Optimal Time Domain Equalization Design for Maximizing Data Rate of Discrete Multi-Tone Systems

Ph. D. Qualifying Proposal

Telecommunications and Information Systems Engineering Area

Department of Electrical and Computer Engineering

The University of Texas at Austin

Submitted by **Milos Milosevic, B.S.E.E., M.S.E.E.**

E-mail: milos@ece.utexas.edu

Advisor:

Dr. Brian L. Evans

Committee members:

Dr. Ross Baldick (Dept. of ECE)

Dr. Gustavo de Veciana (Dept. of ECE)

Dr. Edward J. Powers (Dept. of ECE)

Dr. Robert A. van de Geijn (Dept. of CS)

April 18, 2002

Contents

List of Figures

List of Tables

Summary

The traditional discrete multi-tone equalizer is a cascade of a time domain equalizer (TEQ) as a single finite impulse response (FIR) filter, a fast Fourier transform (FFT) multicarrier demodulator, and a frequency domain equalizer (FEQ) as a one-tap filter bank. The TEQ shortens the transmission channel impulse response (CIR) to mitigate inter-symbol interference (ISI). Maximum Bit Rate (MBR) and Minimum ISI (Min-ISI) methods achieve higher data rates at the TEQ output than previously published methods. As an alternative to the traditional equalizer, the per-tone equalizer (PTE) moves the TEQ into the FEQ and customizes a multi-tap FEQ for each tone.

In this report, we propose a *time domain* TEQ filter bank (TEQFB) and single TEQ designs that demonstrate better data rates *at the FEQ output* than MBR, Min-ISI, and least-squares PTE methods with standard CIRs, transmit filters, and receive filters. We propose a new achievable upper bound on bit rate based the performance of the optimal TEQFB. We propose an iterative fractional programming algorithm, which can produce a single FIR TEQ that achieves on average more than 99% of the performance of the optimal TEQFB for the standard carrier serving area loops 1-8, and standard transmit and receive filters. We develop a model for the signal-to-noise ratio (SNR) based on the desire to obtain a circularly convolved data frame and CIR at the FFT input, while including near-end crosstalk (NEXT), white Gaussian noise, and the digital noise floor (DNF). DNF was not included in prior work although it can represent a significant source of noise in cases when the signal is highly attenuated. NEXT was also not included in prior work on TEQ design although it is present in practical DMT systems as high interference from other communication lines in close proximity. The proposed SNR model and subsequent TEQ design take these noise sources into

account and help mitigate their effect on the bit rate.

The expected contributions of the dissertation research are:

1. a new model for the SNR (*completed work*, submitted in [1]),

2. data rate optimal time domain per-tone TEQ filter bank (TEQFB) algorithm (*completed work*, submitted in [1]),
   (a) proof of optimality of TEQFB (*completed work*, submitted in [1]),
   (b) a new achievable upper bound on bit rate performance (*completed work*, submitted in [1]),
   (c) computationally less intensive algorithm for TEQFB then obtained straight from the equations (*current work*),
   (d) release of the TEQFB design software as a part of a new release of the MATLAB DMT TEQ Design Toolbox [2] (*future work*),

3. data rate maximization single TEQ design algorithm (*completed work*, submitted in [1])
   (a) computationally less intensive algorithm for single TEQ then obtained straight from the equations (*current work*),
   (b) release of the software performing single TEQ design as a part of a new release of the MATLAB DMT TEQ Design Toolbox [2] (*future work*),
   (c) data rate maximization time domain TEQ filter bank optimized for groups of tones with the expectation of performance similar to TEQFB but at lower cost (*future work*), and
   (d) depending on the results from item (3c), add item (3d) to a new release of the MATLAB DMT TEQ Design Toolbox (*future work*).

Proposed milestones to complete research:

1. July 2002 - (2c),(3a) completed,

2. Aug 2002 - (2d),(3b) completed,

3. Oct 2002 - (3c) completed, and

4. Nov 2002 - (3d) completed if deemed valuable.

1 Introduction

Discrete multi-tone (DMT) is a multicarrier modulation method in which the available bandwidth of a communication channel, such as twisted-pair copper media, is divided into numerous subchannels or bins via a fast Fourier transform (FFT). Data simultaneously flows downstream from a central office to a remote terminal, and upstream in the opposite direction. The DMT technique has been adopted in the US by the American National Standards Institute (ANSI) T1.413-1998 standard (Asymmetric Digital Subscriber Loop ADSL standard) [3], and internationally by the International Telecommunications Union G.DMT (G.992.1) [4] and G.Lite (G.992.2) [5] ADSL standards. DMT is figuring in the very-high-rate DSL (VDSL) standard proposals [6, 7, 8]. In the ANSI ADSL standard, DMT is used to generate up to 250 separate 4.3125 kHz wide downstream subchannels from 26 kHz to 1.1 MHz. Likewise, DMT is used to generate 26 upstream subchannels from 26 kHz to 138 kHz. Modulation by the inverse fast Fourier transform (IFFT) and demodulation by the FFT create nearly orthogonal subchannels. Each subchannel is nearly independent of the other subchannels and the degree of independence increases with the number of subchannels [9].

Each subchannel can support a specific number of bits given the power level of the signal, the desired bit error rate (BER) and the total power of the noise in that subchannel. The total number of bits transmitted is the sum over the bits transmitted in each subchannel. A spectrally shaped channel impulse response (CIR), i.e. one that has a duration longer than a single sample, removes the orthogonality between subchannels so that they cannot be fully separated at the receiver. It causes both inter-carrier interference (ICI) and inter-symbol interference (ISI).

If the length L of the CIR is less than or equal to $\nu + 1$, then adding a guard period of ν

samples at the beginning of a DMT symbol will prevent the occurrence of ISI. If we choose the guard period to be a copy of the last ν samples of a DMT symbol, then we eliminate the ICI as well. In the latter case, the guard period is called the cyclic prefix (CP) and represents the solution adopted in the ADSL standards and proposed for VDSL. If a DMT symbol contains N samples, then the guard period ν samples lowers the data rate by a factor of $N/(N+\nu)$. For ADSL downstream transmission, $N = 512$ and $\nu = 32$ samples, whereas for ADSL upstream transmission, $N = 64$ and $\nu = 4$ samples. In VDSL, N is up to 8192 and ν is the same fraction ($\frac{1}{16}$) of N as in ADSL.

The equalization in DSL is based on a two-step process. In the first equalization step, a finite impulse response filter (FIR) block, called a time domain equalizer (TEQ), is tasked with eliminating the ISI and ICI from the received DMT signal. The convolution of the TEQ and CIR ideally results in a transmission channel whose shortened impulse response (SIR) has an extent that is smaller than or equal to $\nu + 1$ samples. Linear convolution of the SIR and a DMT symbol is converted into their circular convolution by the virtue of having repeated DMT symbol samples present in the CP [3]. Demodulation, which is performed using an FFT, then transforms this circular convolution into a multiplication of complex sequences in the frequency domain [10]. In the second equalization step, division by the frequency domain response of the SIR (known as frequency domain equalization or FEQ), fully equalizes the signal. This fully equalized signal can then be processed further for the proper decoding of the transmitted message.

Many TEQ design methods have been published based on the traditional receiver architecture of a single FIR TEQ followed by an FFT and a single-tap FEQ. Minimum mean-square error (MMSE) is one of the earliest, widely used, and mathematically tractable solutions [11]. The

equalizer taps are designed to minimize the mean-square of the residual error between the SIR and a target impulse response (TIR) of length $\nu + 1$ samples. The TEQ taps obtained by using this approach do not necessarily increase the data rate because minimizing MSE is not tightly coupled with increasing data rate. MMSE solutions tend to put deep nulls[1] in the SIR, which in the presence of a digital noise floor (DNF) renders some of the subchannels unable to carry data [12, 13].

The Maximum Shortening SNR (MSSNR) method [14] attempts to maximize the ratio of the energy present in the target SIR window of length ν samples to the energy outside of the window. The method is based on an observation that reducing the energy of SIR outside the target window will reduce ISI and ICI. While this is true, the method does not maximize bit rate directly, and SSNR may not be the best tool for TEQ design as concluded by its authors [14]. A recent paper by Daly *et al.* [15] showed that the MSSNR design is equivalent to the widely used Minimum Mean-Squared Error MMSE method.

Maximum Bit Rate (MBR) and Minimum ISI (Min-ISI) methods [12] are based on a similar idea as [14], which is used to define a new model for the signal-to-noise ratio (SNR). This SNR model includes ISI and additive white Gaussian noise (AWGN) and is used to determine the bit rate. The SNR is defined in the frequency domain for each subchannel. MBR tries to maximize the nonlinear bit rate equation through the use of the newly defined subchannel SNR via the Broyden-Fletcher-Goldfarb-Shanno quasi-Newton algorithm in Matlab's optimization toolbox. The authors in [12] conclude that the MBR procedure is computationally expensive and therefore not well suited for real-time implementation on a programmable digital signal processor (DSP). Nevertheless, the procedure does maximize the

[1]Null in this report refers to a region of high attenuation in the spectrum of the signal

bit rate for the traditional receiver architecture at the TEQ output as it achieves the matched filter bound. The authors in [12] also define the Min-ISI method, which minimizes the sum of subchannel noise power present in the new SNR definition with the constraint that the sum of subchannel signal power is constant. In simulation, the Min-ISI method achieved more than 95% of the bit rate of the MBR method [12]. The Min-ISI method has been implemented on several fixed-point programmable DSPs [12].

An alternate receiver architecture is proposed in [16]. Since the traditional equalizer equalizes all subchannels "in a combined fashion," which limits equalization performance, the authors of [16] propose to transfer the TEQ operations to the frequency domain by moving the TEQ into the FEQ. The combined TEQ-FEQ would yield a multi-tap FEQ structure in which each subchannel (tone) could be separately equalized. This per-tone equalizer (PTE) could be implemented as a vector dot product of the sliding FFT coefficients for that subchannel and the vector of complex-valued FEQ coefficients. Hence, the single-tap FEQ, which multiplies each subchannel FFT coefficient by the inverse of the equalized channel frequency response, is replaced by a vector dot product for each subchannel. In [16], the authors also propose various groupings of subchannels, in which each group has a complex equalizing filter assigned to it. By grouping subchannels, they trade achievable bit rate for reduced implementation complexity.

In this report, we develop a new SNR model based on desire to obtain a circularly convolved data frame and the CIR at the input of the demodulating FFT block while including the effects of near-end crosstalk (NEXT), AWGN and the DNF. We arrive at the optimal time domain per tone TEQ filter bank (TEQFB) by employing Goertzel filters [17, 18] at the receiver during data transmission. We propose an iterative fractional programming algorithm that produces a single TEQ FIR that achieves on average more than 99% of the performance

of TEQFB for the CIR loops containing standard carrier serving area loops 1-8 and high pass transmit and receive filters. We propose a method of assessing performance of TEQ based on the performance of TEQFB. We show that both TEQFB and TEQ outperform MBR, Min-ISI and least-squares (LS) PTE designs. The contributions of this paper are: (1) a new subchannel SNR model, (2) data rate optimal per-tone TEQ filter bank, and (3) data rate maximization single TEQ design algorithm. In the report, we refer to the proposed optimal per tone TEQ filter bank as "TEQFB" and to a single TEQ as "TEQ".

This report is organized as follows. Section 2 summarizes background information on the achievable data rate of a DMT system. Section 3 presents the proposed SNR model. Section 4 derives the proposed method for design of the optimum subchannel TEQ and a proposed method for the near-optimum design of a TEQ. Both of the proposed methods are compliant with the various ADSL standards. Section 5 presents simulation results from the work already completed. Section 6 draws conclusions from the completed work. Section 7 outlines future research work.

2 Achievable data rate in DMT systems

The achievable rate of a white Gaussian transmission channel is given by its capacity in bits per real dimension per transmission as [19]

$$b_{\mathrm{G}} = \frac{1}{2}\log_2\left(1 + \frac{P_{\mathrm{s}}}{P_{\mathrm{n}}}\right) \text{ (bits/s/Hz)} \tag{2.1}$$

Here, b_{G} is the number of bits per transmission, P_s is the signal power and P_n is the noise power. We can define SNR $= P_{\mathrm{s}}/P_{\mathrm{n}}$. Practical coding/modulation methods cannot achieve the rate given in (2.1). The difference between the rate in (2.1) and the best achievable rate in practice can be characterized by a quantity called the SNR gap (Γ) in (dB) [9, 20]. Γ is a function of the modulation method and the target probability of bit error per dimension, P_{e}. In DMT, data is modulated in the complex (two-dimensional) plane and every subchannel can have a different SNR gap Γ_k. We will assume that the target probability of error in all subchannels is the same. Thus, we can set $\Gamma_k = \Gamma \ \forall k$. For a coded quadrature amplitude modulation system (QAM),

$$\Gamma = 9.8 + \gamma_{\mathrm{m}} - \gamma_{\mathrm{c}} \text{ (dB)} \tag{2.2}$$

where γ_{m} is the desired system margin, and γ_{c} is the gain (efficiency) of the coding method. In G.DMT ADSL, typically, $P_e = 10^{-7}$, $\gamma_{\mathrm{m}} = 6$ dB, and $\gamma_{\mathrm{c}} \approx 4.2$ dB; hence, $\Gamma \approx 11.6$ dB.

