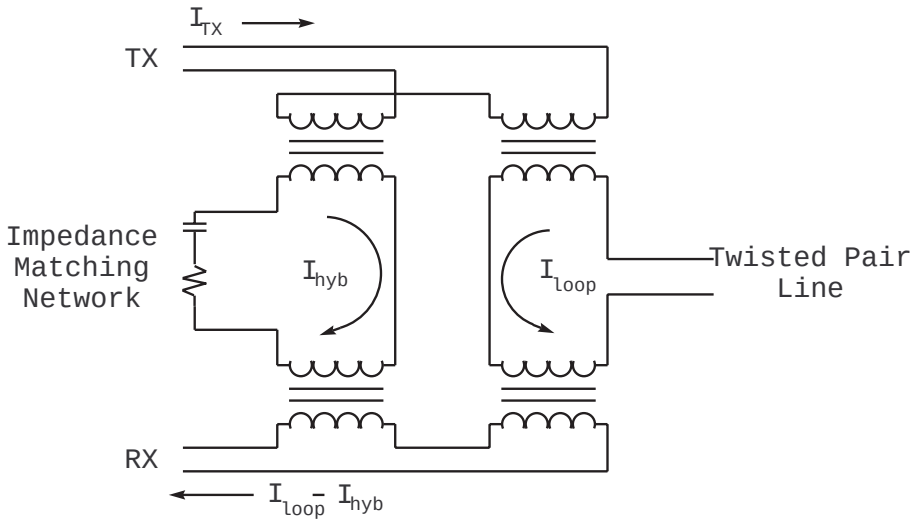

FIGURE 3.2: Simple Hybrid Circuit

"TX" side is coupled to the "RX" side of the hybrid ($I_{hyb} \neq I_{TX}$), thus giving rise to an echo ($I_{echo} = I_{hyb} - I_{TX}$). Furthermore, echo is not only an attenuated version of the input signal, but a filtered input signal due to the balancing network's lumped elements. The line impedance is largely dependent on the line type and length. In general, it is very difficult to compensate for all variations of line impedances to eliminate echo. However, there is such work as [21] to adaptively match the impedance to minimize echo. In this study, we do not include such adaptive impedance matching, and we treat the echo path

response as a fixed "channel" for a given subscriber loop. Therefore, the transmitted signal is passed through an "*echo channel*" and added to the received signal. Since modelling of the hybrid requires intense circuit modelling, this is beyond the scope of this thesis. For the purpose of testing the echo canceller, a simple 10th order Infinite Impulse Response (IIR) model was generated [7] [8]. Two generated sample impulse responses are shown in Fig.(3.3) For ADSL, transmit and receive sampling rates are different. Since the echo

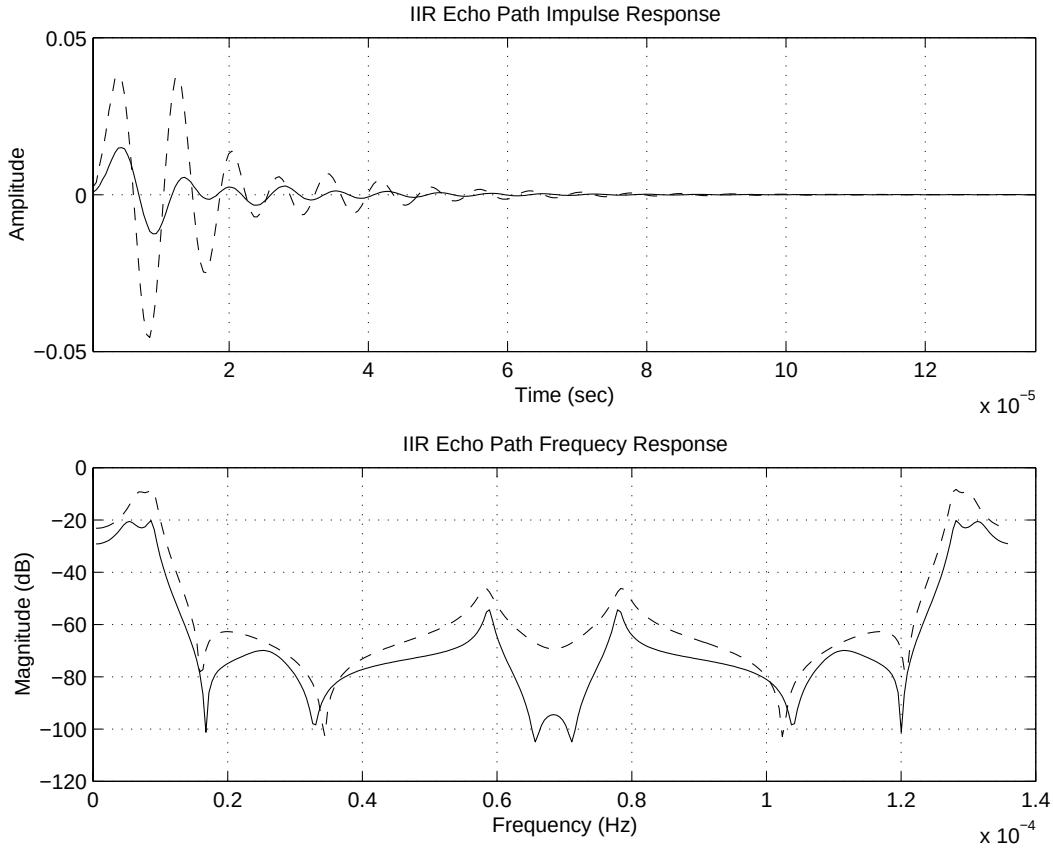

FIGURE 3.3: Sample Echo Path Response

canceller operates *both* on the transmit signal and the receive signal, multirate signal processing is involved. The echo path channel is relatively stationary but it may change

with time, temperature, or during on/off hook transition of a telephone. Therefore, echo canceller must continuously adapt itself to avoid loss of data.

3.2. Common Framework for ADSL Echo Cancellation

Echo canceller for ADSL has been rapidly developing area in the past decade [3] [5] [4] [7] [8] [6]. In this thesis, the latest architecture proposed by [6] is used. First, we will present the common framework for understanding ATU-R echo canceller by assuming symmetric rate. Then, the common framework is extended to ATU-R multirate echo canceller.

We model the echo signal as a linear convolution of cyclically prefixed data symbols x_i and the echo channel impulse response h. The length of echo channel is assumed to be $M \leq N$, where N is the FFT size. Therefore, we have,

$$\mathbf{h} = [h(1) \quad h(2) \quad \cdots \quad h(M) \quad \underbrace{0 \quad \cdots \quad 0}_{N-M \quad Zeros}]^T \tag{3.1}$$

$$\mathbf{x}_i = [x_i(N-L+1) \quad x_i(N-L+2) \cdots x_i(N) \quad x_i(1) \quad x_i(2) \cdots x_i(N)]^T \tag{3.2}$$

where L is the length of cyclic prefix and $\mathbf{x}_i$ is the ith frame transmitted. Then, $N+L$ samples received echo frame can be written as

$$\mathbf{y}_i = \mathbf{T}_{i,i-1} \cdot \mathbf{h} + \mathbf{n}_i \tag{3.3}$$

where $\mathbf{y}_i$ is the received frame vector given by,

$$\mathbf{y}_i = [y(N-L+1) \quad y_i(N-L+2) \quad \cdots \quad y_i(N) \quad y_i(1) \quad y_i(2) \quad \cdots \quad y_i(N)]^T \tag{3.4}$$

$\mathbf{T}_{i,i-1}$ is the convolution matrix,

$$\mathbf{T}_{i,i-1} = \begin{bmatrix} x_i(N-L+1) & x_{i-1}(N) & x_{i-1}(N-1) & \cdots & x_{i-1}(1) & x_{i-1}(N) & \cdots & x_{i-1}(N-L+2) \\ x_i(N-L+2) & x_i(N-L+2) & x_{i-1}(N) & \cdots & x_{i-1}(2) & x_{i-1}(1) & \cdots & x_{i-1}(N-L+3) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ x_i(N) & x_i(N-1) & x_i(N-2) & \cdots & x_i(N-L+1) & x_{i-1}(N) & \cdots & x_{i-1}(N-L+1) \\ x_i(1) & x_i(N) & x_i(N-1) & \cdots & x_i(N-L+2) & x_i(N-L+1) & \cdots & x_{i-1}(N-L+2) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ x_i(N) & x_i(N-1) & x_i(N-2) & \cdots & x_i(N-L+1) & x_i(N-L) & \cdots & x_i(1) \end{bmatrix} \tag{3.5}$$

and $\mathbf{n}_i$ is the distorted data from the far-end transmitter plus a white additive Gaussian noise.

