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

Miloš Milošević, *Student Member, IEEE*, Lúcio F. C. Pessoa, *Senior Member, IEEE*, Brian L. Evans, *Senior Member, IEEE* and Ross Baldick, *Member, IEEE*

Abstract

The traditional discrete multi-tone equalizer is a cascade of a time domain equalizer (TEQ) as a single finite impulse response filter, a multicarrier demodulator as a fast Fourier transform (FFT), 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 paper, we propose a *time domain* TEQ filter bank (TEQFB) and single TEQ 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. The contributions of this paper are: (1) a model for the signal-to-noise ratio (SNR) at the FFT output that includes ISI, near-end crosstalk, white Gaussian noise, analog-to-digital converter quantization noise and the digital noise floor; (2) data rate optimal time domain per-tone TEQ filter bank and the upper bound on bit rate performance it achieves; and (3) data rate maximization single TEQ design algorithm.

SP EDICS: *SP 3-TDSL: Telephone Networks and Digital Subscriber Loops.*

Miloš Milošević *, B. L. Evans and R. Baldick are with the Dept. of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712-1084, L. F. C. Pessoa is with Motorola, Inc.,7700 West Parmer Lane, MD: TX32-PL30, Austin, TX 78729. E-mail: {milos,bevans,baldick}@ece.utexas.edu and Lucio.Pessoa@motorola.com. B. L. Evans was supported by a gift from the Motorola Semiconductor Products Sector, Austin, Texas, and The State of Texas Advanced Technology Program under grant 003658-0614-2001.

I. 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) [1], and internationally by the International Telecommunications Union G.DMT (G.992.1) [2] and G.Lite (G.992.2) [3] ADSL standards. DMT is figuring in the very-high-rate DSL (VDSL) standard proposals [4], [5], [6]. 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.

Fig. 1 shows a simplified block diagram of a DMT system. Data transmission is organized into groups of N samples called symbols. The input bit stream on the transmitter is mapped using Quadrature Amplitude Modulation (QAM) into a complex-valued vector $\mathbf{X}^i$ at time i. Each entry of $\mathbf{X}^i$ is modulated by the IFFT block into a different frequency band (subchannel) with the carrier frequency lying in the center of the band.

Each subchannel is nearly independent of the other subchannels and the degree of independence increases with the number of subchannels [7]. 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 guard period of ν samples is added before digital-to-analog (D/A) conversion and transmission. A spectrally shaped channel impulse response (CIR) longer than $\nu + 1$ 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 would 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 would 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 of 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 CP is removed after the analog-to-digital (ADC) conversion.

Fig. 1 shows two scenarios for demodulation and equalization. The equalization in ADSL is based on a two-step process. Traditionally, in the first equalization step, a finite impulse response filter (FIR) block shown in Fig. 1 in a), 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 shortened channel impulse response that has an extent smaller than or equal to $\nu + 1$ samples. Then, linear convolution of the shortened channel impulse response and a DMT symbol is converted into their circular convolution by the virtue of having repeated DMT symbol samples present in the CP [1]. Demodulation, which is performed using an FFT, then transforms this circular convolution

into a multiplication of complex-valued sequences in the frequency domain [8]. In the second equalization step, division by the frequency domain response of the shortened channel impulse response (known as frequency domain equalization or FEQ), fully equalizes the signal. This fully equalized signal is then decoded using a QAM decoder resulting in an estimate of the transmitted complex-valued symbol $\mathbf{Y}^i$. The estimated symbol $\mathbf{Y}^i$ is then mapped into a corresponding binary sequence in the parallel-to-serial conversion.

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-squared error (MMSE) is one of the earliest, widely used, and mathematically tractable solutions [9]. The equalizer taps are designed to minimize the mean-squared of the residual error between the shortened channel impulse response and a desired target impulse response of length $\nu + 1$ samples during training. In [9] the authors discuss TEQ solutions under two constraints on the resulting equalized channel: unit-tap constraint (UTC) and unit-energy constraint (UEC). The TEQ taps obtained by using the MMSE 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 shortened channel impulse response, which in the presence of a digital noise floor (DNF) renders some of the subchannels unable to carry data [10], [11].

The Shortening Signal-to-Noise Ratio (SSNR) method [12] attempts to maximize the ratio of the energy present in the target shortened channel impulse response 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 [12].

Maximum Bit Rate (MBR) and Minimum ISI (Min-ISI) methods [10] are based on a similar idea as in the SSNR method in [12], which is used to define a new model for the SNR. In the MBR and Min-ISI methods, the SNR model includes ISI and additive white Gaussian noise (AWGN) and is used to determine the bit rate. MBR tries to maximize the nonlinear bit rate equation by using the Broyden-Fletcher-Goldfarb-Shanno quasi-Newton algorithm, which is available in Matlab's optimization toolbox. The authors in [10] conclude that the MBR procedure is computationally expensive and therefore not well suited for real-time implementation on a programmable digital signal processor. Nevertheless, the procedure does maximize the bit rate for the traditional receiver architecture at the TEQ output with respect to the SNR model employed. The authors in [10] also define the Min-ISI method, which minimizes the sum of noise power present in the new SNR definition with the constraint that the power of the signal is constant. In simulation, the Min-ISI method achieved more than 95% of the bit rate of the MBR method [10]. The Min-ISI method has been implemented on several fixed-point programmable digital signal processors [10].