A DMT system has $N/2 - 1$ subchannels, where N is the IFFT size. When N is large, the subchannels can be considered independent in the presence of Gaussian noise [9]. The data rate in bits per symbol in the k^{th} subchannel becomes

$$b_k = \log_2\left(1 + \frac{\mathrm{SNR}_k}{\Gamma}\right) \tag{2.3}$$

A DMT system has $N/2-1$ subchannels but only a portion of those carry data. Some of the subchannels may not be of interest; e.g., in ADSL, subchannels 0-6 are reserved for voice service and Integrated Services Digital Network (ISDN) compatibility while subchannel 64 is reserved for the pilot tone used for synchronization [3, 4, 5]. Accordingly, we define a set of subchannels of interest, $\mathcal{I}$, such that $\mathcal{I} \subseteq \{0,\ 1, \cdots,\ N/2-1\}$. Then, the number of bits per DMT symbol is given by

$$b_{\mathrm{DMT}}(\mathcal{I}) = \sum_{k \in \mathcal{I}} \log_2 \left(1 + \frac{\mathrm{SNR}_k}{\Gamma}\right) \tag{2.4}$$

Equation (2.4) could result in non-integer bit values, but ADSL and VDSL standards allow only integer bit loading on subchannels. Non-integer number of bits could be loaded if constellations of dimensionality higher than two are considered. This situation arises when Trellis coding is used [21]. Thus, the number of bits that would be actually used in a DMT system is

$$b^{\mathrm{int}}_{\mathrm{DMT}}(\mathcal{I}) = \sum_{k \in \mathcal{I}} \left\lfloor \log_2 \left(1 + \frac{\mathrm{SNR}_k}{\Gamma}\right) \right\rfloor \tag{2.5}$$

where $\lfloor . \rfloor$ means the closest smaller integer. The main objective of this paper is to model (2.5) as a function of TEQ filter coefficients and then design an efficient optimization method that can maximize data rate on DMT systems. The flooring operation makes (2.5) mathematically difficult to handle, and we would like to deal instead with (2.4). In general, the maximizer of (2.4) may not maximize (2.5). We will explore this issue in Section 3.1 and give evidence as to why in our case, it is reasonable to assume that the maximizers will be the same. Nonetheless, both functions depend on the SNR_k quantities, as the value of Γ is fixed for the particular modulation method. Both (2.4) and (2.5) are increasing functions of SNR_k. In the next section, we develop a new model of the $\mathrm{SNR_k}$ dependent on the TEQ taps.

3 Proposed SNR model

In the following derivation, only major ideas are discussed and specific details and definitions are presented in the Appendix A.

Let $\mathbf{u}_i$ be the i^{th} $N \times 1$ sample DMT symbol to be decoded. Preceding and following this symbol are $\mathbf{u}_{i-1}$ and $\mathbf{u}_{i+1}$ DMT symbols. Let ν be the length of the CP and $\mathbf{h} = [h_0, h_1, \cdots, h_{N-1}]^{\mathrm{T}}$ be $N \times 1$ transmission CIR. Let $\mathbf{w} = [w_0, w_1, \cdots, w_{M-1}]^{\mathrm{T}}$ be an $M \times 1$ TEQ, where M is some predetermined length. Let Δ be the transmission delay incurred by the signal from the transmitter to the FFT block and let

$$\mathbf{U}_{\mathrm{ISI}}^{\Delta} = \mathbf{U}_{i-1}^{\Delta} + \mathbf{U}_{i}^{\Delta} + \mathbf{U}_{i+1}^{\Delta} \tag{3.1}$$

be the convolutional matrix of the DMT symbol i including contributions of symbols $i-1$ and $i+1$. Let $\mathbf{H}$ be the convolution matrix of the CIR and TEQ (A.2). We define the vector

$$\mathbf{q}_k = \left[1,\ e^{j2\pi k/N}, \cdots,\ e^{j2\pi(N-1)k/N}\right]^{\mathrm{T}} \tag{3.2}$$

such that the inner product of $\mathbf{q}_k^{\mathrm{H}}$ with an N-point vector gives the k^{th} FFT coefficient of that vector, where $(\cdot)^{\mathrm{H}}$ means the Hermitian (conjugate transpose) operator.

DMT time domain samples are a superposition of a number of sinusoids (up to 256 in ADSL and up to 4096 in VDSL) with "random" amplitudes and phase shifts of varying frequencies. DMT time domain samples can therefore be approximated using the Central Limit Theorem [22] as being independent, identically distributed (i.i.d.) according to $\mathcal{N}(0, \sigma_s^2)$, where $\mathcal{N}$ represents normal distribution and σ_s^2 is the transmit signal variance, which is measured by the

transmit power with respect to 100Ω resistance.

Let $\mathbf{G}_{\text{AWGN}}$ and $\mathbf{G}_{\text{NEXT}}$ be the AWGN and NEXT convolution matrices defined in A.3, respectively, with the TEQ impulse response. The samples of AWGN are i.i.d. according to $\mathcal{N}(0, \sigma^2_{\text{AWGN}})$ where σ^2_{AWGN} is the noise variance, which is measured by the power of the noise with respect to 100Ω resistance. The samples of the NEXT have spectrally shaped power spectral density (PSD) as defined in [3] for various types of interferers and denoted using the Toeplitz variance matrix $\mathbf{\Sigma}_{\text{NEXT}}$.

Define D_k as the DNF term in the subchannel k due to finite precision in digital processing, which is independent of the TEQ but affects the subchannel SNR terms, $\text{SNR}_k(\mathbf{w})$, which are a function of the TEQ impulse response $\mathbf{w}$ (FIR coefficients). We assume that the samples of DNF are i.i.d. according to $\mathcal{U}(0, \sigma^2_{\text{DNF}})$, where $\mathcal{U}$ stands for uniform distribution and $\sigma^2_{\text{DNF}} = 2^{-2(p+1)}/3$, where p is the number of bits of precision of digital processing (e.g. the word length of the A/D converter). This noise can become a dominant source of noise in a subchannel where the signal has been attenuated severely.

We write the received data point in the k^{th} subchannel as

$$Y^k_{\text{R}}(\mathbf{w}) = \mathbf{q}^{\text{H}}_k \mathbf{U}^{\Delta}_{\text{ISI}} \mathbf{H}\mathbf{w} + \mathbf{q}^{\text{H}}_k \mathbf{G}_{\text{AWGN}}\mathbf{w} + \mathbf{q}^{\text{H}}_k \mathbf{G}_{\text{NEXT}}\mathbf{w} + D_k \tag{3.3}$$

where $k \in \{0, \cdots, \frac{N}{2} - 1\}$. The received data contains the noise due to the ISI, ICI, AWGN, NEXT and DNF, and suffers from the effects of the channel. Now, we see the dependence of the received signal on the TEQ.

Next, we will express the desired signal as a function of the TEQ taps. The ideal received signal has no noise present and is "formatted" to fit the demodulation scheme. In the case of

DMT, this means that the received symbol has minimal noise present due to AWGN, NEXT, and ISI and that the strength of the signal is much higher than the DNF.

Ideally, after the TEQ, one would prefer to have mimicked a circular convolution of the signal and the CIR using the CP. The desirable circular convolution of the i^{th} symbol and the CIR in the k^{th} subchannel, after the TEQ and FFT, can be written as

$$Y_{\text{D}}^{k}(\mathbf{w}) = \mathbf{q}_{k}^{\text{H}} \left[\mathbf{U}_{i}^{\Delta}\right]_{\text{circ}} \mathbf{H}\mathbf{w},\ k \in \{0, \cdots, N/2-1\} \tag{3.4}$$

So, now the received data $Y_{\text{R}}^{k}(\mathbf{w})$ can be rewritten as

$$Y_{\text{R}}^{k}(\mathbf{w}) = Y_{\text{D}}^{k}(\mathbf{w}) + \underbrace{\left(Y_{\text{R}}^{k}(\mathbf{w}) - Y_{\text{D}}^{k}(\mathbf{w})\right)}_{\text{ISI, ICI, AWGN, NEXT, DNF}} \tag{3.5}$$

We rewrite $\text{SNR}_{k}(\mathbf{w})$ for all k as

$$\text{SNR}_{k}(\mathbf{w}) = \frac{E\left\{\left[Y_{\text{D}}^{k}(\mathbf{w})\right]^{H} Y_{\text{D}}^{k}(\mathbf{w})\right\}}{E\left\{\left[Y_{\text{R}}^{k}(\mathbf{w}) - Y_{\text{D}}^{k}(\mathbf{w})\right]^{\text{H}} \left[Y_{\text{R}}^{k}(\mathbf{w}) - Y_{\text{D}}^{k}(\mathbf{w})\right]\right\}} \tag{3.6}$$

where $E[\cdot]$ is the statistical expectation operator. We have defined the SNR as the ratio of the desired signal, which excludes the effects of the noise including the ISI and ICI, to the difference between the received data and the desirable data. If the received data were to contain more noise, then the denominator of this quotient would grow, which would reduce the SNR. If there were no noise due to ISI and ICI, then the denominator would be reduced to the contributions of the AWGN, NEXT, and DNF.

Derive $\tilde{\mathbf{A}}_{k}$ (A.5) as

$$\tilde{\mathbf{A}}_{k} = \sigma_{s}^{2} \mathbf{H}^{\text{T}} \mathbf{Q}_{k}^{\text{circ}} \left[\mathbf{Q}_{k}^{\text{circ}}\right]^{\text{H}} \mathbf{H} \tag{3.7}$$

Similarly, derive $\tilde{\mathbf{B}}_k$ (A.6) as,

$$\tilde{\mathbf{B}}_k = 2\sigma_s^2 \left(\mathbf{H}_u^{\mathrm{T}} \mathbf{V}_k \mathbf{V}_k^{\mathrm{H}} \mathbf{H}_u + \mathbf{H}_{\mathrm{b}}^{\mathrm{T}} \mathbf{W}_k \mathbf{W}_k^{\mathrm{H}} \mathbf{H}_{\mathrm{b}} \right) + \mathbf{Q}_k^{\mathrm{noise}} \left\{ \sigma_{\mathrm{AWGN}}^2 \mathbf{I} + \mathbf{\Sigma}_{\mathrm{NEXT}} \right\} \left[\mathbf{Q}_k^{\mathrm{noise}} \right]^{\mathrm{H}} + \frac{\sigma_{\mathrm{DNF}}^2}{\mathbf{w}^{\mathrm{T}} \mathbf{w}} \mathbf{I} \quad (3.8)$$

where $\mathbf{I}$ is the identity matrix. $\tilde{\mathbf{A}}_k$ and $\tilde{\mathbf{B}}_k$ are Hermitian symmetric. Now (3.6) becomes

$$\mathrm{SNR}_k(\mathbf{w}) = \frac{\mathbf{w}^{\mathrm{T}} \tilde{\mathbf{A}}_k \mathbf{w}}{\mathbf{w}^{\mathrm{T}} \tilde{\mathbf{B}}_k \mathbf{w}} \quad (3.9)$$

$\mathrm{SNR}_k(\mathbf{w})$ is a ratio of quadratic functions of $\mathbf{w}$. The "same" quadratic form has been derived in [12, 14], but here the matrices $\tilde{\mathbf{A}}_k$ and $\tilde{\mathbf{B}}_k$ are derived using more sources of interference with a more intuitive approach.

3.1 Achievable data rate in terms of proposed SNR model

By substituting (3.9) into (2.4),

$$b_{\mathrm{DMT}}(\mathbf{w}, \mathcal{I}) = \sum_{k \in \mathcal{I}} \log_2 \left[\frac{\mathbf{w}^{\mathrm{T}} (\Gamma_k \tilde{\mathbf{B}}_k + \tilde{\mathbf{A}}_k) \mathbf{w}}{\mathbf{w}^{\mathrm{T}} (\Gamma_k \tilde{\mathbf{B}}_k) \mathbf{w}} \right] \quad (3.10)$$

where we include the dependence of b_{DMT} on $\mathbf{w}$. For further notational convenience, let $\mathbf{A}_k = \Gamma_k \tilde{\mathbf{B}}_k + \tilde{\mathbf{A}}_k$ and $\mathbf{B}_k = \Gamma_k \tilde{\mathbf{B}}_k$. Thus,

$$b_{\mathrm{DMT}}(\mathbf{w}, \mathcal{I}) = \sum_{k \in \mathcal{I}} \log_2 \left(\frac{\mathbf{w}^{\mathrm{T}} \mathbf{A}_k \mathbf{w}}{\mathbf{w}^{\mathrm{T}} \mathbf{B}_k \mathbf{w}} \right) \quad (3.11)$$

Matrices $\mathbf{A}_k$ and $\mathbf{B}_k$ are positive definite, since both the numerator and denominator are always positive numbers for $\mathbf{w} \neq \mathbf{0}$, which is what is expected given that both represent power. The maximization problem is

$$\mathbf{w}_{\text{opt}} = \arg\max_{\mathbf{w}} \{ b_{\text{DMT}}(\mathbf{w}, \mathcal{I}) | \mathbf{w} \in \mathcal{S} \} \tag{3.12}$$

where $\mathcal{S}$ is a set of constraints we choose to impose on $\mathbf{w}$ and $\mathcal{I} \subseteq \{0, \cdots, N/2-1\}$. Given that (3.11) is scale invariant (i.e., $b_{\text{DMT}}(\mathbf{w}, \mathcal{I}) = b_{\text{DMT}}(\alpha\mathbf{w}, \mathcal{I}), \forall \alpha \neq 0$), the magnitude of $\mathbf{w}$ is irrelevant[1]. Therefore, without loss of generality, we adopt the unit norm constraint set $\mathcal{S} = \{\mathbf{w} : \mathbf{w}^{\text{T}}\mathbf{w} = \|\mathbf{w}\|^2 = 1\}$. Notice that, $\tilde{\mathbf{B}}_k$ becomes independent of $\mathbf{w}$ over this constraint set since the last term in (3.8) becomes $\sigma^2_{\text{DNF}}\mathbf{I}$.