Since $M > N$, there is ISI caused by the symbol transmitted at instance $i-1$ as indicated at the upper right diagonals of $\mathbf{T}_{i,i-1}$. However, if the echo path channel impulse response length $M < L$, linear convolution becomes circular convolution and the following circulant matrix $\mathbf{C}_i$ is equivalent to $\mathbf{T}_{i,i-1}$.

$$\mathbf{C}_i = \begin{bmatrix} x_i(N-L+1) & x_i(N) & \cdots & x_i(1) & x_i(N) & \cdots & x_i(N-L+2) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ x_(N) & x_i(N-1) & \cdots & x_i(N-L+1) & x_i(N) & \cdots & x_i(1) \\ x_(1) & x_i(N) & \ddots & \cdots & \cdots & \ddots & x_i(2) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ x_i(N) & x_i(N-1) & \cdots & \cdots & \cdots & \cdots & x_i(1) \end{bmatrix} \tag{3.6}$$

The unique property of the circulant matrix is that if we take the FFT of $\mathbf{C}_i$, the resultant matrix is a diagonal matrix composed of frequency domain frame $\mathbf{X}$ [6]. Therefore, we can simplify the convolution by element by element multiplication in the frequency domain as follows.

$$\mathbf{Y}_i = \mathbf{F}_N \mathbf{C}_i \mathbf{F}_N^{-1} \mathbf{F}_N \mathbf{h} \tag{3.7}$$

$$= diag(\mathbf{X}_i)\mathbf{H} \tag{3.8}$$

where $\mathbf{F}_N$ denotes the FFT matrix, $\mathbf{H}$ is the frequency domain echo path response, and $diag(\mathbf{X}_i)$ denotes diagonal matrix with complex symbols $\mathbf{X}_i$ along the diagonal.

With this background, we identify two basic components of ADSL echo canceller. First, Circular Echo Synthesis (CES) will operate on $\mathbf{T}_{i,i-1}$ matrix to modify the received frame to appear circularly convolved with the echo path impulse response. After cyclic prefix stripping and FFT, FRequency domain Echo Canceller (FREC) will operate on the frequency domain frame $\mathbf{X}_i$ to remove the echo signal.

3.2.1. Circular Echo Synthesis

We saw in the previous argument that the received frame matrix $\mathbf{T}_{i,i-1}$ is not circulant. Therefore, taking the FFT of this matrix will result in non-diagonal matrix, implying that we no longer have parallel independent subchannels. The purpose of CES is to minimize the components in the off diagonal elements such that the result of FFT is diagonal, or to make $\mathbf{T}_{i,i-1}$ appear circulant. We can mathematically separate the diagonal and non-diagonal parts as,

$$\begin{aligned} \mathbf{Y}_i &= \mathbf{F}_N(\mathbf{T}_{i,i-1} + \mathbf{C}_i - \mathbf{C}_i)\mathbf{h} && (3.9) \\ &= diag(\mathbf{X}_i)\mathbf{H} + \mathbf{F}_N\mathcal{X}_{i,i-1}\mathbf{h} && (3.10) \end{aligned}$$

where $\mathcal{X}_{i,i-1} = \mathbf{T}_{i,i-1} - \mathbf{C}_i$. If we can estimate $\mathbf{F}_N\mathcal{X}_{i,i-1}\mathbf{h}$ and subtract it from the received frame, we have orthogonalized the received echo frame to operate in the frequency domain. Note, that in theory, it is possible to ignore this step and obtain a square matrix in the frequency domain, but the computational burden in dealing with such matrices is high. Recalling the structure of $\mathbf{T}_{i,i-1}$, the matrix $\mathcal{X}_{i,i-1}$ is an upper triangular matrix with all other elements zero.

The uniqueness of the echo cancellation problem is that the transmit or reference signal is completely known. Therefore, it is a relatively simple task to construct $\mathcal{X}_{i,i-1}$ from the transmitted symbols. This is usually referred to as "*Data-Driven Echo Canceller*" [3]. Then, given an adequate estimate of echo path impulse response, w_i, we can subtract the off diagonal components from the received symbols. That is,

$$\begin{aligned} \hat{\mathbf{Y}}_i &= \mathbf{F}_N(\mathbf{T}_{i,i-1}\mathbf{h} - \mathcal{X}_{i,i-1}\mathbf{w}_i) && (3.11) \\ &\approx diag(\mathbf{X}_i)\mathbf{H} \end{aligned}$$

given that $\mathbf{w}_i \approx \mathbf{h}$.

We will illustrate the idea behind the CES through a simple example. Assume that two successive symbols have been transmitted with $L = 2$, $N = 4$, and $M = 4$. The

transmitted symbols are,

$$\begin{aligned}\mathbf{x}(i) &= [x_{-1}(2) \quad x_{-1}(3) \quad x_{-1}(0) \quad x_{-1}(1) \quad x_{-1}(2) \quad x_{-1}(3) \\ &\qquad x_0(2) \quad x_0(3) \quad x_0(0) \quad x_0(1) \quad x_0(2) \quad x_0(3)] \\ &for \qquad i = -8, -7, \cdots, -1, 0, 1, 2, 3\end{aligned} \tag{3.12}$$

The index i is chosen such that $i = 0$ is at $x_0(0)$. The channel impulse response is,

$$\mathbf{h} = [h(0) \quad h(1) \quad h(2) \quad h(3)] \tag{3.13}$$

The noiseless output of the echo path channel for $0th$ symbol becomes,

$$\begin{aligned}\mathbf{y}_0 &= \mathbf{T}_{0,-1}\mathbf{h} \\ \begin{bmatrix} y_0(-2) \\ y_0(-1) \\ y_0(0) \\ y_0(1) \\ y_0(2) \\ y_0(3) \end{bmatrix} &= \begin{bmatrix} x_0(2) & x_{-1}(3) & x_{-1}(2) & x_{-1}(1) \\ x_0(3) & x_0(2) & x_{-1}(3) & x_{-1}(2) \\ x_0(0) & x_0(3) & x_0(2) & x_{-1}(3) \\ x_0(1) & x_0(0) & x_0(3) & x_0(2) \\ x_0(2) & x_0(1) & x_0(0) & x_0(3) \\ x_0(3) & x_0(2) & x_0(1) & x_0(0) \end{bmatrix} \begin{bmatrix} h(0) \\ h(1) \\ h(2) \\ h(3) \end{bmatrix}\end{aligned} \tag{3.14}$$

The cyclic prefix samples, $y_0(-2)$ and $y_0(-1)$, are discarded before taking the FFT. We can see that for $y_0(0)$, there is a contribution of $x_{-1}(3)$ that is preventing $\mathbf{T}_{0,-1}$ from becoming a circulant matrix. In CES, we first find,

$$\begin{aligned}\mathcal{X}_{0,-1} &= \mathbf{T}_{0,-1} - \mathbf{C}_0 \\ &= \begin{bmatrix} 0 & 0 & 0 & x_{-1}(3) - x_0(1) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\end{aligned} \tag{3.15}$$

Next, by multiplying $\mathcal{X}_{0,-1}$ with the estimate of the echo path channel $\mathbf{w}_i$, and subtracting it from the received frame vector $\mathbf{y}_0$, we obtain,

$$\hat{\mathbf{y}}_0 = \mathbf{y}_0 - \mathcal{X}_{0,-1}\mathbf{w}_i \tag{3.16}$$

If $\mathbf{w}_i \approx \mathbf{h}$ or more importantly, if $\mathbf{w}_i(3) \approx \mathbf{h}(3)$, the new received frame is orthogonalized into parallel independent subchannels in the frequency domain.

Now the remaining problem in the echo cancellation is to acquire a good estimate of the echo path channel $\mathbf{h}$. This can be done by any adaptation techniques such as Least Mean Square (LMS) or Recursive Least Squares (RLS). The existing method uses Block Frequency Domain LMS method which is outlined next.

3.2.2. Frequency Domain Echo Cancellation

After CES is performed, the received frame is demodulated by FFT. The advantage here is that no extra hardware is required to obtain the frequency domain symbols because FFT is already part of the system. Secondly, block frequency domain processing is faster and simpler because implementation of large adaptive FIR would involve large latency and cost prohibiting computations [23]. Although this is a form of fast convolution, it differs from traditional overlap and save method and overlap and add method [30]. Moreover, if overlap and save or overlap and add methods were used, modifications to the receive path must be made to accommodate the technique (such as doubling the FFT size). The idea of orthogonalizing the echo frame prior to FFT allows us to use the existing FFT in the receive path without system modification.