An alternate receiver architecture is proposed in [13], [14] and shown in Fig. 1 in the receiver in b). Since the traditional equalizer equalizes all subchannels "in a combined fashion," [13] which may limit equalization

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

performance, the authors of [13] 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. The optimality of the per-tone equalizer (PTE) structure was shown in [13]. This 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 [13], the authors also propose various groupings of subchannels, in which each group has a complex-valued equalizing filter assigned to it.

In this paper, we develop a subchannel Signal-to-Noise Ratio (SNR) model where the desired signal is formed as the circularly convolved data symbol and the channel impulse response at the input of the FFT, and noise is the difference between the received and the desired signal. The received signal also includes the near-end crosstalk (NEXT), additive white Gaussian noise, analog-to-digital converter (ADC) quantization noise and the digital noise floor (DNF) due to finite precision arithmetic. Following the approach we published in [15], [16] we arrive at the optimal time domain per-tone TEQ filter bank (TEQFB) and compute the FFT by using Goertzel filters [17], [18] at the receiver during data transmission. TEQFB architecture equalizes the signal on a per-tone basis in the time-domain as opposed to the PTE which does the per-tone equalization in the frequency domain. Nonetheless, both architectures equalize on a per-tone basis and can be seen from that point of view as a dual of each other. We also propose an iterative fractional programming algorithm to initialize the single FIR coefficients used in the traditional TEQ architecture that achieves on average more than 99% of the performance of TEQFB for the tested standard ADSL carrier serving area loops. We propose a method of assessing performance of TEQ based on the performance of TEQFB. We present simulation results showing that TEQFB achieves higher data rate than UTC and UEC MMSE, MBR, Min-ISI and least-squares (LS) PTE methods [2]. The proposed single FIR TEQ design achieves higher data rate in our simulations than UTC and UEC MMSE, MBR and Min-ISI, while it closely matches LS PTE data rate. The contributions of this paper are: (1) a model for the signal-to-noise ratio (SNR) at the FFT output that includes ISI, near-end crosstalk, white Gaussian noise, analog-to-digital converter quantization noise and the digital noise floor; (2) data rate optimal time domain per-tone TEQ filter bank and the upper bound on bit rate performance it achieves; and (3) data rate maximization single TEQ design algorithm.

This paper is organized as follows. Section II summarizes background information on the achievable data rate of a DMT system. Section III presents the proposed SNR model. Section IV derives the proposed method for design of the TEQFB and a proposed method for the near-optimum design of a single TEQ and discusses their computational complexity. Both of the proposed methods are compliant with the various ADSL standards. Section V presents simulation results. Section VI concludes this paper.

[2] See Section V for the explanation of PTE initialization used in this paper.

II. 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 of transmission as [19]

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

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 $\mathrm{SNR} = P_{\mathrm{s}}/P_{\mathrm{n}}$. Practical coding/modulation methods cannot achieve the rate given in (1). The difference between the rate in (1) and the best achievable rate in practice can be characterized by a quantity called the SNR gap (Γ) in dB [7], [20]. Γ is a function of the modulation method and the target probability of bit error per dimension, P_{e}. For a coded quadrature amplitude modulation system (QAM),

$$\Gamma = 9.8\mathrm{dB} + \gamma_{\mathrm{m}} - \gamma_{\mathrm{c}} \tag{2}$$

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 [7]. The data rate in bits per symbol in the k^{th} subchannel becomes

$$b_k = \log_2\left(1+\frac{\mathrm{SNR}_k}{\Gamma_k}\right) \tag{3}$$

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 $\Gamma_k = \Gamma\ \forall k$.

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

$$\mathcal{I} \subseteq \{0,\ 1, \cdots,\ N/2-1\} \tag{4}$$

Then, the number of bits per DMT symbol that can be reliably transmitted for the target BER is given by

$$b_{\mathrm{DMT}}(\mathcal{I}) = \sum_{k\in\mathcal{I}} \log_2\left(1+\frac{\mathrm{SNR}_k}{\Gamma}\right) \tag{5}$$

Equation (5) could result in non-integer bit values, but ADSL and VDSL standards allow only integer bit loading[3]. Thus, the number of bits that would be actually used in a DMT system is

$$b^{\mathrm{int}}_{\mathrm{DMT}}(\mathcal{I}) = \sum_{k\in\mathcal{I}} \left\lfloor \log_2\left(1+\frac{\mathrm{SNR}_k}{\Gamma}\right)\right\rfloor \tag{6}$$