Now we return to the issue of whether the maximizer of (3.11) will be the maximizer of

$$b^{\text{int}}_{\text{DMT}}(\mathbf{w}, \mathcal{I}) = \sum_{k \in \mathcal{I}} \left\lfloor \log_2 \left(\frac{\mathbf{w}^{\text{T}}\mathbf{A}_k\mathbf{w}}{\mathbf{w}^{\text{T}}\mathbf{B}_k\mathbf{w}} \right) \right\rfloor \tag{3.13}$$

Figures 3.1 and 3.2 plot the values of both (3.11) and (3.13) for all possible values of $\mathbf{w}$ of length 2 and 3 taps, respectively, for a CIR involving the carrier serving area (CSA) loop 3 and transmit and receive filters. The maxima of both (3.11) and (3.13) happen for the same $\mathbf{w}$ in both 2- and 3-dimensional space. This is also true for the other seven CSA loops. Thus maximizing (3.11) maximizes (3.13) in these two cases; however, we cannot guarantee that this will be the case in general.

[1]This is not the case for the DNF noise if σ^2_{DNF} is not small or $\mathbf{w}$ imposes a null, but we assume that these cases are not generated by the TEQ design method proposed in this paper.

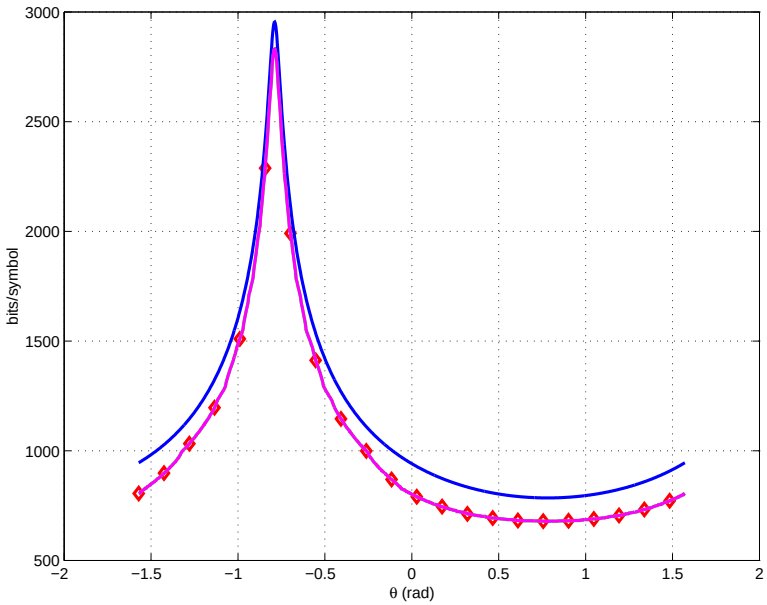

Figure 3.1: The evaluation of b_{DMT} (solid) and $b_{\mathrm{DMT}}^{\mathrm{int}}$ (solid with diamonds) for all possible values of $\mathbf{w} = [\sin(\theta), \cos(\theta)]^{\mathrm{T}}$, $\theta = \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, $M = 2$ for the CIR containing CSA loop 3 with Δ=15, $N = 512$ and $\nu = 32$ with input power = 0.2472 W, AWGN power =-140 dBm/Hz, NEXT modelled as 49 ADSL disturbers and θ sampled with 1081 points. Notice that $b_{\mathrm{DMT}}^{\mathrm{int}}$ follows b_{DMT} and that they share the single global maximizer in this case.

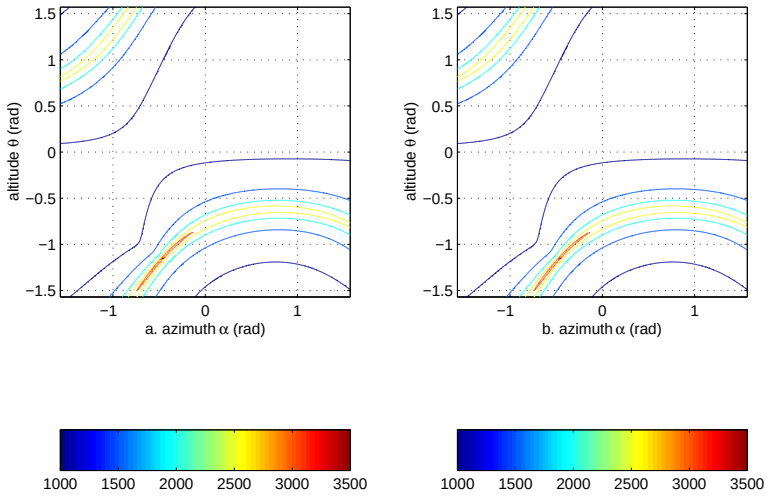

Figure 3.2: The evaluation of a. $b_{\mathrm{DMT}}^{\mathrm{int}}$ and b. b_{DMT} for all possible values of $\mathbf{w} = [\sin(\alpha)\sin(\theta), \cos(\alpha)\sin(\theta), \cos(\theta)]^{\mathrm{T}}$ for the azimuth angle $\alpha = \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ and altitude angle $\theta = \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ for the CIR containing CSA loop 3 with $M = 3$, Δ=15, $N = 512$ and $\nu = 32$ with input power = 0.2472 W, AWGN power =-140 dBm/Hz, NEXT modelled as 49 ADSL disturbers and both α and θ sampled at 1081 points. Note that the maxima of both b_{DMT} and $b_{\mathrm{DMT}}^{\mathrm{int}}$ occur for $\alpha = -0.4741$ rad and $\theta = -1.1606$ rad

3.2 Impact of Δ on TEQ design

The delay that marks the amount of time that it takes the signal to reach the receiver from the time it has been transmitted is visible in the CIR as the number of samples with small amplitude before the samples of the CIR containing most of the energy of the CIR. Any additional filter such as TEQ may add some delay to the composite response by shifting the energy to later samples. Jointly, the delays of CIR and TEQ make up the transmission delay Δ referred to first in (3.1) and then used throughout in the definition of the proposed SNR measure. One can think of Δ as a parameter which selects the beginning of a symbol from a stream of samples coming into the receiver. Notice that once Δ is defined, all of the matrices that compose the SNR are defined, as well as (3.6). Thus, defining Δ prior to TEQ design also defines (3.11) and its optimal solution. A different Δ may result in a higher value for (3.11). As of yet there is no known way to search for the optimal Δ without searching exhaustively through all possible values of Δ where $\Delta \in [0, N-1]$ and solving the optimization problem. This report does not propose a better or less costly method. Developing a heuristic for the choice of optimal Δ remains an open research problem.

4 Proposed TEQ Design

There are $N/2-1$ possible data carrying subchannels in DMT. We will now propose a method of finding the optimal $\mathbf{w}$ for every one of them.

4.1 TEQ for a single subchannel

Maximizing the number of bits allocated in a single channel, $b_k(\mathbf{w})$, involves maximizing the argument of the log function. Since the log is a monotonically increasing function for a non-negative argument, maximizing its non-negative argument will also maximize the function. Mathematical notation for this statement is

$$b_k^{\text{opt}} = \max_{\mathbf{w}_k:\|\mathbf{w}_k\|^2=1} \left[\log_2\left(\frac{\mathbf{w}_k^{\mathrm{T}}\mathbf{A}_k\mathbf{w}_k}{\mathbf{w}_k^{\mathrm{T}}\mathbf{B}_k\mathbf{w}_k}\right)\right] = \log_2\left[\max_{\mathbf{w}_k:\|\mathbf{w}_k\|^2=1}\left(\frac{\mathbf{w}_k^{\mathrm{T}}\mathbf{A}_k\mathbf{w}_k}{\mathbf{w}_k^{\mathrm{T}}\mathbf{B}_k\mathbf{w}_k}\right)\right]$$

From [23], the maximization of a single ratio can be transformed into

$$p_k(\mathbf{w}_k,\lambda_k) = \max_{\mathbf{w}_k:\|\mathbf{w}_k\|^2=1}\left\{\mathbf{w}_k^{\mathrm{T}}\mathbf{A}_k\mathbf{w}_k - \lambda_k\mathbf{w}_k^{\mathrm{T}}\mathbf{B}_k\mathbf{w}_k\right\} \tag{4.1}$$

where λ_k is a scalar. To solve (4.1), we compute the derivative of the argument of the maximum operator with respect to $\mathbf{w}_k$ and set the derivative to zero, which yields

$$\mathbf{A}_k^r\mathbf{w}_k = \lambda_k\mathbf{B}_k^r\mathbf{w}_k \tag{4.2}$$

Here $(\cdot)^r$ denotes the real part. This is the well-known generalized eigenvalue problem [24] and the solution is the generalized eigenvector $\mathbf{w}_k^{\text{opt}}$ corresponding to the largest generalized

eigenvalue λ_k^{opt} of $(\mathbf{A}_k^r, \mathbf{B}_k^r)$:

$$\lambda_k^{\text{opt}} = \frac{\left(\mathbf{w}_k^{\text{opt}}\right)^{\text{T}} \mathbf{A}_k^r \mathbf{w}_k^{\text{opt}}}{\left(\mathbf{w}_k^{\text{opt}}\right)^{\text{T}} \mathbf{B}_k^r \mathbf{w}_k^{\text{opt}}} = \frac{\left(\mathbf{w}_k^{\text{opt}}\right)^{\text{T}} \mathbf{A}_k \mathbf{w}_k^{\text{opt}}}{\left(\mathbf{w}_k^{\text{opt}}\right)^{\text{T}} \mathbf{B}_k \mathbf{w}_k^{\text{opt}}} \tag{4.3}$$

Hence, $b_k^{\text{opt}} = \log_2(\lambda_k^{\text{opt}})$. If an optimal TEQ were found for every subchannel, then the bit allocation for every one of those subchannels would be maximized which would lead to

$$b_{\text{DMT}}^{\text{opt}} = \sum_{k \in \mathcal{I}} \log_2 \left[\frac{(\mathbf{w}_k^{\text{opt}})^{\text{T}} \mathbf{A}_k \mathbf{w}_k^{\text{opt}}}{(\mathbf{w}_k^{\text{opt}})^{\text{T}} \mathbf{B}_k \mathbf{w}_k^{\text{opt}}} \right] \tag{4.4}$$

We could realize (4.4) as the optimal per-tone TEQ filter bank (TEQFB) with $|\mathcal{I}|$ TEQs where $|*|$ denotes the cardinality of the set. The incoming signal stream would be split into $|\mathcal{I}|$ signals with each signal going to its own TEQ and FFT blocks.

Obviously, this procedure is computationally very expensive. A preferred implementation would be to employ a bank of Goertzel filters [17] with each one computing a single point DFT. If we denote $y_k(n)$ as the received signal after every TEQ ($k \in \mathcal{I}$), then the corresponding frequency response Y_k for the k^{th} subcarrier would be efficiently computed as follows:

$$G_k(n) = y_k(n) + 2\cos\left(\frac{2\pi k}{N}\right) G_k(n-1) - G_k(n-2) \tag{4.5}$$

$$Y_k = \left[G_k(N) - G_k(N-1)\cos\left(\frac{2\pi k}{N}\right) \right] + j \left[G_k(N-1)\sin\left(\frac{2\pi k}{N}\right) \right]$$

where $G_k(-1) = G_k(-2) = 0$ and $n = 0,\ 1, \cdots,\ N$. In either case, computational complexity is certainly higher than using a single FFT block; however, using Goertzel filters provides us with the upper limit on what we can achieve in terms of maximizing bit allocation and is used for

that purpose in this paper. Theoretically, any other arrangement of TEQ filters, be it a single TEQ or multiple, can only perform at par with the TEQFB or worse.

We found the optimal solution for the problem of maximizing (4.1) which allows for fractional bit loading. If we take into account that only integer bit loading is allowed in ADSL and VDSL, then we can say that possibly more than one $\mathbf{w}$ achieves maximum integer bit loading for the k^{th} subchannel. The set of solutions is:

$$\mathcal{J}_k = \{\mathbf{w} : \ \lfloor b_k(\mathbf{w}) \rfloor = \lfloor b_k(\mathbf{w}_k^{opt}) \rfloor, \ \mathbf{w} \in \mathcal{S}\} \tag{4.6}$$

Ideally, we would like to find a TEQ that will perform as well on the subchannels of interest, i.e., those carrying data. The set of these filters is

$$\mathcal{J} = \{\mathbf{w} : \ \mathbf{w} \in \bigcap_{k \in \mathcal{I}} \mathcal{J}_k\} \tag{4.7}$$

Set $\mathcal{J}$ may contain many points due to the integer-only loading and/or the behavior of b_{DMT}, or may be an empty set. The task of any maximization procedure is to find the best possible solution, i.e., one that provides a bit allocation as close as possible to (4.4).