The purpose of Frequency Domain Echo Canceller (FREC) is to acquire an estimate of the echo path impulse response based on the knowledge of the transmitted signal and the received signal. The estimate of the echo path impulse response is first used in CES to orthogonalize the received frame matrix, and second in frequency domain to subtract estimated echo signal from the received signal. In [6], LMS was used as the adaptation scheme for its simplicity. In LMS, estimated echo signal is subtracted from the received signal to obtain an error signal. The error signal is then used in update equation to obtain the new estimate of the echo path impulse response. Using the updated echo path impulse

response, new estimate of echo signal is obtained and the procedure is repeated. The error signal is obtained by,

$$\mathbf{E}_i = \mathbf{F}_N(\mathbf{y}_i - \mathcal{X}_{i,i-1}\mathbf{w}_i) - diag(\mathbf{X}_i)\mathbf{W}_i \tag{3.17}$$

where $\mathbf{w}_i$ is the time domain estimated echo path impulse response for frame i and $\mathbf{W}_i$ is the corresponding frequency domain echo path response. The first term on the right hand side is the FFT of the CES output, or desired signal which we would like to cancel, and the second term on the right is the estimated echo signal in the frequency domain. The update equation becomes,

$$\mathbf{W}_{i+1} = \mathbf{W}_i + \mu diag(\mathbf{X}_i^*)\mathbf{E}_i \tag{3.18}$$

where μ is the step size parameter.

For our analysis, the tap weight estimate $\mathbf{W}_i$ is assumed to be updated every frame period and the corresponding time domain tap weight estimate $\mathbf{w}_i$ is also assumed to be computed via IFFT every frame period. In actual implementation, trade-off must be made in the amount of computations allowed during the frame period versus frequency of the update. For example, if the computation required for update requires three frame periods, tap weights can only be updated every third frame rather than every frames. Fig.(3.4) shows the block diagram of the common framework ADSL echo canceller presented thus far.

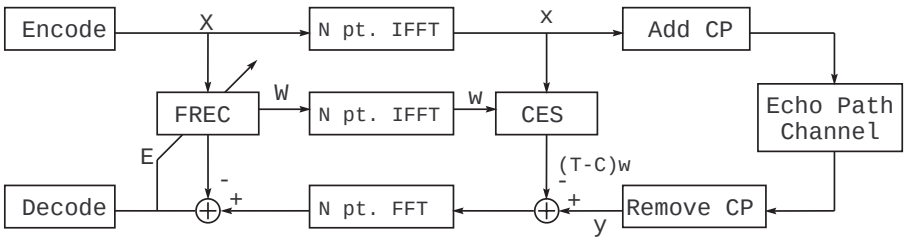

FIGURE 3.4: Common Framework ADSL Echo Canceller

3.3. Multirate ATU-R Echo Cancellation

As described in ADSL overview section, the downstream and the upstream data rates used in ADSL are asymmetric. The upstream bandwidth is exactly one eighth of the downstream bandwidth. Hence, when the upstream echo is present in the downstream, the upstream signal is effectively upsampled by a factor of eight, and likewise downstream echo signal into upstream is downsampled by a factor of eight. In this section, the problem of echo cancellation is addressed when the transmit bandwidth is k times (in ADSL $k = 8$) smaller than the receive bandwidth and the frame alignment is synchronized.

Ideally, only those bandwidth in which the transmit bandwidth and the receive bandwidth overlap will be affected by the echo signal. However, due to nonideality in Digital-to-Analog Converters (DAC), some spectral leakage will be present in higher frequencies. Since the receiver's Analog-to-Digital Converter (DAC) is designed to take k time more bandwidth than the transmit signal bandwidth, all spectral leakage present within the receive bandwidth must be properly cancelled. To account for such leakage, ideally upsampled transmit signal without lowpass filtering (which would be a worst case DAC nonideality) is used to cancel the echo signal. Fortunately, this leads to a simple solution when we apply the common framework echo canceller. The time domain reference signal needed to construct $\mathcal{X}_{i,i-1}$ is generated by simply zero padding the time domain transmit signal by k times. The frequency domain reference signal needed to update the tap coefficients is generated by simply replicating the original transmit spectrum by k times.

Applying these ideas to our common framework echo canceller, we obtain the echo canceller structure shown in Fig.(3.5). In time domain, the transmit samples are upsampled by a factor of k to form the $\mathcal{X}_{i,i-1}$ matrix. $\mathcal{X}_{i,i-1}$ is now a sparse upper triangle matrix and therefore, in computing the CES output, we only need $1/k$ number of multiplies compared to the common framework.

In frequency domain, the transmit symbols $\mathbf{X}_i$ are replicated k times to form the

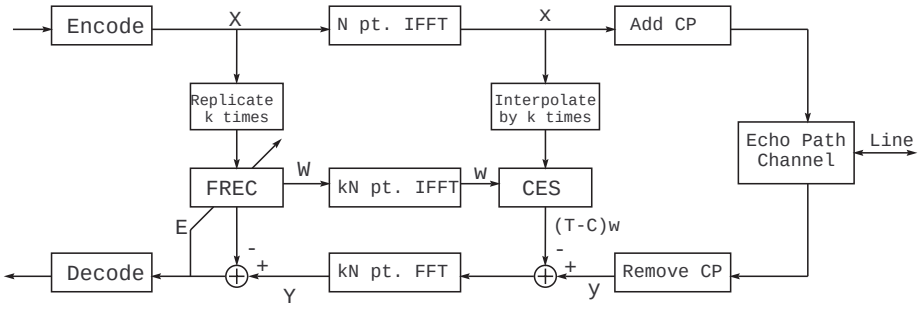

FIGURE 3.5: ATU-R Echo Canceller Block Diagram

adjusted data vector,

$$\mathbf{X}_{i,k} = \underbrace{[\mathbf{X}_i \quad \mathbf{X}_i \quad \cdots \quad \mathbf{X}_i]}_{\text{replicate k times}} \tag{3.19}$$

so that error signal $\mathbf{E}_i$ and update equation becomes,

$$\mathbf{E}_i = \mathbf{F}_{kN}(\mathbf{y}_i - \mathcal{X}_{i,i-1}\mathbf{w}_i) - diag(\mathbf{X}_{i,k})\mathbf{W}_i \tag{3.20}$$

$$\mathbf{W}_{i+1} = \mathbf{W}_i + \mu diag(\mathbf{X}^*_{i,k})\mathbf{E}_i \tag{3.21}$$

Note that the difference from the common framework echo canceller is the extra processing performed to obtain $\mathcal{X}_{i,i-1}$ and $\mathbf{X}_{i,k}$.

3.4. Initialization for ATU-R Echo Canceller

During the startup process of the modem, initialization is performed. The echo canceller's initial tap weights are found during this process and the adaptation presented thus far is performed when two way communication begins. Here, we present the ATU-R echo canceller initialization as we focus on the ATU-R echo canceller performance.

ATU-R echo canceller is initialized by transmitting up to 512 frames of repeated sequence generated by Pseudo Random Binary Sequence (PRBS) generator [16]. The

generator polynomial is given by,

$$p(D) = 1 + D^4 + D^9 \tag{3.22}$$

with initial conditions {111111111}. Every two bits of PRBS generator output is grouped to map on to 4-QAM. First bit of the pair determines the polarity of the imaginary part, and the second bit of the pair determines the polarity of the real part. This is done for all channels except the $0th$ and the $32nd$ channel which is assigned purely real polarity for Hermitian symmetry. Additionally, these frames are transmitted without the cyclic prefix and ATU-C transmitter is silenced.

Upon reception of these frames, ATU-R receiver accumulates all the received frequency domain symbols, and at the completion of up to 512 frames, accumulated sum of received frame is divided by the transmitted PRBS frame.

$$\mathbf{H}_{init} = \frac{\frac{1}{512}\sum_{i=0}^{511} Y_i}{\mathbf{X}_{PRBS}} \tag{3.23}$$

where Y_i is the ith received frequency domain frame and $\mathbf{X}_{PRBS}$ is the transmitted frame. Since cyclic prefix was not used, received frames appear circularly convolved with the echo path channel. Therefore, we may simply divide by the transmitted frequency domain frame to obtain the echo path response.

4. CONJUGATE GRADIENT METHOD

Previously proposed ADSL echo canceller as described in chapter 3 uses LMS method for adaptive update routine. It is well known that LMS method sacrifices convergence speed for its simplicity [22]. However, due to various line impairments discussed in section 3.1, it is critical that we adapt to new condition as soon as possible to avoid loss of data transmission. In an attempt to improve the ATU-R echo canceller, we apply a simplified form of Conjugate Gradient (CG) method such that the convergence speed is increased with moderate increase in the computational complexity.

CG method has been well known in linear algebra where it has been applied for solving a large system of equations [24] [37]. CG method is an iterative method in the class of gradient based techniques and it can be considered as an extension of steepest decent method (LMS). Difficulty associated with steepest decent method is in obtaining the optimal step size parameter. Various methods exist in computing the step size adaptively (e.g. Variable Step size LMS), but the computational cost is still high and some extent of heuristic bounds must be predefined [22]. More recently, there has been an increasing interest in the use of CG method for adaptive filtering applications [25] [26] [27] [28] [29]. In this thesis, we apply a simplified form of CG method tailored for use in adaptive filtering with explicit computation of optimal step size parameter such that the convergence speed is significantly increased over LMS.