4.2 Optimal TEQ for the subchannels of interest

The problem we are considering is

$$\mathbf{w}_{\text{opt}} = \arg\max_{\mathbf{w}} \{b_{\text{DMT}}(\mathbf{w}, \mathcal{I}) | \mathbf{w} \in \mathcal{J}\} \tag{4.8}$$

The bit rate (3.11) is a difficult function to optimize. One can choose to parameterize the problem according to the constraint set $\mathcal{S}$, by imposing a constraint that $\mathbf{w}$ is a point on the M-dimensional unit hyper-sphere, given by

$$w_0 = \prod_{l=1}^{M-1} \sin(\theta_l) \text{ and } \mathrm{w_m} = \cos(\theta_\mathrm{m}) \prod_{\mathrm{l=m+1}}^{\mathrm{M-1}} \sin(\theta_\mathrm{l}),\ \mathrm{m} = 1, 2, \cdots, \mathrm{M}-1 \tag{4.9}$$

The problem may be solved by performing an exhaustive search on the set

$$\mathcal{T} = \{(\theta_1, \theta_2, \cdots, \theta_{M-1}) \text{ and } \theta_\mathrm{m} \in [-\pi/2, \pi/2],\ \mathrm{m} = 1, 2, \cdots, \mathrm{M}-1\} \tag{4.10}$$

When $M = 2$, this approach leads to a simple line search on $\theta_1 \in [-\pi/2, \pi/2]$. In this case, there is a single mode corresponding to the global maximum of the function and is found for a particular value of θ_1. Therefore, the solution can be found using virtually any method that can handle a line search and a single maximum. However, for $M > 2$ the structure changes to a multimodal structure with several maxima as demonstrated in Figure 3.2. We present a method that will find the global maximum provided that we have a "good" starting point. We also present a method for finding a good starting point. The method guarantees the closest local maximum to our starting point, without the oscillatory behavior associated with standard steepest ascent algorithms [25].

We apply Almogy and Levin's method [26], which is based on the Dinkelbach approach [27], to find the roots of the first derivative of (3.11). Almogy and Levin intended to maximize a sum-of-ratios problem

$$b(x) = \max_{x \in R} \sum_{i=1}^{n} \frac{f_i(x)}{g_i(x)} \tag{4.11}$$

which they transformed into the parametric problem

$$H_n(q) = \max_{x \in R} \sum_{i=1}^{n} [f_i(x) - q_i g_i(x)] \tag{4.12}$$

where, by analogy to the single-ratio maximization Dinkelbach approach (which is the same as Lagrangian multipliers [25]), they defined

$$q_i = \frac{f_i(x_p)}{g_i(x_p)} \tag{4.13}$$

where x_p is the solution of the maximization in the previous step. In a single-ratio problem ($n = 1$) the solution is reached at the zero of $H_1(q)$ where q is maximized and the point x maximizing the current iteration is the same as x_p. $H_1(q)$ is a convex, non-increasing function of q with a single root. In that problem q increases with every iteration, while $H_1(q)$ decreases. By analogy to Dinkelbach method, Almogy and Levin solve $H_n(q) = 0$ to find the optimal solution of the sum-of-ratios problem. This approach is erroneous as shown by Falk and Palocsay [28] and validated in our experiments in using Almogy and Levin's iteration to maximize a version of (3.11). That said, function $H_n(q)$ is still a convex, non-increasing function of q with a single crossing [23]. In the sum-of-ratios problem, finding a zero of $H_n(q)$ does not maximize $b(x)$. A method to maximize a sum-of-ratios is an active research topic in the fractional programming community for which no definitive solution exists yet [23, 29]. However, Almogy and Levin's iteration does find the roots of the nonlinear $H_n(q)$ efficiently; hence, we use their idea with modifications specific to our problem, to find an optimal root of the first derivative of (3.11) that corresponds to the closest local maximum to the set (4.7). As indicated above, given a good initial point, the local maximum can be a global maximum. Our

experimental results presented in the following section support that. The first derivative of (3.11) is

$$\frac{db_{\mathrm{DMT}}(\mathbf{w},\mathcal{I})}{d\mathbf{w}} = \frac{2}{\ln 2}\sum_{k\in\mathcal{I}} r_k(\mathbf{w})\left[\mathbf{A}_k^r - \lambda_k(\mathbf{w})\mathbf{B}_k^r\right]\mathbf{w} \tag{4.14}$$

where

$$r_k(\mathbf{w}) = \frac{1}{\mathbf{w}^{\mathrm{T}}\mathbf{A}_k\mathbf{w}} \text{ and } \lambda_k(\mathbf{w}) = \frac{\mathbf{w}^{\mathrm{T}}\mathbf{A}_k\mathbf{w}}{\mathbf{w}^{\mathrm{T}}\mathbf{B}_k\mathbf{w}} \tag{4.15}$$

Notice that $b_{\mathrm{DMT}}(\mathbf{w},\mathcal{I}) = \sum_{k\in\mathcal{I}} \log_2\left[\lambda_k(\mathbf{w})\right]$, thus increasing $\lambda_k(\mathbf{w})$ increases $b_{\mathrm{DMT}}(\mathbf{w},\mathcal{I})$. Now we can write

$$\mathbf{C}_k(\mathbf{w}) = r_k(\mathbf{w})\left[\mathbf{A}_k^r - \lambda_k(\mathbf{w})\mathbf{B}_k^r\right] \text{ and } \mathbf{C}(\mathbf{w},\mathcal{I}) = \sum_{k\in\mathcal{I}} \mathbf{C}_k(\mathbf{w}) \tag{4.16}$$

This leads to an equation similar to (4.12) in which $r_k(\mathbf{w})$, $\lambda_k(\mathbf{w})$ and $\mathbf{C}_k(\mathbf{w})$ are projected according to (4.15) and (4.16) during the iterative procedure that finds the optimal root of (4.14). Let

$$H(\lambda) = \max_{\mathbf{w}\in\mathcal{S}} \mathbf{w}^{\mathrm{T}}\mathbf{C}(\mathbf{w},\mathcal{I})\mathbf{w} = \max_{\mathbf{w}\in\mathcal{S}} \sum_{k\in\mathcal{I}} r_k\left[\mathbf{w}^{\mathrm{T}}\mathbf{A}_k^r\mathbf{w} - \lambda_k\mathbf{w}^{\mathrm{T}}\mathbf{B}_k^r\mathbf{w}\right]$$

where $\lambda = [\cdots, \lambda_k, \cdots]^{\mathrm{T}}, k \in \mathcal{I}$. Now we see the similarities between (4.12) and (4.17). Our method uses Almogy and Levin's method as basis, but modified with specific requirements of our problem to find a zero of (4.17) corresponding to the closest maximum of (3.11). In the process, $\lambda_k(\mathbf{w})$ will always increase, thereby increasing (3.11).

4.3 TEQ design algorithm

For the initial point of the optimization algorithm, choose $\mathbf{w}_k^{\mathrm{opt}}$ from TEQFB which gives the highest data rate, i.e. choose initial $\mathbf{w}^{\mathrm{opt}} = \mathbf{w} = \mathbf{w}_k^{\mathrm{opt}}$. Set iteration counter $i = 0$, smoothing

1. $r_k = \alpha r_k + (1-\alpha)\frac{1}{\mathbf{w}^{\mathrm{T}}\mathbf{A}_k\mathbf{w}}, \forall k \in \mathcal{I}$
2. $\lambda_k = \alpha\lambda_k + (1-\alpha)\frac{\mathbf{w}^{\mathrm{T}}\mathbf{A}_k\mathbf{w}}{\mathbf{w}^{\mathrm{T}}\mathbf{B}_k\mathbf{w}}, \forall k \in \mathcal{I}$
3. Compute $\mathbf{C}(\mathbf{w},\mathcal{I}) = \sum_{k\in\mathcal{I}} r_k \left[\mathbf{A}_k^r - \lambda_k \mathbf{B}_k^r\right]$
4. $\mathbf{w}_{\mathrm{new}} = \arg\max\{\mathbf{v}^{\mathrm{T}}\mathbf{C}(\mathbf{w},\mathcal{I})\mathbf{v}, \mathbf{v} \in \mathcal{S}\}$
5. If $\|\mathbf{w}_{\mathrm{new}} - \mathbf{w}\|_\infty < \epsilon$ OR $i > i_{\max}$ then return $\mathbf{w}_{\mathrm{opt}}$.
6. If $b_{\mathrm{DMT}}(\mathbf{w}_{new},\mathcal{I}) < b_{\mathrm{DMT}}(\mathbf{w},\mathcal{I})$ set $\alpha = (1+\alpha)/2$, ELSE $\mathbf{w}_{\mathrm{opt}} = \mathbf{w}_{\mathrm{new}}$.
7. $\mathbf{w} = \mathbf{w}_{\mathrm{new}}$
8. $i = i + 1$
9. Go back to step 1 and repeat.

Figure 4.1: Proposed TEQ Design algorithm

factor $\alpha = 0$ and values r_k and λ_k to zero for all k. The algorithm proceeds as shown in Figure 4.1.

The key steps in the proposed TEQ design algorithm are the computations of $b_{\mathrm{DMT}}(\mathbf{w}_k^{\mathrm{opt}},\mathcal{I}), k \in \mathcal{I}$ and $\mathbf{w}_{\mathrm{new}}$, $i = 1, 2, \cdots$, which are solved by finding the generalized eigenvector corresponding to the largest eigenvalue of a pair of matrices. Many algorithms find the largest generalized eigenvalue and many of them are publicly available. Using ideas from the inverse power method and conjugate gradients [30], we have derived an efficient algorithm for solving these key steps which can easily be implemented on most digital signal processors available today, however a detailed explanation of it is beyond the scope of this report. Further research will include full development of this new and efficient algorithm.

5 Simulation Results

The simulation results compare the performance of the proposed TEQ design methods with Min-ISI, MBR and LS PTE initialized using least-squares methods. We use the eight standard downstream CSA loops [3] convolved with transmit and receive filters as the test CIR. The transmit and receive filters are modelled as first-order high-pass infinite impulse response (IIR) filters, which are designed to separate ADSL from the voice band (0-4 kHz)(Fig. 5.1). All CIR consist of 512 samples sampled at 2.208 MHz. Upstream was not simulated for the purposes of this paper as the downstream TEQ problem deals with equalization over a much wider bandwidth and as such is a more challenging problem. We, thus, use the FFT size $N = 512$ standard in downstream ADSL with $\Gamma = 11.8$ dB. All of the powers used in the simulations are defined with respect to a 100 Ω resistance. The power of the signal is 0.2475 W spread equally over all of the subchannels. AWGN power is equal to -140 dBm/Hz over the bandwidth of 1.104 MHz with the NEXT source being modelled as 49 ADSL disturbers. The PSD of the NEXT is defined in the ADSL standard [3]. Delay Δ is taken as a free parameter in the optimization and is varied over all possible values with the best achieved data rate taken to be the best result. Thus, once Δ is chosen the optimization proceeds with the search for either TEQFB or a TEQ.

Single TEQ FIR structures designed by Min-ISI and MBR methods are obtained using the Matlab DMT TEQ Design Toolbox [2]. The coding gain assumed equals 4.2 dB while the margin used equals 6 dB for all data rates reported [1]. LS PTE structures of [16] were designed using least-squares methods that do not require knowledge of the channel model.

[1]While the bit loading in ADSL is designed to yield the bit-error rate (BER) of 10^{-7} at 0 dB margin the measured SNR is used to determine bit loading) in practice often additional safety is sought in the form of margin to budget for unforseen increases in noise or insufficiently accurate SNR measurements.

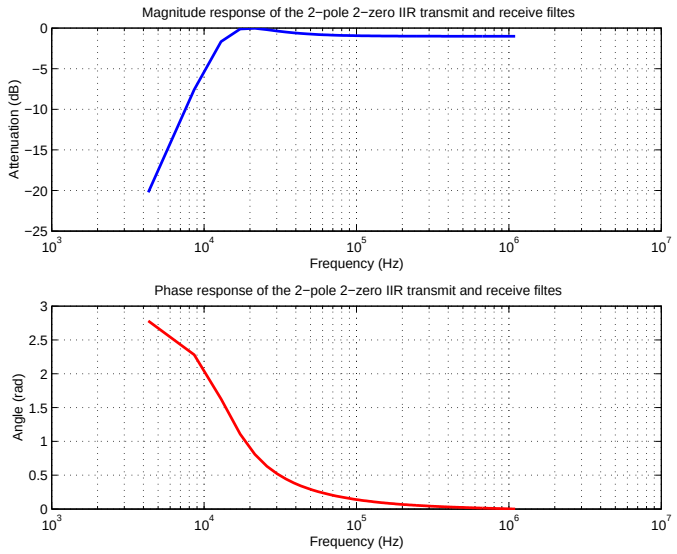

Figure 5.1: Phase and Magnitude Response of the IIR filter used to model transmit and receive filters. A double zero is located at 1, while conjugate symmetric poles are located at $0.9799 \pm$ j0.0317. Transmit and receive high-pass filters are meant to separate the telephone band from the ADSL band.