First, general adaptive filtering background is given. Second, general proof of CG method is outlined. Third, proposed Modified CG method is outlined for use in ATU-R Echo Canceller and compared with LMS method. Finally, modified CG method is compared with Variable Step size LMS (VS-LMS).

4.1. Adaptive Filtering Preliminary

A general illustration of adaptive FIR filtering mechanism is shown in Fig.(4.1). The input signal to the filter and the filter coefficients are defined by,

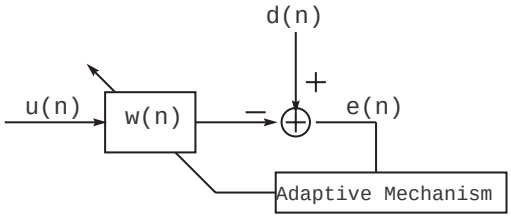

FIGURE 4.1: General Illustration of Adaptive Filtering

$$\mathbf{u}(n) = [u(n) \quad u(n-1) \quad u(n-2) \quad \cdots \quad u(n-N+1)]^T \tag{4.1}$$

$$\mathbf{w}(n) = [w_o(n) \quad w_1(n) \quad \cdots \quad w_{N-1}(n)]^T \tag{4.2}$$

where N is the size of the FIR filter, the index in the parenthesis denotes the *time* or the *iteration* the filtering is being performed, and the lower index of the filter coefficients denotes the position of the filter tap.

It is well known in adaptive filtering that by using the Mean Squared Error (MSE) as the measure of performance, the cost function can be written as [22],

$$J(w) = \sigma_d^2 - \mathbf{w}^H\mathbf{p} - \mathbf{p}^H\mathbf{w} + \mathbf{w}^H\mathbf{R}\mathbf{w} \tag{4.3}$$

where σ_d^2 is the variance of the desired signal $d(n)$, $\mathbf{w}$ represents the tap coefficients, $\mathbf{p}$ is the cross-correlation between the desired signal $d(n)$ and the input signal vector $\mathbf{u}(n)$, and $\mathbf{R}$ is the autocorrelation matrix of the input signal vector $\mathbf{u}(n)$. H denotes the Hermitian transpose of the matrix. The autocorrelation matrix $\mathbf{R}$ is almost always positive definite.

Then, we may rewrite Eq.(4.3) as [22],

$$J(\mathbf{w}) = \sigma_d^2 - \mathbf{p}^H \mathbf{R}^{-1}\mathbf{p} + (\mathbf{w} - \mathbf{R}^{-1}\mathbf{p})^H \mathbf{R}(\mathbf{w} - \mathbf{R}^{-1}\mathbf{p}) \tag{4.4}$$

and we immediately see that,

$$\min J(\mathbf{w}) = \sigma_d^2 - \mathbf{p}^H \mathbf{R}^{-1}\mathbf{p} \tag{4.5}$$

for,

$$\mathbf{w}_{opt} = \mathbf{R}^{-1}\mathbf{p} \tag{4.6}$$

This is the well known equation for Wiener filter and it is equivalent to the problem of minimizing a quadratic cost function. It can be shown that the gradient of the cost function J can be written as [22],

$$\nabla J = -2E[\mathbf{u}^*(n)e(n)] \tag{4.7}$$

where E denotes the expectation operator and $e(n)$ is the error signal given by,

$$e(n) = d(n) - \mathbf{w}^T(n)\mathbf{u}(n) \tag{4.8}$$

In the method of Steepest Decent, the update equation for the filter coefficients becomes,

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu}{2}\nabla J \tag{4.9}$$

In order to make this applicable, statistical expectation is replaced by instantaneous estimate of the gradient in LMS to obtain,

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu\mathbf{u}^*(n)e(n) \tag{4.10}$$

While LMS is very simple, it also suffers from many disadvantages. The convergence of the filter is sensitive to the choice of step size parameter μ which is bounded by [22],

$$0 < \mu < \frac{2}{\lambda_{max}} \tag{4.11}$$

where λ_{max} is the maximum eigenvalue of the input data autocorrelation matrix $\mathbf{R}$. Hence μ is typically chosen by iterative experimentation until satisfactory convergence and excess

mean square error is obtained. This may not be suitable in some applications since a very small step size μ must be chosen to ensure safe operating condition for uncertain environments.

Now we illustrate and prove the CG method applied to general problem of minimizing the cost function $f(x) = \frac{1}{2}x^T Ax - b^T x$.

4.2. Method of Conjugate Gradient : General Overview

CG method was proposed and its performance was analyzed for purely quadratic problem [24].

$$\min\{\frac{1}{2}x^T Ax - b^T x\} \tag{4.12}$$

where A is an n by n symmetric positive definite matrix and b and x are vectors of order n with x being the variable of interest and T denoting the transpose. We may find the minima of $f(x)$ by taking the gradient

$$\nabla f(x) = \begin{bmatrix} \frac{\partial}{\partial x_1} f(x) \\ \vdots \\ \frac{\partial}{\partial x_n} f(x) \end{bmatrix} \tag{4.13}$$

and setting this equal to zero. Since A is a symmetric positive definite matrix, we obtain,

$$\nabla f(x) = Ax - b = 0 \tag{4.14}$$

$$Ax = b \tag{4.15}$$

Therefore, finding the minimum solution to $f(x)$ is equivalent to solving the system of equations $Ax = b$ which requires inversion of matrix A. The fundamental principle in which the method of conjugate gradient is derived is stated by the following definition [24].

Definition 1 *Given a symmetric matrix A, two vectors d_1 and d_2 are said to be A-orthogonal, or conjugate with respect to A, if $d_1^T A d_2 = 0$ and $d_1 \neq d_2$.*

Note that matrix A does not have to be a positive definite matrix for the definition of A-orthogonality. From Definition 1, if $A = 0$, any two vectors are conjugate, while if $A = I$, I being the identity matrix, conjugacy is equivalent to standard notion of orthogonality.

Suppose there exist constants $\alpha_i, i = 0, 1, \cdots, k$ such that

$$\alpha_0 d_0 + \cdots + \alpha_k d_k = 0 \tag{4.16}$$

where all $d_i, i = 0, 1, \cdots, k$ are A-orthogonal vectors with respect to positive definite matrix A. Premultiplying by $d_i^T A$ yields,

$$\alpha_i d_i^T A d_i^T = 0 \tag{4.17}$$

Since $d_i^T A d_i^T > 0$, we have $\alpha_i = 0$. With this, we can state the following theorem [24].

Theorem 1 *If A is positive definite and the set of nonzero vectors $d_0, d_1, \cdots, d_k$ are A-orthogonal, then these vectors are linearly independent.*

From this theorem, it must be that for n by n positive definite matrix A, the solution x_{opt} to the equation $Ax_{opt} = b$ can be expanded in the following form,

$$x_{opt} = \alpha_0 d_0 + \alpha_1 d_1 + \cdots + \alpha_{n-1} d_{n-1} \tag{4.18}$$

for some constants α's. In fact, premultiplication of both sides by $d_i^T A$ yields,

$$\alpha_i = \frac{d_i^T A x_{opt}}{d_i^T A d_i} = \frac{d_i^T b}{d_i^T A d_i} \tag{4.19}$$

where the final fraction is composed of all known quantities. Therefore, we may compactly write,

$$x_{opt} = \sum_{i=0}^{n-1} \frac{d_i^T b}{d_i^T A d_i} d_i \tag{4.20}$$

We may also view Eq.(4.20) as a result of n step iterative process where at each iteration, a new scaled direction vector $\alpha_i d_i$ is added to the initial guess vector x_0 with $x_0 \in [d_0, d_1, \cdots, d_{n-1}]$. Thus, we may write,

$$x_{opt} - x_0 = \alpha_0 d_0 + \alpha_1 d_1 + \cdots + \alpha_{n-1} d_{n-1} \tag{4.21}$$

Again, we premultiply by $d_k^T A$ to get,

$$\alpha_k = \frac{d_k^T A(x_{opt} - x_0)}{d_k^T A d_k} \tag{4.22}$$

Also, for some iteration k we may write,

$$x_k - x_0 = \alpha_0 d_0 + \alpha_1 d_1 + \cdots + \alpha_{k-1} d_{k-1} \tag{4.23}$$

and premultiplication by $d_k^T A$ yields,

$$d_k^T A(x_k - x_0) = 0 \tag{4.24}$$

Expanding Eq.(4.22) and substituting Eq.(4.24) gives,

$$\alpha_k = \frac{d_k^T A(x_{opt} - x_k)}{d_k^T A d_k} \tag{4.25}$$

Observe that $Ax_{opt} - Ax_k = b - Ax_k = -\nabla f(x_k)$. Let,

$$g_k = \nabla f(x_k) = Ax_k - b \tag{4.26}$$

Finally, substituting Eq.(4.26) into Eq.(4.25) yields,

$$\alpha_k = -\frac{d_k^T g_k}{d_k^T A d_k} \tag{4.27}$$

Now, we may state the conjugate directions theorem [24].