The bit rates reported are calculated using the estimated SNR and following the ADSL standard for 10^{-7} BER. Subchannels of interest (7 to 256 from [3], excluding 64, which is reserved for the pilot tone) are loaded with a randomly chosen two bit constellation point at the transmitter. The symbols are convolved with the CIR, passed through the TEQ block designed by one of the methods, and then passed through an FFT block after which the phase and magnitude distortion is removed and slicing operation is performed. The slicing operation compares the complex value received in a particular subchannel with the expected value so that SNR measurement can be derived from the power of the error averaged over 1000 symbols.

We calculate the bound on the SNR estimation using 1000 symbols to be approximately ± 0.5 dB which means that our estimate of the SNR is within 0.5 dB of the true SNR. Correspondingly, the reported bit rates can vary approximately ± 60 kbps. The SNR reported (and consequently, bit rate) was calculated using double-precision arithmetic.

Proposed TEQ and TEQFB , Min-ISI, MBR and LS PTE design methods are evaluated using

this procedure, which establishes a common testing platform. In the past, some TEQ design methods have used objective functions that have been derived from their formulation of the problem (such as MSE, Geometric SNR, etc.) to report the validity of their design. However, these designs often failed to increase the bit rate to its full potential although their objective functions projected success. The testing platform used in this paper is the most objective method of testing a TEQ design as it relates the success of the design to the increase in data rate achievable on a real system.

In Figure 5.2 we show how the achieved bit rate varies with the change in the number of TEQ taps M ranging from 2 to 32 as an example for the CIR containing CSA loop 2. The bit rate grows significantly from 2 to 3 TEQ taps with the upward slope present with the further increase in the number of TEQ taps but significantly moderated. The proposed design of TEQFB performs the best, which is not surprising as it is guaranteed to be optimal by design and is upper envelope of the achieved bit rate curves of other methods. TEQFB outperforms LS PTE although both seek to equalize the signal on the per tone basis. LS PTE performs poorly for M ranging from 2-6 after which its performance can be compared to other methods under consideration. The proposed TEQ design performs closely to TEQFB and outperforms Min-ISI, MBR and LS PTE.

Table 5.1 lists the data rate achieved with the proposed optimal TEQFB for the CIR including CSA loops 1-8 . These data rates represent the maximum data rate that can be achieved as a function of TEQ for the given signal and noise power levels.

Table 5.2 shows the achieved data rates of the proposed TEQ design method, MBR, Min-ISI and LS PTE for the CIR including CSA loops 1-8 as percentage (%) of the data rate of the proposed TEQFB shown in Table 5.1. Experiments are performed for each CSA loop with M

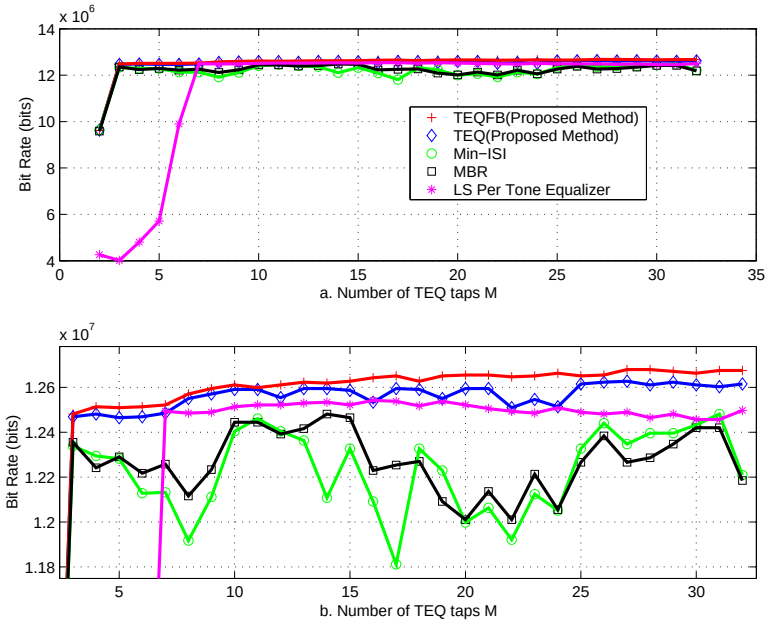

Figure 5.2: Data rates achieved for different number of TEQ taps M for the CIR containing CSA loop 2 for $N = 512$ and $\nu = 32$ with input power = 0.2472 W, AWGN power =-140 dBm/Hz, NEXT modelled as 49 ADSL disturbers. a. Entire graph and b. Detail of the higher data rates. Note that the reported rates are up to ± 60 Kbps away from true values

Table 5.1: Highest achieved bit rates in Mbps of the proposed TEQFB for the CIR involving standard CSA loops 1-8 with transmit and receive high-pass filters for $N = 512$ and $\nu = 32$ with input power = 0.2472 W, AWGN power =-140 dBm/Hz, NEXT modelled as 49 ADSL disturbers and optimal Δ.

CIR containing	**Proposed TEQFB**
CSA loop 1	11.417
CSA loop 2	12.680
CSA loop 3	10.995
CSA loop 4	11.288
CSA loop 5	11.470
CSA loop 6	10.861
CSA loop 7	10.752
CSA loop 8	9.615

Table 5.2: Average achieved bit rates for various TEQ design methods averaged over TEQ sizes $M \in \{2, \cdots, 32\}$ as a percentage of the achieved bit rate of the proposed optimal TEQFB for the CIR involving standard CSA loops 1-8 with transmit and receive high-pass filters for $N = 512$ and $\nu = 32$ with input power = 0.2472 W, AWGN power =-140 dBm/Hz, NEXT modelled as 49 ADSL disturbers.

CIR containing	**Proposed Single TEQ**	**MBR**	**Min-ISI**	**LS PTE**
CSA loop 1	99.60%	97.29%	97.51%	99.04%
CSA loop 2	99.59%	97.02%	97.33%	90.41%
CSA loop 3	99.55%	97.82%	97.33%	99.15%
CSA loop 4	99.28%	98.14%	98.18%	96.73%
CSA loop 5	99.57%	97.71%	97.17%	97.40%
CSA loop 6	99.47%	97.72%	98.31%	95.28%
CSA loop 7	98.80%	96.27%	96.25%	95.36%
CSA loop 8	98.65%	97.38%	97.47%	99.10%
Average	**99.31%**	**97.42%**	**97.45%**	**96.56%**

ranging from 2 to 32 for standard $N = 512$ and $\nu = 32$. Achieved data rates for tabulated methods were expressed as a percentage of the performance of TEQFB for the same M and these percentages were averaged to arrive at a single percentage number for each CSA loop. These are the numbers tabulated in the columns for each method. Finally, each column was averaged in the numbers listed in the last row to arrive at the performance across CSA loops. TEQFB is always higher than the compared methods. The proposed TEQ design method achieves higher percentage for each CIR than Min-ISI, MBR and LS PTE. Proposed TEQ final average almost 2% higher than either MBR or the Min-ISI, and 2.75% higher than LS PTE . For a data rate of 11 Mbps a 2% improvement amounts to 220 kbps.

In Figure 5.3 we compare how the achieved bit rate varies with the change in the CP length ν from 1 to 32 for a fixed number of TEQ taps, $M = 3$. We use $M = 3$ because results in Figure 5.2 show that is the smallest M to achieve very high data rates. The bit rate steadily increases with the increase in ν. The proposed design of TEQFB outperforms other methods. The proposed TEQ design performs closely to TEQFB and outperforms both Min-ISI, MBR and

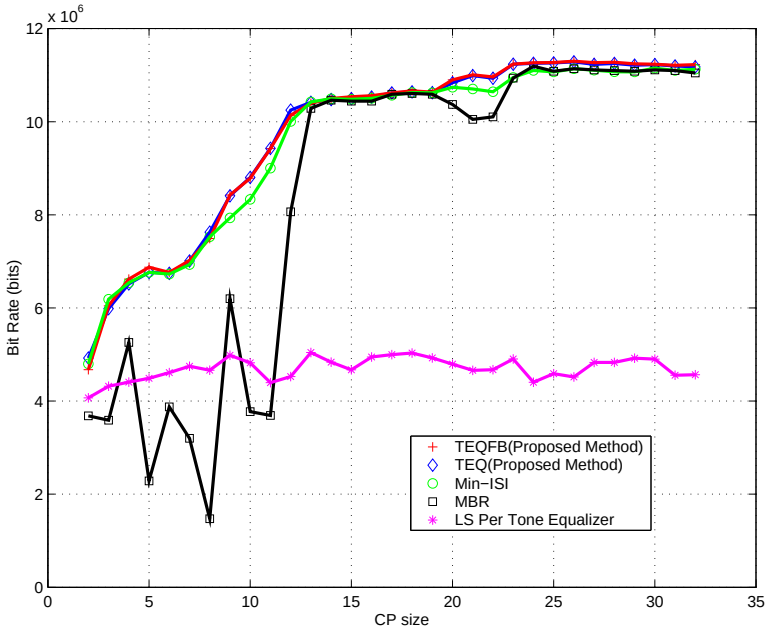

Figure 5.3: Data rates achieved for different sizes of the cyclic prefix for the CIR containing CSA loop 5 for $N = 512$ and $M = 3$ with input power = 0.2472 W, AWGN power =-140 dBm/Hz, NEXT modelled as 49 ADSL disturbers.

LS Per Tone TEQ. Poor performance of LS PTE can be linked to Figure 5.2 where for $M = 3$ still does not approach the data rates of other methods. Evidently, the critical parameter in PTE design is the length of the TEQ filters and not necessarily CP length.

Figure 5.4 shows how the bit rate changes as a function of transmission delay Δ for the proposed single TEQ design method. The CIRs tested include CSA loop 1 for TEQ length $M = \{3, 10, 30\}$, $N = 512$, $\nu = 32$. Generally, the bit rate rises over a small range of Δ, plateaus for a range of Δ and declines for the rest of the measured delays. Increasing the number of TEQ taps brings about an increase in the data rate but most prominently increases the range of Δ over which the bit rate function plateaus. This reduces the sensitivity of the design to the choice of Δ as the near maximum can be achieved for a larger number of Δ. The fact that the bit rate as a function of Δ has a high degree of "regularity" may aid future TEQ design methods when it comes to the choice of Δ.

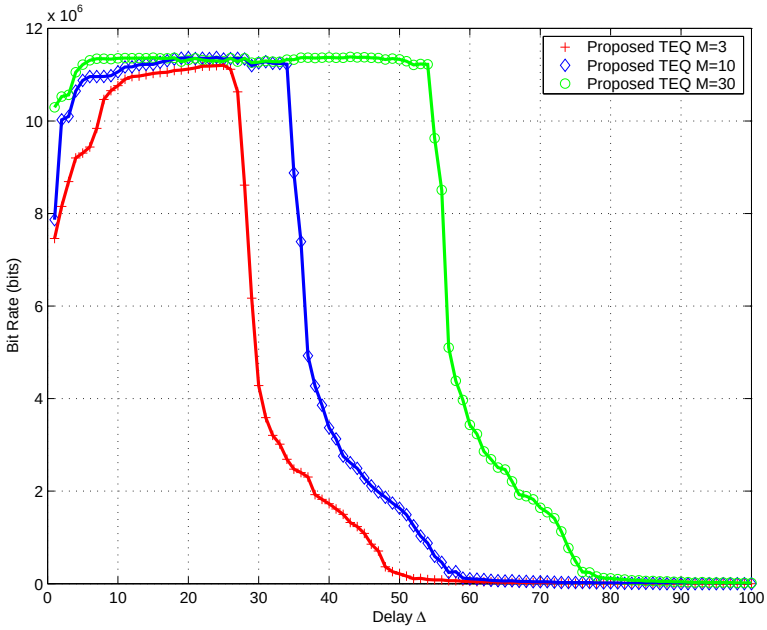

Figure 5.4: Bit rate achieved as a function of Δ for the CIR including CSA loop 1 using $M = 3, 10, 30$ for $N = 512$ and $\nu = 32$ with input power = 0.2472 W, AWGN power =-140 dBm/Hz, NEXT modelled as 49 ADSL disturbers.

Figure 5.5 compares the SNR achieved with the proposed single TEQ design, MBR and Min-ISI. The figure gives insight into why the performance of the proposed TEQ design outperforms the compared methods. Both MBR and Min-ISI in this particular example tend to put nulls in the SNR thus reducing the data rate. Our experimental observations suggest that a successful TEQ has a flat magnitude response over most of the spectrum and a null at the position of the highest ISI whereas a less successful design will put a number of nulls elsewhere in the spectrum.