Theorem 2 (Conjugate Direction) *Let $d_i, i = 0, 1, \cdots, n-1$ be a set of nonzero A-orthogonal vectors. For any $x_0 \in [d_0, d_1, \cdots, d_{n-1}]$, x_k is generated according to*

$$x_{k+1} = x_k + \alpha_k d_k \tag{4.28}$$

where

$$\begin{aligned} \alpha_k &= -\frac{d_k^T g_k}{d_k^T A d_k} \\ g_k &= Ax_k - b \end{aligned} \tag{4.29}$$

which converges to the optimal solution x_{opt} after n steps.

This theorem presents the underlying principle based on A-orthogonal expansion of the solution to Eq.(4.15). Note that this theorem assumes that the direction vectors, d_i's are available.

Now we turn our attention to finding the direction vectors d_i's. To do so, few more theorems need to be proved. First, substitute Eq.(4.28) into Eq.(4.26) to get,

$$g_{k+1} = g_k + \alpha_k A d_k \tag{4.30}$$

Premultiplying both sides by d_k^T and using Eq.(4.27),

$$d_k^T g_{k+1} = 0 \tag{4.31}$$

which states that current gradient (g_k) of the estimated solution x_k is orthogonal (and not A-orthogonal) to the previous iterations direction vector d_k. In general, we may state this in the following theorem [24].

Theorem 3 (Expanding Subspace) *In the method of conjugate directions, the gradients g_k for $k = 0, 1, \cdots, n$ satisfy,*

$$g_k^T d_i = 0 \qquad for \quad i < k \tag{4.32}$$

In order to find the new direction vector, we apply Gram-Schmidt orthogonalization procedure by subtracting from g_{k+1} all components in appropriately scaled d_k as follows.

$$d_{k+1} = -g_{k+1} + \beta_k d_k \tag{4.33}$$

Note that the second term on the right hand side need not be summation of all previous direction vectors as in standard Gram-Schmidt orthogonalization due to Expanding Subspace theorem. To find the coefficients β_k, we again premultiply both sides by $d_k^T A$ to get,

$$\beta_k = \frac{d_k^T A g_{k+1}}{d_k^T A d_k} \tag{4.34}$$

Unfortunately, this formula requires two full matrix multiplication in both numerator and denominator due to presence of matrix A. We may simplify this by using Eq.(4.30) in

Eq.(4.33) to get,

$$Ad_k = \frac{1}{\alpha_k}(g_{k+1} - g_k) \tag{4.35}$$

Premultiplication of both sides by g_{k+1} yields,

$$g_{k+1}^T A d_k = \frac{1}{\alpha_k} g_{k+1}^T g_{k+1} \tag{4.36}$$

Here, $g_{k+1}g_k = 0$ from Expanding Subspace Theorem. Finally substituting Eq.(4.36) and Eq.(4.27) into Eq.(4.34) yields,

$$\beta_k = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k} \tag{4.37}$$

Furthermore, the matrix multiplication in the denominator of Eq.(4.27) can be simplified by defining,

$$y_k = x_k - g_k \tag{4.38}$$

$$\begin{aligned} p_k &= \nabla f(y_k) = Ay_k - b \\ &= g_k - Ag_k \end{aligned} \tag{4.39}$$

and premultiplying Eq.(4.33) by $d_k^T A$ gives

$$Ad_k = -Ag_k \tag{4.40}$$

Thus, combining Eq.(4.39) and Eq.(4.40) into Eq.(4.27) yields,

$$\alpha_k = \frac{d_k g_k}{d_k^T (g_k - p_k)} \tag{4.41}$$

Finally, we may state the Conjugate Gradient algorithm [24].

Conjugate Gradient Algorithm Starting with any x_0 with dimension n, set $d_0 = -g_0 = b - Ax_0$, $y_0 = x_0 - g_0$, and $p_0 = Ay_0 - b$. Perform the following routine for $k = 0, 1, \cdots, n-1$.

$$\alpha_k = -\frac{d_k^T g_k}{d_k^T (g_k - p_k)} \tag{4.42}$$

$$x_{k+1} = x_k + \alpha_k d_k \tag{4.43}$$

$$\beta_k = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k} \tag{4.44}$$

$$d_{k+1} = -g_{k+1} + \beta_k d_k \tag{4.45}$$

This algorithm is efficient to solve system of equations $Ax = b$ given A and b. Due to our initial assumption that the solution can be represented by some set of linear independent direction vectors, the algorithm is shown to converge in finite steps [24].

4.3. Modified Conjugate Gradient

The general CG method is efficient in solving system of equation $\mathbf{w}_{opt} = \mathbf{R}^{-1}\mathbf{p}$ as shown in previous section. However, for real time adaptive filtering problems, there are two problems associated with it. First, the autocorrelation matrix $\mathbf{R}$ must be estimated. Second, general CG method is a double loop routine implying that for each received sample (or frame), the CG algorithm must be repeated for some finite amount of times. Both of these operations are computationally intensive and it may not be possible to perform in real time.

To simplify the operation for real time applications, the first problem is overcome by taking statistical average of the received data to estimate the autocorrelation matrix $\mathbf{R}$ in the method proposed by [25]. However, this method does not overcome the double loop structure of the algorithm. In this thesis, we turn from the general CG method to obtain Modified CG method which will overcome both problems.

4.3.1. Modified Conjugate Gradient Method

A general approach using the CG method for adaptive filtering application has been proposed by [25]. In this method, the gradient is taken by averaging over some finite window as given by,

$$\nabla J(n) = -(\frac{2}{n_w}) \sum_{j=n-n_w+1}^{n} \mathbf{u}^*(j)e(j) \tag{4.46}$$

$$= -(\frac{2}{n_w}) \sum_{j=n-n_w+1}^{n} \mathbf{u}^*(j)[d(j) - \mathbf{w}^T(n)\mathbf{u}(j)]$$

where n_w is the number of previous windows to average over. Also note that in computing $e(j)$, current estimate of the filter coefficients $\mathbf{w}(n)$ is used. This averaging is primarily to avoid using noisy measurement to estimate the conjugate directions. If at any point, the conjugate direction is obscured by noise, this estimation error tends to propagate resulting in wrong estimate.

The author noticed that by using instantaneous gradient estimate without averaging, a simplified form can be obtained. Moreover, similar to LMS, only the first gradient is used as the direction, thereby eliminating the possibility of noise propagation. Expansion of the gradient reveals,

$$-\mathbf{g}(n) = \nabla J = -2\mathbf{u}^*(n)[d(n) - \mathbf{w}^T(n)\mathbf{u}(n)] \tag{4.47}$$

and the second gradient $\mathbf{p}(n)$ can be found by,

$$\begin{aligned} \mathbf{y}(n) &= \mathbf{w}(n) - \mathbf{g}(n) \\ \mathbf{p}(n) &= \nabla f(\mathbf{y}(n)) = -2\mathbf{u}^*(n)[d(n) - \mathbf{y}^T(n)\mathbf{u}(n)] \\ &= -2\mathbf{u}^*(n)[e(n) - g(n)^T\mathbf{u}(n)] \end{aligned} \tag{4.48}$$

This method yields only one conjugate direction which is equivalent to the steepest decent. However, CG formulation allows us to compute optimal step size explicitly at every iterations. Modified CG method applied to general adaptive filtering is outlined below.

Modified Conjugate Gradient Algorithm :

1. Initialize $\mathbf{w}(0)$
2. Repeat the following computations for $n = 0, 1, \cdots$
   (a) $e(n) = d(n) - \mathbf{w}^T(n)\mathbf{u}(n)$
   (b) $\mathbf{g}(n) = 2\mathbf{u}^*(n)e(n)$
   (c) $\mathbf{p}(n) = 2(e(n) - \mathbf{g}^T(n)\mathbf{u}(n))\mathbf{u}^*(n)$

(d) $\alpha(n) = -\frac{\mathbf{g}^T(n)\mathbf{g}(n)}{\mathbf{g}^T(n)(\mathbf{g}(n)-\mathbf{p}(n))}$

(e) $\mathbf{w}(n+1) = \mathbf{w}(n) - \alpha(n)\mathbf{g}(n)$

This method is very similar to Variable Step size LMS (VS-LMS) [22]. However, VS-LMS methods proposed in the past does not overcome the design of heuristic coefficient selection. Benefits brought by the simplicity of this algorithm and its compatibility with LMS are significant. Furthermore, double loop structure of general CG method is now reduced to single loop structure making Modified CG method more suitable for real time applications.

4.3.2. Modified Conjugate Gradient Method in ADSL Echo Canceller

For our ADSL echo canceller, all adaptations are performed in the frequency domain (FREC). The illustration is repeated in Fig.(4.2). Replacing the LMS in FREC with

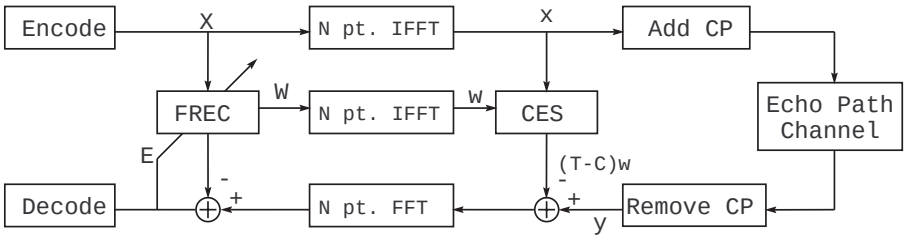

FIGURE 4.2: Common Framework ADSL Echo Canceller

Modified CG method, we obtain the following update routine.

Modified Conjugate Gradient Algorithm :

1. Initialize $\mathbf{W}(0)$

2. Repeat the following computations for $i = 0, 1, \cdots$

   (a) $\mathbf{E}_i = \mathbf{F}_N(\mathbf{y}_i - \mathcal{X}_{i,i-1}\mathbf{w}_i) - \mathbf{X}_i \cdot \mathbf{W}_i$

   (b) $\mathbf{G}_i = -2\mathbf{X}_i^* \cdot \mathbf{E}_i$

   (c) $\mathbf{P}_i = -2\mathbf{X}_i^* \cdot (\mathbf{E}_i - \mathbf{G}_i \cdot \mathbf{X}_i)$

   (d) $\alpha_i = \frac{\mathbf{G}_i^T \mathbf{G}_i}{\mathbf{G}_i^T(\mathbf{G}_i - \mathbf{P}_i)}$

   (e) $\mathbf{W}_{i+1} = \mathbf{W}_i - \alpha_i \mathbf{G}_i$

where $\cdot$ denotes element by element multiplication of two vectors. The update routine using LMS method is repeated here for illustration.

LMS Algorithm :

1. Initialize $\mathbf{W}(0)$

2. Repeat the following computations for $i = 0, 1, \cdots$

   (a) $\mathbf{E}_i = \mathbf{F}_N(\mathbf{y}_i - \mathcal{X}_{i,i-1}\mathbf{w}_i) - \mathbf{X}_i \cdot \mathbf{W}_i$

   (b) $\mathbf{W}_{i+1} = \mathbf{W}_i - \mu \mathbf{X}_i^* \mathbf{E}_i$

We can immediately see that per iteration computation is significantly larger for Modified CG method. In fact, Modified CG method requires 6N multiplications where N is the size of the tap weight coefficient vector $\mathbf{W}_i$ and the computation of $\mathbf{E}_i$ is excluded since this is a common term obtained in both methods. It is also assumed that multiplication by 2 is done by shift operation in hardware. LMS method, on the other hand, requires 2N multiplication. However, the improvement obtained with Modified CG method more than accounts for the computational complexity as we shall see in Chapter 5.

4.3.3. Comparison with VS-LMS Algorithm

LMS algorithm with an adaptive step size in an effort to increase the convergence behaviour of LMS was suggested in [38] and was further elaborated by [39]. Here, the

step size adjustment factor is found by trying to minimize the cost function with respect to the step size parameter μ. That is,

$$\nabla_{\mu}(n) = \frac{\delta J(n)}{\delta \mu} \tag{4.49}$$

This algorithm may be viewed as exciting the slow *modes* of the error signal such that the overall convergence is increased. The VS-LMS algorithm is outlined below using the notation from section 4.1.

VS-LMS Algorithm :

1. Initialize $\mathbf{w}(0)$

2. Repeat the following computations for $i = 0, 1, \cdots$

   (a) $e(n) = d(n) - \mathbf{w}^T(n)\mathbf{u}(n)$

   (b) $\mathbf{w}(n+1) = \mathbf{w}(n) + \mu(n)\mathbf{u}^*(n)e(n)$

   (c) $\mu(n+1) = (\mu(n) + \alpha\Re(\phi^T(n)\mathbf{u}^*(n)e(n)))_{\mu-}^{\mu+}$

   (d) $\phi(n+1) = (\mathbf{I} - \mu(n)\mathbf{u}(n)\mathbf{u}^T(n))\phi(n) + \mathbf{u}^*(n)e(n)$

Here, $\mu-$ and $\mu+$ are lower and upper limits of the step size at which point it must be truncated. These bounds must be selected through experimentation such that the algorithm does not become numerically unstable. Moreover, we may see that the computation required for VS-LMS algorithm is in the order of N^2 from the 4th line. Per iteration computation for VS-LMS algorithm is found to be $N^2 + 5N$ which is significantly larger than 6N for Modified CG algorithm.

4.3.4. Misadjustment

One of the common measure of the assessing the performance of adaptive filter technique is the *misadjustment* or the *excess mean square error*. Using generalized notation

from section 4.1, the error term may be expanded as,

$$\begin{aligned} e(n) &= d(n) - \mathbf{w}^T(n)\mathbf{u}(n) \\ &= d(n) - (\mathbf{w}_0^T(n) + \epsilon(n))^T\mathbf{u}(n) \\ &= e_0(n) - \epsilon^T(n)\mathbf{u}(n) \end{aligned} \tag{4.50}$$

where $e_0(n)$ is the optimal error using the Wiener solution and $\epsilon(n)$ is the tap misadjustment with respect to the Wiener solution. Using Eq.(4.50) in the cost function $J(n)$,

$$\begin{aligned} J(n) &= E[|e^2(n)|] \\ &= J_{min} + E[\epsilon^T(n)\mathbf{u}(n)\mathbf{u}^T(n)\epsilon(n)] \end{aligned} \tag{4.51}$$

where J_{min} is the *minimum mean squared error* (MMSE) produced by the optimal Wiener filter and the second term is the *excess mean square error*. In fact, for echo cancellation application, $d(n) = \mathbf{h}_{echo}^T(n)\mathbf{u}(n)$, where $\mathbf{h}_{echo}(n)$ is the sampled echo path impulse response. Then, we may rewrite Eq.(4.51) as,

$$J(n) = E[\epsilon^T(n)\mathbf{u}(n)\mathbf{u}^T(n)\epsilon(n)] \tag{4.52}$$

assuming that $\mathbf{h}_{echo}(n) = \mathbf{w}_0(n)$. Therefore, MMSE is in fact *zero* for echo cancellation problems. With the presence of background noise, we can expect that the residual echo level to fall indefinitely until the background noise level. We may further rearrange the excess mean squared error to get,

$$J_{ex}(n) = tr[\mathbf{R}\mathbf{K}(n)] \tag{4.53}$$

where tr denotes the trace of the matrix, $\mathbf{R}$ is the autocorrelation of the input data, and $\mathbf{K}(n)$ is the correlation of the tap misadjustment $\epsilon(n)$. Thus, excess mean squared error is caused by the variance of autocorrelation $\mathbf{R}$ and variance of tap misadjustment $\mathbf{K}$.

5. EXPERIMENTAL RESULTS

5.1. General Results

ATU-R echo canceller was implemented in Matlab. Two sample echo path impulse responses were generated to simulate a change in the twisted pair environment. First, initialization was performed using the procedure presented in section 3.4. During this time, far side signal is silenced and only the echo signal is present in the receiver path. Following the initialization, two way communication starts which is assumed to be a random 16-QAM signal. This is to simulate the actual *double talk* environment in which the echo canceller must adapt in real time (Fig.(5.1)). Modified CG and LMS adaptation are compared during double talk environment after the initial filter coefficients are obtained from the initialization procedure. At frame 50, the channel is changed and the adaptability of proposed method was examined. The echo impulse response and frequency domain echo response is shown in Fig.(3.3). This may represent customer adding a phone line, or pick/hang up action on the phone. The adaptation is continued until frame 200.

Care must be taken when examining the performance of an echo canceller. Usually, performance of an adaptive filter is examined experimentally by observing the MSE or the tap weight evolution. In echo canceller, or in any general problem of adaptive noise canceller, the output signal is of the desired signal and depending on the desired signal to echo signal ratio, we may not observe the adaptation behaviour at the output. This is illustrated in Fig.(5.1). To cope with such difficulty, the received signal $z(t)$ is subtracted from the final error signal to isolate the residual echo signal.