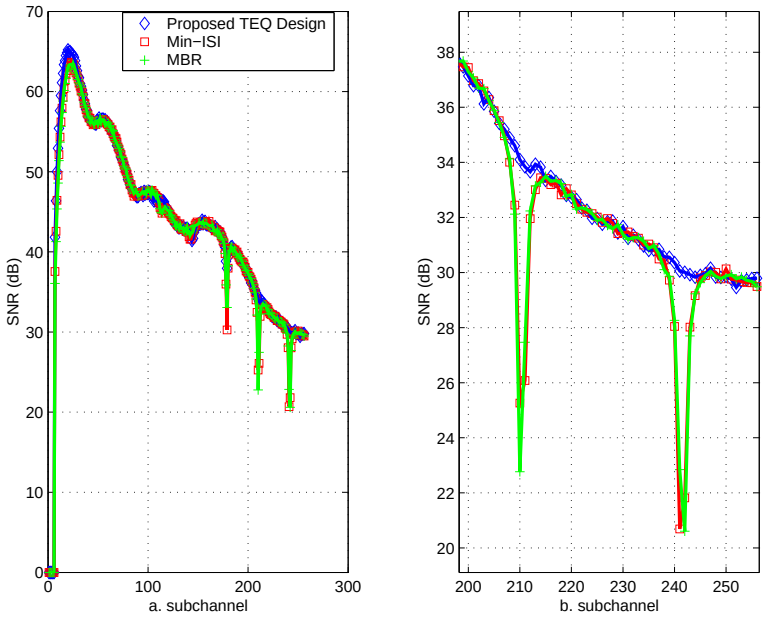

Figure 5.5: SNR achieved using the proposed TEQ, MBR and Min-ISI for CSA loop 4 with $M = 18$, $N = 512$, $\nu = 32$, input power = 0.2472 W, AWGN power =-140 dBm/Hz, and NEXT modelled as 49 ADSL disturbers; both MBR and Min-ISI put nulls in the SNR while the proposed single TEQ design does not. Nulls will limit the data rate.
a. SNR for all carriers of interest, b. Detail of nulls

6 Conclusions From Completed Work

In this report we propose a new SNR measure as a function of TEQ based on the desire to have a circular convolution of the signal and the transmission channel present at input to the FFT. The noise measure is defined as the deviation of the received signal from this ideal and it includes the contribution of the non-signal dependent sources such as AWGN, NEXT and DNF. The SNR is a ratio of quadratic functions in TEQ $\mathbf{w}$. We use the SNR measure in the definition of the bit rate which is what we attempt to maximize using TEQ. We propose two methods to maximize the bit rate using TEQ. The first method, TEQFB, involves the optimal per-tone equalization in the time domain where the bit rate in each subchannel of interest is maximized with a separate TEQ. Each TEQ is arrived at as the eigenvector corresponding to the largest generalized eigenvalue of the signal and noise matrices pair for the particular subchannel. During data transmission this TEQ filter bank would employ Goertzel filters to arrive at the frequency domain point corresponding to the particular subchannel. This approach reduces the number of computations needed. The second method aims to arrive at a single TEQ that will perform as close as possible to TEQFB. We propose a good starting point being the subchannel TEQ that achieves highest bit rate and then proceed to minimize the derivative of the function using Almogy-Levin iteration. The results show that proposed TEQ design achieves on average 99.31% of the proposed TEQFB performance where the average is taken over the standard CSA loops 1-8 and over TEQ with taps numbering 2-32 which outperforms state of the art MBR and Min-ISI by 2% and LS PTE by 2.75%. Our experimental observations suggest that a successful TEQ has a flat magnitude response over most of the spectrum and a null at the position of the highest ISI whereas a less successful design will put a number of nulls elsewhere in the spectrum.

7 Proposed Work

The expected contributions of the dissertation research are:

1. a new model for the SNR (*completed work*, submitted in [1]),

2. data rate optimal time domain per-tone TEQ filter bank (TEQFB) algorithm (*completed work*, submitted in [1]),
   (a) proof of optimality of TEQFB (*completed work*, submitted in [1]),
   (b) a new achievable upper bound on bit rate performance (*completed work*, submitted in [1]),
   (c) computationally less intensive algorithm for TEQFB then obtained straight from the equations (*current work*),
   (d) release of the software performing TEQFB designs as a part of a new release of the MATLAB DMT TEQ Design Toolbox [2] (*future work*),

3. data rate maximization single TEQ design algorithm (*completed work*, submitted in [1])
   (a) computationally less intensive algorithm for single TEQ then obtained straight from the equations (*current work*),
   (b) release of the software performing single TEQ design as a part of a new release of the MATLAB DMT TEQ Design Toolbox [2] (*future work*),
   (c) data rate maximization time domain TEQ filter bank optimized for groups of tones with the expectation of performance similar to TEQFB but at lower cost (*future work*), and
   (d) depending on the results from item (3c), add item (3d) to a new release of the MATLAB DMT TEQ Design Toolbox (*future work*).

Proposed milestones to complete research:

1. July 2002 - (2c),(3a) completed,

2. Aug 2002 - (2d),(3b) completed,

3. Oct 2002 - (3c) completed, and

4. Nov 2002 - (3d) completed if deemed valuable.

Bibliography

[1] M. Milosevic, L. F. C. Pessoa, B. Evans, and R. Baldick, "Optimal time domain equalization design for maximizing data rate of discrete multi-tone systems," *IEEE Transactions on Signal Processing*, Submitted Feb. 2002.

[2] G. Arslan, B. Lu, and B. L. Evans, "Matlab DMT TEQ design toolbox," Tech. Rep. available at http://signal.ece.utexas.edu/∼ arslan/dmtteq/dmtteq.html, Embedded Signal Processing Lab, The University of Texas at Austin, July 2001.

[3] ANSI, "Network and customer installation interfaces: Asymmetric digital subscriber line (ADSL) metallic interface," in *American National Standard for Telecommunications*, no. T1E1.413, 1998.

[4] ITU-T, "Asymmetrical digital subscriber line (ADSL) transceivers," in *Int. Telecommunications Union*, no. G.992.1, 1999.

[5] ITU-T, "Splitterless asymmetric digital subscriber line (ADSL) transceivers," in *Int. Telecommunications Union*, no. G.992.2, 1999.

[6] VDSLalliance, "Very-high-speed digital subscriber lines; system requirements; draft technical report," Tech. Rep. T1E1.4/97-131R2, American National Standard for Telecommunications, 1997.

[7] VDSLalliance, "SDMT VDSL draft standard proposal," Tech. Rep. T1E1.4/97-332, American National Standard for Telecommunications, 1997.

[8] ETSI, "Transmission and multiplexing (TM); access transmission systems on metallic access cables; very high speed digital subscriber line (VDSL); part 2: Transceiver specification," in *European Telecommunications Standards Institute*, no. TS 101 270-2, 1998.

[9] J. Cioffi, "A multicarrier primer," Tech. Rep. T1E1.4/91-157, Amati Comun. Corp. and Stanford University, 1991.

[10] J. Proakis and D. Manolakis, *Digital Signal Processing*, pp. 415–420. Prentice Hall, 3 ed., 1996.

[11] N. Al-Dhahir and J. Cioffi, "Efficiently computed reduced-parameter input-aided MMSE equalizers for ML detection: A unified approach," *IEEE Transactions on Communications*, vol. 42, pp. 903–915, Jan 1996.

[12] G. Arslan, B. Evans, and S. Kiaei, "Equalization for discrete multitone transceivers to maximize bit rate," *IEEE Transactions on Signal Processing*, vol. 29, pp. 845–866, Dec. 2001.

[13] B. Farhang-Boroujeny and M. Ding, "An eigen-approach to design of near-optimum time domain equalizer for DMT tranceivers," in *IEEE Int. Conf. on Communications*, (Vancouver, Canada), pp. 937–941, 1999.

[14] P. Melsa, R. Younce, and C. Rohrs, "Impulse response shortening for discrete multitone transceivers," *IEEE Transactions on Communications*, vol. 44, pp. 1662–1672, Dec 1996.

[15] D. Daly, C. Heneghan, and A. D. Fagan, "A minimum mean-squared error interpretation of residual ISI channel shortening for discrete multitone transceivers," in *Proc. IEEE ICASSP-01*, (Salt Lake City), May 2001.

[16] K. V. Acker, G. Leus, M. Moonen, O. van de Wiel, and T. Pollet, "Per tone equalization for DMT-based systems," *IEEE Transactions on Communications*, vol. 49, pp. 109–119, Jan 2001.

[17] G. Goertzel, *An algorithm for the evaluation of finite trigonometric series*, vol. 65, pp. 34–35. American Math. Monthly, 1958.

[18] A. Oppenheim and R. Schafer, *Discrete-time signal processing.* Prentice-Hall Inc., 1992.

[19] T. Starr, J. Cioffi, and P. Silverman, *Understanding digital subscriber line technology*, pp. 205–206. Prentice-Hall, 1999.

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

[21] D. Bengtsson and D. Landstrm, "Coding in a discrete multitone modulation system," Master's thesis, Lulea University of Technology, University Campus, Porsn, 97187 Lule, Sweden, March 1996.

[22] H. Stark and J. Woods, *Probability, random processes, and estimation theory for engineers*, pp. 213–218. Prentice Hall, 2 ed., 1994.

[23] S. Schaible, "Fractional programming - a recent survey," *Journal of Statistics and Management Systems*, vol. 29, pp. 845–866, March 2001.

[24] J. Demmel, *Applied Numerical Linear Algebra*, pp. 173–176. SIAM, 1 ed., 1997.

[25] D. P. Bertsekas, *Nonlinear Programming*, pp. 54–75, 253. Attena Scientific, 1 ed., 1995.

[26] Y. Almogy and O. Levin, *A Class of Fractional Programming Problems*, vol. 19, pp. 57–67. Operations Research, 1971.

[27] W. Dinkelbach, *On Nonlinear Fractional Programming*, vol. 13, pp. 492–498. Management Science, 1967.

[28] J. E. Falk and S. W. Palocsay, *Optimizing the Sum of Linear Fractional functions*, pp. 221–258. Kluwer Academic Publishers, 1992.

[29] R. Freund and F. Jarre, "Solving the sum-of-ratios problem by an interior-point method," Tech. Rep. 3/99, Bell Labs, 1999.

[30] G. H. Golub and C. F. V. Loan, *Matrix Computations.* John Hoppkins, 1 ed., 1996.

A Appendix

A.1 Definition of signal-dependent matrices: $\mathbf{U}_i^{\Delta}$, $\mathbf{U}_{i-1}^{\Delta}$, $\mathbf{U}_{i+1}^{\Delta}$, and $\left[\mathbf{U}_i^{\Delta}\right]_{\text{circ}}$; all $N \times (N+M-1)$ matrices

Let

$$\mathbf{U}_i^{\Delta} = \left[\begin{array}{cc} \left(\mathbf{U}_i^{\Delta}\right)_L & \left(\mathbf{U}_i^{\Delta}\right)_R \end{array} \right] \tag{A.1}$$

where

$$\left(\mathbf{U}_i^{\Delta}\right)_L = \begin{bmatrix} u_i^{\Delta} & u_i^{\Delta-1} & \cdots & u_i^{0} & u_i^{N-1} & \cdots & u_i^{N-\nu} \\ u_i^{\Delta+1} & u_i^{\Delta} & \cdots & \cdots & u_i^{0} & u_i^{N-1} & \cdots \\ \vdots & \vdots & & & & \vdots & \vdots \\ u_i^{N-1} & u_i^{N-2} & \cdots & \cdots & \cdots & u_i^{0} & u_i^{N-1} \\ 0 & u_i^{N-1} & u_i^{N-2} & \cdots & \cdots & u_i^{1} & u_i^{0} \\ \vdots & & & & & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & u_i^{N-1} \\ \vdots & & & & & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \end{bmatrix} \tag{A.2}$$

and

$$\left(\mathbf{U}_i^{\Delta}\right)_R = \begin{bmatrix} 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ u_i^{N-\nu} & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ u_i^{N-2} & \cdots & u_i^{N-\nu} & 0 & \cdots & 0 \\ u_i^{N-1} & \cdots & u_i^{N-\nu-1} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ u_i^{N-2} & \cdots & u_i^{0} & u_i^{N-1} & \cdots & u_i^{N-\nu} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & u_i^{N-1} & \cdots & u_i^{\Delta-M-1} \end{bmatrix} \tag{A.3}$$

$$\mathbf{U}_{i-1}^{\Delta} = \begin{bmatrix} 0 & \cdots & 0 & u_{i-1}^{N-1} & u_{i-1}^{N-2} & \cdots & u_{i-1}^{\nu+\Delta-M+1} \\ \vdots & & & & & & \vdots \\ 0 & \cdots & \cdots & 0 & u_{i-1}^{N-1} & \cdots & u_{i-1}^{N+\nu-M+1} \\ \vdots & & & & & & \vdots \\ 0 & \cdots & \cdots & 0 & \cdots & 0 & u_{i-1}^{N-1} \\ 0 & \cdots\cdots & 0 & 0 & \cdots & \cdots & 0 \\ \vdots & & & & & & \vdots \\ 0 & \cdots\cdots & 0 & 0 & \cdots & \cdots & 0 \end{bmatrix} \tag{A.4}$$

$$\mathbf{U}_{i+1}^{\Delta} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \cdots & 0 & 0 & 0 \\ \vdots & & & & & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ u_{i+1}^{N-\nu} & 0 & \cdots & \cdots & 0 & \cdots & \cdots \\ \vdots & & & & & & \vdots \\ u_{i+1}^{0} & u_{i+1}^{N-1} & \cdots & u_{i+1}^{N-\nu} & 0 & \cdots & 0 \\ \vdots & & & & & & \vdots \\ u_{i+1}^{\Delta-\nu-1} & \cdots & u_{i+1}^{0} & u_{i+1}^{N-1} & \cdots & u_{i+1}^{N-\nu} & 0 \end{bmatrix} \tag{A.5}$$

$$\left[\mathbf{U}_{i}^{\Delta}\right]_{\mathrm{circ}} = \begin{bmatrix} u_i^{\Delta} & \cdots & u_i^{0} & u_i^{N-1} & \cdots & u_i^{\Delta-M+2} \\ \vdots & & & & \vdots & \\ u_i^{N-1} & \cdots & u_i^{\Delta} & u_i^{\Delta-1} & \cdots & u_i^{N-M+1} \\ u_i^{0} & \cdots & u_i^{\Delta+1} & u_i^{\Delta} & \cdots & u_i^{N-M+2} \\ \vdots & & & & & \vdots \\ u_i^{\Delta-1} & \cdots & u_i^{0} & u_i^{N-1} & \cdots & u_i^{\Delta-M-1} \end{bmatrix} \tag{A.6}$$