Fig.(5.2) shows the received signal spectrum from the far end transmitter, echo signal spectrum, and the ratio of the received signal spectrum and the echo signal spectrum. We may readily see that the echo signal is in fact buried in the received signal. Hence, we do not expect to be able to observe the behaviour of the adaptation by simply looking at the echo canceller output. Fig.(5.3) shows the residual echo signal at the output of

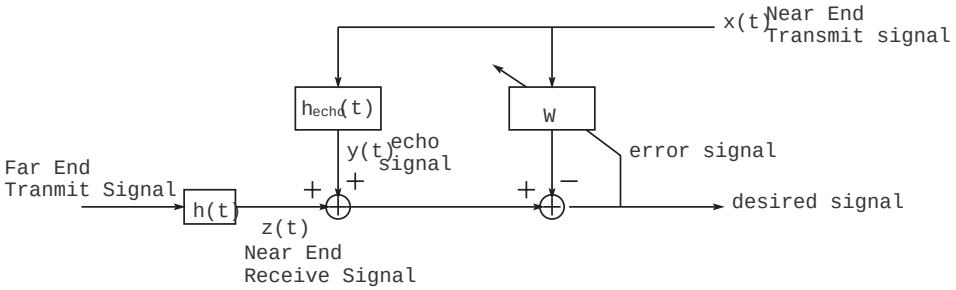

FIGURE 5.1: General Block Diagram of Echo Cancellation Problem

the ATU-R echo canceller for LMS and Modified CG. We may observe the following from Fig.(5.3).

- The first 50 frames remains constant since the echo canceller taps were initialized to the optimal value.
- At frame 50, we see a sudden jump in the residual echo signal due to change in the environment.
- After frame 50, we see the convergence behaviour of the LMS method and Modified CG method.

The Modified CG method converges at approximately 70th frame (or after 20 iterations) where as LMS converges at 150th frame (or after 100 iterations). The step size μ used for LMS was 2^{-6} as suggested in [6]. As described in Chapter 4.3.2, LMS requires 2N multiplications and Modified CG requires 6N multiplications where N is the size of the filter coefficient. Based on this, the required number of multiplications until the convergence is approximately 200N for LMS and 120N for Modified CG. Therefore, we see that nearly half the amount of multiplication is required with 5 times increase in the convergence speed. The results are tabulized in Table.(5.1).

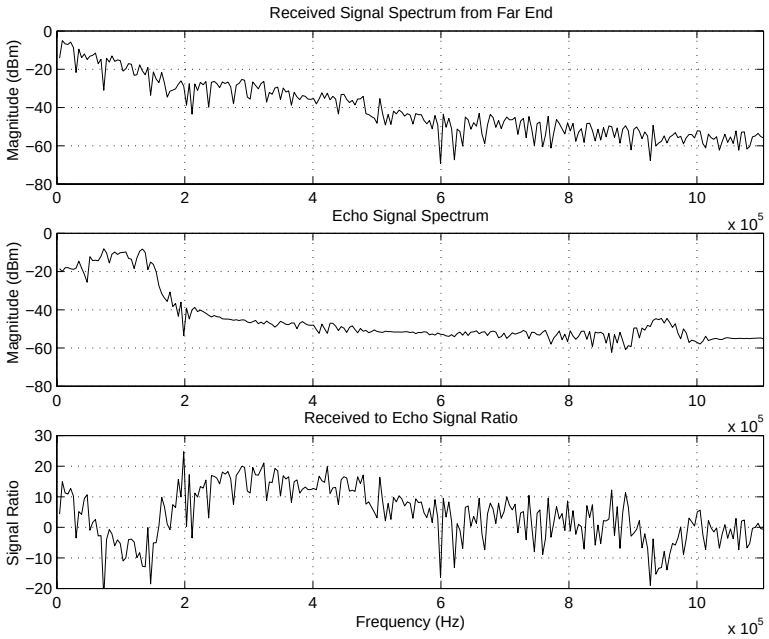

FIGURE 5.2: Received signal and Echo Signal Spectrum

5.2. Fixed Point Simulation Results

For the purpose of product implementation, performance of the algorithm with fixed point effects must be evaluated. High precision computations are expensive in embedded DSP systems, and usually cost performance trade-off must be made for actual implementation. Additionally, certain types of algorithm may be more prone to numerical instability when number of bits are reduced. Here, we examined the performance of both LMS and Modified CG with finite number of bits. Computed tap coefficients are multiplied by $2^{(number\ of\ bits)}$, rounded to nearest integer, and then divided by $2^{(number\ of\ bits)}$ to obtain the fixed point effects. In reality, addition and multiplication algorithms used may also play a role, but this extends beyond our scope of study. Therefore the multiplication and addition was assumed to be performed by the precision as allowed by Matlab, and the coefficient representation was assumed to be fixed point.

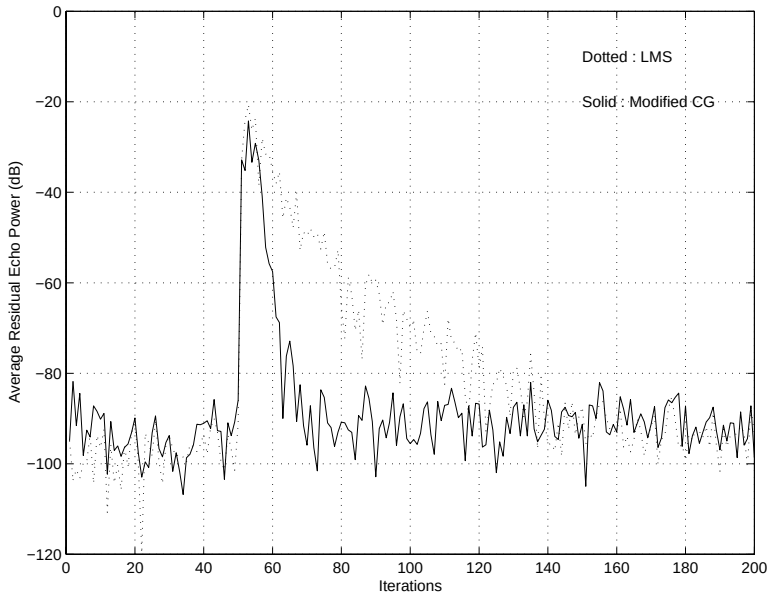

FIGURE 5.3: Average Residual Echo Signal Power

Fig.(5.4) shows the experimental results using fixed point effect described above. The figure illustrates the MSE for two algorithms when 16 bits are used. We see that with 16-bit precision, similar results are obtained as shown in Fig.(5.3). The consequence of reducing the number of bits is increase in the final misadjustment. The top figure in Fig.(5.4) shows that the achievable echo rejection, or residual echo power has decreased to approximately -65dB. However, the convergence behaviour of Modified CG remains the same and that numerical instability is not observed. This shows that Modified CG performs equally or better than LMS with 3 times computational complexity per iteration. Furthermore, Modified CG's compatibility with LMS, which may be done by fixing the step size parameter α, gives us flexibility if LMS like behaviour is desired.

Method	Number of Multiplications per Iteration	Number of Iterations until convergence	Total number of Multiplication until Convergence
LMS	2N	100	200N
Modified CG	6N	20	120N

TABLE 5.1: Comparison Table

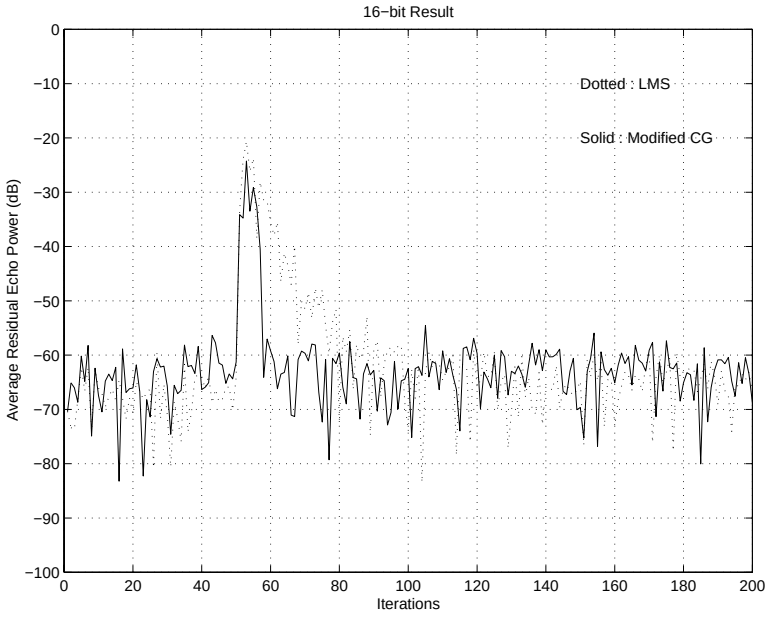

FIGURE 5.4: Average Residual Echo Signal Power with Fixed Point

6. CONCLUSION

6.1. Summary of Results

In this thesis, basics of Asymmetric Digital Subscriber Line (ADSL) was discussed. The problem of echo cancellation in general and details specific to ADSL echo cancellation was discussed based on previously proposed method in [6]. The author noticed that in certain operating environments, Frequency domain Echo Cancellation (FREC) using Least Mean Square (LMS) method may not be sufficient for practical use.