A.2 Definition of transmission channel-dependent matrices; both $(N+M-1)\times M$ matrices

Define

$$\mathbf{H} = \begin{bmatrix} h_0 & 0 & 0 & \cdots & 0 \\ h_1 & h_0 & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ h_{M-1} & \cdots & & \cdots & h_0 \\ \vdots & & & & \vdots \\ h_{N-1} & \cdots & & \cdots & h_{N-1-M} \\ 0 & h_{N-1} & \cdots & \cdots & h_{N-M} \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & h_{N-1} \end{bmatrix} = \begin{bmatrix} (\mathbf{H}_{\mathrm{u}})_{\Delta\times M} \\ (\mathbf{H}_{\mathrm{c}})_{\nu\times M} \\ (\mathbf{H}_{\mathrm{b}})_{(N-\nu-\Delta+M-1)\times M} \end{bmatrix} \tag{A.7}$$

A.3 Definition of noise-dependent matrices: $\mathbf{G}_{\text{AWGN}}$ and $\mathbf{G}_{\text{NEXT}}$; both $(N+M-1)\times M$ matrices

$$\mathbf{G}_{\text{AWGN}} = \begin{bmatrix} n_0^{\text{AWGN}} & n_{-1}^{\text{AWGN}} & \cdots & n_{-M+1}^{\text{AWGN}} \\ n_1^{\text{AWGN}} & n_0^{\text{AWGN}} & \cdots & n_{-M+2}^{\text{AWGN}} \\ \vdots & & & \vdots \\ n_{M-1}^{\text{AWGN}} & \cdots & \cdots & n_0^{\text{AWGN}} \\ \vdots & & & \vdots \\ n_{N-1}^{\text{AWGN}} & \cdots & & n_{N-1-M}^{\text{AWGN}} \\ \vdots & & & \vdots \\ n_{N+M}^{\text{AWGN}} & n_{N+M-1}^{\text{AWGN}} & \cdots & n_{N-1}^{\text{AWGN}} \end{bmatrix} \text{ and } \mathbf{G}_{\text{NEXT}} = \begin{bmatrix} n_0^{\text{NEXT}} & n_{-1}^{\text{NEXT}} & \cdots & n_{-M+1}^{\text{NEXT}} \\ n_1^{\text{NEXT}} & n_0^{\text{NEXT}} & \cdots & n_{-M+2}^{\text{NEXT}} \\ \vdots & & & \vdots \\ n_{M-1}^{\text{NEXT}} & \cdots & \cdots & n_0^{\text{NEXT}} \\ \vdots & & & \vdots \\ n_{N-1}^{\text{NEXT}} & \cdots & & n_{N-1-M}^{\text{NEXT}} \\ \vdots & & & \vdots \\ n_{N+M}^{\text{NEXT}} & n_{N+M-1}^{\text{NEXT}} & \cdots & n_{N-1}^{\text{NEXT}} \end{bmatrix} \tag{A.8}$$

A.4 Definition of FFT-related matrices: $\mathbf{Q}_k^{\text{noise}}$, $\mathbf{V}_k$, $\mathbf{W}_k$ and $\mathbf{Q}_k^{\text{circ}}$

Matrix $\mathbf{Q}_k^{\text{noise}}$ is defined as

$$\mathbf{Q}_k^{\text{noise}} = \begin{bmatrix} 0 & 0 & \cdots & 0 & q_k^0 & \cdots & q_k^{N-1} \\ 0 & 0 & \cdots & q_k^0 & \cdots & q_k^{N-1} & 0 \\ \vdots & & & & & & \vdots \\ q_k^0 & \cdots & q_k^{N-1} & 0 & \cdots & 0 & 0 \end{bmatrix} \tag{A.9}$$

where $q_k^{(\cdot)}$ are members of the vector $\mathbf{q}_k$ (3.2) and $\mathbf{Q}_k^{\text{noise}}$ is $M \times (N+M-1)$ matrix.

$\mathbf{V}_k$ is an upper diagonal matrix defined as

$$\mathbf{V}_k = \begin{bmatrix} q_k^{N-\Delta} & q_k^{N-\Delta+1} & \cdots & q_k^{N-2} & q_k^{N-1} \\ q_k^{N-\Delta+1} & q_k^{N-\Delta} & \cdots & q_k^{N-1} & 0 \\ \vdots & & & \vdots & \vdots \\ q_k^{N-1} & 0 & \cdots & 0 & 0 \end{bmatrix}_{\Delta\times\Delta} \tag{A.10}$$

Also define $\mathbf{V}_k = \left[\mathbf{V}^{\text{u}} | \mathbf{V}^{\text{b}}\right]$ where $\mathbf{V}^{\text{u}}$ is a $\Delta - \nu \times 1$ matrix and $\mathbf{V}^{\text{b}}$ is $\nu \times 1$ matrix.

$\mathbf{W}_k$ is a lower diagonal $(N-\nu-\Delta+M-1)\times(N-\nu-\Delta+M-1)$ matrix defined as

$$\mathbf{W}_k = \begin{bmatrix} 0 & 0 & \cdots & 0 & q_k^0 \\ 0 & 0 & \cdots & q_k^0 & q_k^1 \\ \vdots & & & \vdots & \vdots \\ q_k^0 & q_k^1 & \cdots & \cdots & q_k^{N-\nu-\Delta+M-2} \end{bmatrix} \tag{A.11}$$

Define $ND = N - \Delta$, $MD = M + \Delta$ and $N1 = N - 1$ for notational convenience. Then define

$$\mathbf{Q}_k^{\text{circ}} = \begin{bmatrix} q_k^{ND+1} & \cdots & \cdots & q_k^{N1} & q_0 & \cdots & q_k^{ND} \\ q_k^{ND+2} & \cdots & q_k^{N1} & q_0 & \cdots & \cdots & q_k^{ND+1} \\ \vdots & & & & & & \vdots \\ q_k^0 & q_k^1 & \cdots & \cdots & \cdots & \cdots & q_k^{N1} \\ \vdots & & & & & & \vdots \\ q_k^{MD-3} & q_k^{MD-2} & \cdots & q_0 & \cdots & \cdots & q_k^{N-MD-2} \end{bmatrix} \tag{A.12}$$

as a $N + M - 1 \times N$ matrix.

A.5 Derivation of the signal matrix $\tilde{\mathbf{A}}_k$

$$\begin{aligned} E\left[\left(Y_{\text{D}}^k\right)^{\text{H}} Y_{\text{D}}^k\right] &= E\left[\mathbf{w}^{\text{T}}\mathbf{H}^{\text{T}}\left[\mathbf{U}_i^{\Delta}\right]_{\text{circ}}^{\text{T}} \mathbf{q}_k \mathbf{q}_k^{\text{H}}\left[\mathbf{U}_i^{\Delta}\right]_{\text{circ}} \mathbf{H}\mathbf{w}\right] \\ &= \mathbf{w}^{\text{T}}\mathbf{H}^{\text{T}}\mathbf{Q}_k^{\text{circ}} E\left[\mathbf{u}_i\mathbf{u}_i^{\text{T}}\right]\left[\mathbf{Q}_k^{\text{circ}}\right]^{\text{H}}\mathbf{H}\mathbf{w} \\ &= \sigma_s^2 \mathbf{w}^{\text{T}}\mathbf{H}^{\text{T}}\mathbf{Q}_k^{\text{circ}}\left[\mathbf{Q}_k^{\text{circ}}\right]^{\text{H}}\mathbf{H}\mathbf{w} \\ &= \mathbf{w}^{\text{T}}\tilde{\mathbf{A}}_k\mathbf{w} \end{aligned} \tag{A.13}$$

Note that if the power allocated to subchannels is not equal in all subchannels we would need to include a diagonal matrix that would contain these different powers on the diagonal instead of just σ_s^2.

A.6 Derivation of the noise matrix $\tilde{\mathbf{B}}_k$

Noise is defined using the difference between the received and desired shapes in the frequency domain for every subchannel.

$$Y_{\text{R}}^k - Y_{\text{D}}^k = \mathbf{q}_k^{\text{H}}\left(\mathbf{U}_{i-1}^{\Delta} + \mathbf{U}_{i+1}^{\Delta} + \underbrace{\mathbf{U}_i^{\Delta} - [\mathbf{U}_i^{\Delta}]_{\text{circ}}}_{\mathbf{P}}\mathbf{H} + \mathbf{G}_{\text{AWGN}} + \mathbf{G}_{\text{NEXT}}\right)\mathbf{w} + D_k \tag{A.14}$$

In order to easily deal with matrix $\mathbf{P}$, we are going to break it into two matrices containing the non-zero entries. Define

$$\mathbf{P} = \left[\mathbf{U}_{i,\text{L}}^{\Delta}\right] + \left[\mathbf{U}_{i,\text{U}}^{\Delta}\right] \tag{A.15}$$

where

$$\left[\mathbf{U}_{i,\mathrm{L}}^{\Delta}\right] = \begin{bmatrix} 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & & & & & & \vdots \\ u_i^0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ u_i^1 & u_i^0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ u_i^{\Delta-1} & u_i^{\Delta-2} & \cdots & u_i^0 & 0 & \cdots & 0 \end{bmatrix} \left[\mathbf{U}_{i,\mathrm{U}}^{\Delta}\right] = \begin{bmatrix} 0 & \cdots & 0 & u_i^{NV1} & u_i^{NV1-1} & \cdots & u_i^{DM+2} \\ 0 & \cdots & 0 & 0 & u_i^{NV1} & \cdots & u_i^{DM+3} \\ \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & 0 & u_i^{NV1} \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \tag{A.16}$$

where we define $NV1 = N - \nu - 1$ and $DM = \Delta - M$. Notice that non-zero entries come about because of the contributions of DMT symbol $i-1$ and $i+1$ present in the matrix $\left[\mathbf{U}_{\mathrm{ISI}}^{\Delta}\right]$.

For notational convenience define $Z_k = Y_{\mathrm{R}}^k - Y_{\mathrm{D}}^k$. Then the power in Z_k is

$$\begin{aligned} EP &= E\left[Z_k^{\mathrm{H}} Z_k\right] = \mathbf{w}^{\mathrm{T}} E\left[\mathbf{S}\right]\mathbf{w} + \mathbf{w}^{\mathrm{T}} E\left[\underbrace{\mathbf{G}_{\mathrm{AWGN}}^{\mathrm{T}}\mathbf{q}_k\mathbf{q}_k^{\mathrm{H}}\mathbf{G}_{\mathrm{AWGN}}}_{\mathbf{EP1}}\right]\mathbf{w} + \mathbf{w}^{\mathrm{T}} E\left[\underbrace{\mathbf{G}_{\mathrm{NEXT}}^{\mathrm{T}}\mathbf{q}_k\mathbf{q}_k^{\mathrm{H}}\mathbf{G}_{\mathrm{NEXT}}}_{\mathbf{EP2}}\right]\mathbf{w} \\ &+ E\left[D_k^{\mathrm{H}} D_k\right] \end{aligned} \tag{A.17}$$

Where

$$\mathbf{S} = \underbrace{\mathbf{H}^{\mathrm{T}}\left[\mathbf{U}_{i-1}^{\Delta}\right]^{\mathrm{H}}\mathbf{q}_k\mathbf{q}_k^{\mathrm{H}}\mathbf{U}_{i-1}^{\Delta}\mathbf{H}}_{\mathbf{S1}} + \underbrace{\mathbf{H}^{\mathrm{T}}\left[\mathbf{U}_{i+1}^{\Delta}\right]^{\mathrm{H}}\mathbf{q}_k\mathbf{q}_k^{\mathrm{H}}\mathbf{U}_{i+1}^{\Delta}\mathbf{H}}_{\mathbf{S2}} + \underbrace{\mathbf{H}^{\mathrm{T}}\mathbf{P}^{\mathrm{H}}\mathbf{q}_k\mathbf{q}_k^{\mathrm{H}}\mathbf{P}\mathbf{H}}_{\mathbf{S3}} \tag{A.18}$$

Notice that now we can transform the parts of $(A.18)$ according to definitions in (A.10) as follows.

$$\left[\mathbf{U}_{i+1}^{\Delta}\right]^{\mathrm{H}}\mathbf{q}_k = \begin{bmatrix} \mathbf{V}_k^{\mathrm{b}} & \mathbf{0} & \mathbf{V}_k^{\mathrm{u}} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}_{(N+M-1)\times N} \mathbf{u}_{i+1} \text{ and } \left[\mathbf{U}_{\mathrm{i-1}}^{\Delta}\right]^{\mathrm{H}}\mathbf{q}_{\mathrm{k}} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_k \end{bmatrix}_{(\mathrm{N+M-1})\times \mathrm{N}} \mathbf{u}_{\mathrm{i-1}} \tag{A.19}$$

where $\mathbf{W}_k$ has been defined in (A.11). Then

$$\left[\mathbf{U}_{i,\mathrm{L}}^{\Delta}\right]^{\mathrm{H}}\mathbf{q}_k = \begin{bmatrix} \mathbf{V}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}_{(N+M-1)\times N} \mathbf{u}_i \text{ and } \left[\mathbf{U}_{\mathrm{i,U}}^{\Delta}\right]^{\mathrm{H}}\mathbf{q}_{\mathrm{k}} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_k & \mathbf{0} \end{bmatrix}_{(\mathrm{N+M-1})\times \mathrm{N}} \mathbf{u}_{\mathrm{i}} \tag{A.20}$$

Now we can go forward with the derivation in

$$\begin{aligned} E\left[\mathbf{S2}\right] &= \mathbf{H}^{\mathrm{T}}\begin{bmatrix} \mathbf{V}_k^{\mathrm{b}} & \mathbf{0} & \mathbf{V}_k^{\mathrm{u}} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} E[\mathbf{u}_{i+1}\mathbf{u}_{i+1}^{\mathrm{T}}] \begin{bmatrix} \left[\mathbf{V}_k^{\mathrm{b}}\right]^{\mathrm{H}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \\ \left[\mathbf{V}_k^{\mathrm{u}}\right]^{\mathrm{H}} & \mathbf{0} \end{bmatrix} \mathbf{H} \\ &= \sigma_s^2 \mathbf{H}^{\mathrm{T}} \begin{bmatrix} \mathbf{V}_k^{\mathrm{b}}\left[\mathbf{V}_k^{\mathrm{b}}\right]^{\mathrm{H}} + \mathbf{V}_k^{\mathrm{u}}\left[\mathbf{V}_k^{\mathrm{u}}\right]^{\mathrm{H}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{H} \\ &= \sigma_s^2 \mathbf{H}_{\mathrm{u}}^{\mathrm{T}} \mathbf{V}_k \mathbf{V}_k^{\mathrm{H}} \mathbf{H}_{\mathrm{u}} \end{aligned} \tag{A.21}$$

where $\mathbf{H}_u$ was defined in $(A.7)$. Similarly,

$$\begin{aligned} E\left[\mathbf{S1}\right] &= \mathbf{H}^{\mathrm{T}} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_k \end{bmatrix} E[\mathbf{u}_{i-1}\mathbf{u}_{i-1}^{\mathrm{T}}] \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_k^{\mathrm{H}} \end{bmatrix} \mathbf{H} \\ &= \sigma_s^2 \mathbf{H}^{\mathrm{T}} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_k \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_k^{\mathrm{H}} \end{bmatrix} \mathbf{H} \\ &= \sigma_s^2 \mathbf{H}_{\mathrm{b}}^{\mathrm{T}} \mathbf{W}_k \mathbf{W}_k^{\mathrm{H}} \mathbf{H}_{\mathrm{b}} \end{aligned} \tag{A.22}$$

where $\mathbf{H}_{\mathrm{b}}$ was defined in (A.7). Following a similar procedure, we arrive at the contribution of $\mathbf{w}^{\mathrm{T}} E[\mathbf{S3}]\mathbf{w}$, which is exactly the sum of (A.21) and (A.22). The contribution of the AWGN is given by

$$E\left[\mathbf{EP1}\right] = \mathbf{Q}_k^{\mathrm{noise}} E[\mathbf{n}^{\mathrm{AWGN}} \left(\mathbf{n}^{\mathrm{AWGN}}\right)^{\mathrm{T}}] \left[\mathbf{Q}_k^{\mathrm{noise}}\right]^{\mathrm{H}} = \sigma_{\mathrm{AWGN}}^2 \mathbf{Q}_k^{\mathrm{noise}} \left[\mathbf{Q}_k^{\mathrm{noise}}\right]^{\mathrm{H}} \tag{A.23}$$

while the contribution of the NEXT is given by

$$E\left[\mathbf{EP2}\right] = \mathbf{Q}_k^{\mathrm{noise}} E[\mathbf{n}^{\mathrm{NEXT}} \left(\mathbf{n}^{\mathrm{NEXT}}\right)^{\mathrm{T}}] \left[\mathbf{Q}_k^{\mathrm{noise}}\right]^{\mathrm{H}} = \mathbf{Q}_k^{\mathrm{noise}} \Sigma_{\mathrm{NEXT}} \left[\mathbf{Q}_k^{\mathrm{noise}}\right]^{\mathrm{H}} \tag{A.24}$$

where the NEXT vector is defined as $\mathbf{n}^{\mathrm{NEXT}} = \{n_{-M+1}^{\mathrm{NEXT}}, n_{-M+2}^{\mathrm{NEXT}}, \cdots, n_0^{\mathrm{NEXT}}, \cdots, n_{N-1}^{\mathrm{NEXT}}\}^{\mathrm{T}}$ and similarly, AWGN $\mathbf{n}^{\mathrm{AWGN}} = \{n_{-M+1}^{\mathrm{AWGN}}, n_{-M+2}^{\mathrm{AWGN}}, \cdots, n_0^{\mathrm{AWGN}}, \cdots, n_{N-1}^{\mathrm{AWGN}}\}^{\mathrm{T}}$

Then

$$\begin{aligned} E\left[Z_k^{\mathrm{H}} Z_k\right] &= \mathbf{w}^{\mathrm{T}} 2\sigma_s^2 \left(\mathbf{H}_{\mathrm{u}}^{\mathrm{T}} \mathbf{V}_k \mathbf{V}_k^{\mathrm{H}} \mathbf{H}_{\mathrm{u}} + \mathbf{H}_{\mathrm{b}}^{\mathrm{T}} \mathbf{W}_k \mathbf{W}_k^{\mathrm{H}} \mathbf{H}_{\mathrm{b}}\right) \mathbf{w} + \mathbf{w}^{\mathrm{T}} \sigma_{\mathrm{AWGN}}^2 \mathbf{Q}_k^{\mathrm{noise}} \left[\mathbf{Q}_k^{\mathrm{noise}}\right]^{\mathrm{H}} \mathbf{w} \\ &+ \mathbf{w}^{\mathrm{T}} \mathbf{Q}_k^{\mathrm{noise}} \Sigma_{\mathrm{NEXT}} \left[\mathbf{Q}_k^{\mathrm{noise}}\right]^{\mathrm{H}} \mathbf{w} + \mathbf{w}^{\mathrm{T}} \frac{\sigma_{\mathrm{DNF}}^2}{\mathbf{w}^{\mathrm{T}}\mathbf{w}} \mathbf{I} \mathbf{w} \\ &= \mathbf{w}^{\mathrm{T}} \tilde{\mathbf{B}}_k \mathbf{w} \end{aligned} \tag{A.25}$$

where $\mathbf{I}$ is the $M \times M$ identity matrix. Observe that over the constraint set $\mathcal{S}$, $\mathbf{w}^{\mathrm{T}}\mathbf{w} = 1$, so that $\tilde{\mathbf{B}}_k$ becomes independent of TEQ FIR taps $\mathbf{w}$.

B Background Information

B.1 Vita

Milos Milosevic received B.S.E.E. and M.S.E.E. at the Illinois Institute of Technology in Chicago, IL in 1996 and 1998, respectively. He is pursuing his Ph.D. in electrical engineering at The University of Texas at Austin, TX. He is a senior engineer in the Acquisition Department, Telemetry Section of Schlumberger in Austin, TX, where he designs next generation high-speed communication systems for oilfield services. His professional experiences include a systems engineering position in Motorola's Networking and Communication Systems Division where he worked on the design of a ADSL transceiver based on Motorola 56300 core and DSP algorithm design for the Texas Instruments C6000 DSP core at Texas Instruments Broadband Division. His current research interests are multicarrier communication systems such as ADSL or VDSL.

B.2 List of Publications

B.2.1 Refereed Journal Papers

M. Milosevic , L. F. C. Pessoa, B. L. Evans, and R. Baldick, " Optimal Time Domain Equalization Design for Maximizing Data Rate of Discrete Multi-Tone Systems," Submitted on Jan. 18, 2002, to the IEEE Trans. on Sig. Proc.

M. Milosevic , T. Inoue, P. Molnar, and B. L. Evans, " Fast Unbiased Echo Canceller Update During ADSL Transmission," accepted for publication IEEE Transactions on Communications.

M. N. Wernick, E. J. Infusino, and M. Milosevic, "Fast spatio-temporal image reconstruction for dynamic PET," IEEE Transactions on Medical Imaging, vol. 18, pp. 185-195, March 1999.

B.2.2 Refereed Conference Papers

M. Milosevic , L. F. C. Pessoa, B. L. Evans, and R. Baldick, " DMT Bit Rate Maximization With Optimal Time Domain Equalizer Architecture," Invited Paper, In Progress,IEEE Asilomar Conf. on Signals, Systems, and Computers, Pacific Grove, CA, Nov. 2002.

M. Milosevic, W. Schwartzkopf, T. E. Milner, B. L. Evans, and A. C. Bovik, Low-Complexity Velocity Estimation in High-Speed Optical Doppler Tomography Systems, IEEE Intl. Conf. Image Proc., Tokyo, Japan, Nov. 1999.

B.2.3 Other Publications

M. Milosevic, P. Molnar, T. Inoue, and M. Pendleton, Method and Apparatus for Echo Cancellation Updates In a Multicarrier Transceiver System, U. S. Patent pending, March 2000.

M. Milosevic, Dictionary of Professional Terms used in Audio, Video, Telecommunications and Multimedia, 1st Ed, Next Millennium, Nis, Yugoslavia, 1999.

M. Milosevic, "Tomographic Image Sequence Reconstruction Using Motion Compensation," Master's Thesis, Illinois Institute of Technology, Chicago, IL, May 1998.

B.3 Courses Taken

a. Complete List of Graduate Courses in Chronological Order (including M.S.E.E.)				
Semester	**Course**	**Title**	**Instructor**	**Grade**
Spring 1997	ECE569	Digital Signal Processing II	Dr. Nicholas P. Galatsanos	A
Spring 1997	ECE570	Fiber Optic Communication Systems	Dr. David B. Patterson	A
Spring 1997	ECE404	Communication Systems II	Dr. Joseph L. LoCicero	A
Fall 1997	ECE567	Statistical Signal Processing	Dr. Nicholas P. Galatsanos	A
Fall 1997	ECE508	Signal and Data Compression	Dr. W. Y. Geoffrey Chan	A
Fall 1997	ECE513	Communication Eng. Fundamentals	Dr. Guillermo E. Atkin	A
Spring 1998	ECE449	Object Oriented Programming	Mr. Christopher Hield	A
Spring 1998	ECE597	Multivariate Statistical Analysis	Dr. Miles N. Wernick	A
Fall 1998	EE 381K	14-Multidimensional Digital Signal Processing	Dr. Brian L. Evans	A
Spring 1999	EE 381K	2-Digital Communications	Dr. Sayfe Kiaei	A
Summer 1999	EE N382M	2-System Design Metrics	Dr. Tony Ambler	A
Fall 1999	EE 381J	Probability & Stochastic Processes I	Dr. Aristotle Arapostathis	B
Spring 2000	EE 380N	11-Optimization In Eng. Systems	Dr. Ross Baldick	B
Summer 2000	M N365C	Real Analysis I	Dr. William T. Eaton	A
Fall 2000	M 383E	Numerical Analysis: Linear Algebra	Dr. Robert A. van de Geijn	A
Spring 2001	EE 381K	7-Information Theory	Dr. Takis Konstantopoulos	A
Fall 2001	M 391C	Wavelets Theory and Applications	Dr. John E. Gilbert	A
b.1 Graduate ECE Courses in TISE area				
Spring 1997	ECE569	Digital Signal Processing II	Dr. Nicholas P. Galatsanos	A
Spring 1997	ECE570	Fiber Optic Communication Systems	Dr. David B. Patterson	A
Spring 1997	ECE404	Communication Systems II	Dr. Joseph L. LoCicero	A
Fall 1997	ECE567	Statistical Signal Processing	Dr. Nicholas P. Galatsanos	A
Fall 1997	ECE508	Signal and Data Compression	Dr. W. Y. Geoffrey Chan	A
Fall 1997	ECE513	Communication Eng. Fundamentals	Dr. Guillermo E. Atkin	A
Fall 1998	EE 381K	14-Multidimensional Digital Signal Processing	Dr. Brian L. Evans	A
Spring 1999	EE 381K	2-Digital Communications	Dr. Sayfe Kiaei	A
Fall 1999	EE 381J	Probability & Stochastic Processes I	Dr. Aristotle Arapostathis	B
Spring 2000	EE 380N	11-Optimization In Eng. Systems	Dr. Ross Baldick	B
Spring 2001	EE 381K	7-Information Theory	Dr. Takis Konstantopoulos	A
b.2 Graduate Non-TISE EE Courses				
Spring 1998	ECE449	Object Oriented Programming	Mr. Christopher Hield	A
Spring 1998	ECE597	Multivariate Statistical Analysis	Dr. Miles N. Wernick	A
Summer 1999	EE N382M	2-System Design Metrics	Dr. Tony Ambler	A
b.3 Graduate Non-EE Courses				
Summer 2000	M N365C	Real Analysis I	Dr. William T. Eaton	A
Fall 2000	M 383E	Numerical Analysis: Linear Algebra	Dr. Robert A. van de Geijn	A
Fall 2001	M 391C	Wavelets Theory and Applications	Dr. John E. Gilbert	A