In an effort to improve the performance of the ADSL FREC, a new algorithm called Modified Conjugate Gradient (CG) method was proposed. This method is derived from previously proposed CG method applied to adaptive filtering and further simplified to obtain an implementable solution with increased performance compared to LMS method. Analysis shows that per iteration computation required for Modified CG method is 6N while LMS method requires 2N. It was also shown that the Modified CG method can also be interpreted as a Variable Step size LMS (VS-LMS) using different criteria. Per iteration computation required for VS-LMS was found to be $N^2 + 5N$. Furthermore, VS-LMS requires bounds on step size to be defined experimentally so that the algorithm does not become unstable. On the contrary, all quantities in Modified CG method are explicitly computed for optimal performance.

Matlab simulation was performed for ATU-R echo canceller using both LMS method and Modified CG method. Modified CG method was shown to provide 5 times increase in convergence speed while trading 3 times increase in the computational complexity. Overall, the total computations required until convergence is reduced by 40% over conventional LMS method. It was also shown that with fixed point effects, the algorithm is numerically stable and provides consistent results.

6.2. Further Research

Due to nonlinear operations involved in computation of CG method in general, it is a very difficult task to prove the exact convergence and stochastic behaviour of the algorithm. It is yet to be analyzed exactly how CG method behaves in adaptive filtering environment.

In echo cancellation, problem of double talk has significant importance. We may halt the two way communication to initialize the tap coefficients of the echo canceller. However, for more efficient operation, we would like to have the echo canceller adapt to changes in double talk environment. In general, the performance of adaptive algorithm in echo canceller during double talk is dependent on the far end signal's statistics. In many cases, the far side signal may be colored and partially correlated to the near side signal. Under such conditions, the adaptive algorithm of the echo canceller may exhibit erroneous behaviour, but the exact consequences are not well known. It would be interesting to see the exact statistical dependence of an adaptive algorithm to the far side signal for more robust design.

BIBLIOGRAPHY

1. G.T.Hawley, "System Considerations for the Use of xDSL Technology for Data Access," IEEE Communications Mag., Vol.35, No.3, pp 56-60, March, 1997.

2. J.M.Cioffi, "A Multicarrier Primer," T1E1.4 91-157.

3. J.M.Cioffi and J.A.C.Bingham, "A Data-Driven Multitone Echo Canceller," IEEE GLOBECOM, Vol 2, pp 57-61, 1991.

4. J.M.Cioffi and J.A.C.Bingham, "A Data-Driven Multitone Echo Canceller," IEEE Trans. on Comm., Vol 42, No.10, pp 2853-2869, Oct. 1994.

5. M.Ho, J.M.Cioffi, and J.A.C.Bingham, "High-Speed Full-Duplex Echo Cancellation for Discrete Multitone Modulation," Intl. Conf. Comm., pp 772-776, 1993.

6. M.Ho, J.M.Cioffi, J.A.C.Bingham, "Discrete Multitone Echo Cancellation," IEEE Trans. on Comm., Vol 44, No.7, pp 817-825, July, 1996.

7. R.C.Younce, P.J.W.Melsa, and S.Kapoor, "Echo Cancellation for Asymmetrical Digital Subscriber Lines," Intl. Conf. Comm., pp 301-306, 1994.

8. D.C.Jones, "Frequency Domain Echo Cancellation for Discrete Multitone Asymmetric Digital Subscriber Line Transceivers," IEEE Trans. on Comm., Vol 43, No.2/3/4, pp 1663-1672, Feb/Mar/Apr, 1995.

9. B.P.Lathi, "Modern Digital and Analog Communication Systems," Oxford University Press, New York, 2nd Ed., 1995.

10. J.S.Chow, "Finite-Length Equalization for Multi-Carrier Transmission Systems," Ph.D Thesis, Standford University, June, 1992.

11. J.S.Chow and M.Ho, "Modified Glissando Sequence for DMT Echo Canceller Initialization," T1E1.4 Standards Contribution, 93-181, Aug., 1993.

12. S.B.Wicker, "Error Control Systems for Digital Communication and Storage," Prentice Hall, New Jersey, 1995.

13. L.Wei, "Trellis-Coded Modulation with Multidimensional Constellations," IEEE Trans. on Information Theory, Vol IT-33, No.4, pp 483-501, July 1987.

14. A.Peled and A.Ruiz, "Frequency Domain Data Transmission using Reduced Computational Complexity Algorithms," IEEE ICASSP, Denver, pp 964-967, April, 1980.

15. A.Ruiz, J.M.Cioffi, and S.Kasturia, "Discrete Multiple Tone Modulation with Coset Coding for the Spectrally Shaped Channel," IEEE Trans. on Comm., Vol 40, No.6, pp 1012-1029, June 1992.

16. J.Bingham and F.Van Der Putten, "T1.413 Issue 2," T1E1.4, 97-007R6, September 26, 1997.

17. K.Murano, S.Unagami, and F.Amano, "Echo Cancellation and Applications," IEEE Communications Mag., pp 49-55, January 1990.

18. M.R.Asharif and F.Amano, "Acoustic Echo-Canceler Using the FBAF Algorithm," IEEE Trans. on Comm., Vol 42, No.12, pp 3090-3094, Dec., 1994

19. M.M.Sondhi and D.A.Berkley, "Silencing the Echoes on the Telephone Network," Proceedings of the IEEE, Vol 68, No.8, pp 948-963, Aug., 1980.

20. S.B.Weinstein, "A Passband Data-Driven Echo Canceller for Full-Duplex Transmission on Two-Wire Circuits," IEEE Trans. on Comm., Vol Com-25, No.7, pp 654-666, Jul., 1977.

21. A.S.Munshi, D.A.Johns, and A.S.Sedra, "Adaptive Impedance Matching," IEEE Intl. Symposium on CAS, Vol 2, pp 69-72, 1994.

22. S.Haykin, "Adaptive Filter Theory," Prentice Hall, New Jersey, 3rd Ed., 1996.

23. J.J.Shynk, "Frequency-Domain and Multirate Adaptive Filtering," IEEE SP Mag., pp 14-32, Jan., 1992

24. D.G.Luenberger, "Linear and Nonlinear Programming," Addison-Wesley, MA, 1984.

25. G.K.Boray and M.D.Srinath, "Conjugate Gradient Technique for Adaptive Filtering," IEEE Trans. on CAS-I, Vol 39, No.1, pp 1-10, Jan., 1992.

26. Z.Fu and E.M.Dowling, "Conjugate Gradient Eigenstructure Tracking for Adaptive Spectral Estimation," IEEE TRans. on SP, Vol 43, No.5, pp 1151-1160, May, 1995.

27. J.S.Lim and C.K.Un, "Block Conjugate Gradient Algorithms for Adaptive Filtering," Signal Processing, Vol 55, pp 65-77, 1996.

28. P.S.Chang and A.N.Willson, Jr, "Adaptive Spectral Estimation Using the Conjugate Gradient Algorithm," IEEE ICASSP, Vol 5, pp 2979-2982, 1996.

29. R.A.Soni and W.K.Jenkins, "Channel Equalization with Adaptive Filtering and the Pre-Conditioned Conjugate Gradient Algorithm," IEEE Intl. Symp. on CAS, Vol 4, pp 2284-2287, 1997.

30. A.V.Oppenheim and R.W.Schafer, "Discrete-Time Signal Processing," Prentice Hall, New Jersey, 1989.

31. I.Kalet, "The Multitone Channel," IEEE Trans. on Comm., Vol 37, No.2, pp 119-124, Feb., 1989.

32. J.G.Proakis and D.G.Manolakis, "Digital Signal Processing : Principles, Algorithms, and Applications," Prentice Hall, New Jersey, 3rd Ed., 1996.

33. J.G.Proakis, "Digital Communications," McGraw-Hill, New York, 3rd Ed., 1995.

34. B.Hirosaki, S.Hasegawa, and A.Sabato, "Advanced Groupband Data Modem Using Orthogonally Multiplexed QAM Technique," IEEE Trans. on Comm., Vol COM-34, No.6, pp 587-592, June, 1986.

35. B.R.Saltzberg, "Performance of an Efficient Parallel Data Transmission System," IEEE Trans. on Comm. Tech., Vol COM-15, No.6, pp 805-811, Dec., 1967.

36. S.B.Weinstein and P.M.Ebert, "Data Transmission by Frequency-Division Multiplexing Using the Discrete Fourier Transform," IEEE Trans. on Comm., Vol COM-19, No.5, pp 628-634, Oct., 1971.

37. G.H.Golub and C.F.Van Loan, "Matrix Computations," The Johns Hopkins University Press, London, 3rd Ed., 1996.

38. A.Benveniste, M.Metivier, and P.Priouret, "Adaptive Algorithms and Stochastic Approximation," Springer-Verlag, New York, 1990

39. H.J.Kushner and J.Yang, "Analysis of adaptive Step Size SA Algorithms for parameter tracking," IEEE Trans. on Auto. Control, Vol 40, pp 1403-1410, 1995.