[3]A 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].

where $\lfloor . \rfloor$ means the closest smaller integer. The primary objective of this paper is to model (6) 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 (6) mathematically difficult to handle, and we would like to deal instead with (5). In general, the maximizer of (5) may not maximize (6). We explore this issue in Section III-A and show empirical evidence that the solutions of the two maximizers are often indistinguishable in the problem considered in this paper. Nonetheless, both functions depend on the SNR_k quantities, as the value of SNR gap Γ is fixed for the particular modulation method. Equation (5) is a monotonically increasing and (6) is a monotonically non-decreasing function of SNR_k. In the next section, we develop a new model of the SNR for each subchannel that models the dependency on the TEQ properties and the improvements that we can make by designing the TEQ taps to maximize (5).

III. Proposed SNR model

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, respectively. 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.

In the following discussion, only major ideas are discussed and specific details are presented in the Appendix. Let ν be the length of the CP and $\mathbf{h} = [h_0, h_1, \cdots, h_{N-1}]^{\mathrm{T}}$ be the $N \times 1$ channel impulse response. G.DMT provides for symbol-length (N samples) identification of the channel impulse response at a sampling rate of $f_{\mathrm{s}} = 2.208$ MHz which amounts to roughly 232 μs for $N = 512$ samples. The symbol rate f_{sym} is 4 kHz.

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 receiver FFT block and let

$$\mathbf{U}_{\mathrm{ISI}}^{\Delta} = \mathbf{U}_{i-1}^{\Delta} + \mathbf{U}_{i}^{\Delta} + \mathbf{U}_{i+1}^{\Delta} \tag{7}$$

be the convolutional matrix of the DMT symbols $i-1$, i and $i+1$. Let $\mathbf{H}$ be the convolution matrix of the CIR and TEQ (see Appendix B). We define the vector

$$\mathbf{q}_k = \left[1,\ e^{j2\pi k/N}, \cdots,\ e^{j2\pi(N-1)k/N}\right]^{\mathrm{T}} \tag{8}$$

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.

Let $\mathbf{G}_{\mathrm{AWGN}}$ and $\mathbf{G}_{\mathrm{NEXT}}$ be the AWGN and NEXT convolution matrices defined in Appendix C, respectively, based on the convolution of AWGN and NEXT noise with the TEQ impulse response. The samples of AWGN are i.i.d. according to $\mathcal{N}(0, \sigma_{\mathrm{AWGN}}^2)$ where σ_{AWGN}^2 is the noise variance, which is measured by the power of the noise

with respect to 100Ω resistance. Samples of the NEXT have spectrally shaped power spectral density (PSD) as defined in [1] for various types of interferers and denoted using the Toeplitz variance matrix $\mathbf{\Sigma}_{\mathrm{NEXT}}$.

Let $\overline{\mathbf{w}} = \mathbf{w} + \mathbf{n}_{\mathrm{w}}$, where $\mathbf{w}$ is the TEQ in infinite precision, $\overline{\mathbf{w}}$ is the TEQ in fixed-point representation, and $\mathbf{n_w}$ is the corresponding quantization noise. Let $\overline{\mathbf{y}} = \mathbf{y} + \mathbf{n_y}$, where $\mathbf{y}$ is the received signal before the ADC, $\overline{\mathbf{y}}$ is the received signal after the ADC, and $\mathbf{n_y}$ is the corresponding quantization noise. Starr, Cioffi and Silverman [19, p. 299] recommend using 14-bit ADC, while ADSL transceivers commonly use 16-bit ADCs. Fixed-point DSPs commonly use more than 16-bit fixed-point arithmetic. Gersho and Gray [23, p. 152] point out that in a high resolution case one can approximate the probability density function of the quantization noise to be uniform. After some algebra, the convolution of $\mathbf{w}$ and $\mathbf{y}$ can be written as

$$\overline{\mathbf{w}} * \overline{\mathbf{y}} = \mathbf{w} * \mathbf{y} + \underbrace{\mathbf{w} * \mathbf{n_y}}_{\mathbf{e}_1} + \underbrace{\overline{\mathbf{y}} * \mathbf{n_w}}_{\mathbf{e}_2} \tag{9}$$

where the error terms are the loss incurred by performing operations using digital processing. Here, $\mathbf{e}_1$ depends on the TEQ, while $\mathbf{e}_2$ does not.

Thus, define $\mathbf{G}_{\mathrm{ADC}}$ as the the ADC noise convolution matrix (Appendix C) with the TEQ. ADC noise samples are i.i.d. according to $\mathcal{U}(0, \sigma^2_{\mathrm{ADC}})$, where $\mathcal{U}$ stands for uniform distribution and $\sigma^2_{\mathrm{ADC}} = 2^{-2p}/12$, where p is the number of bits in the ADC.

The $\mathbf{e}_2$ captures the fixed-point noise that we will call DNF and whose value in subchannel k is D_k. DNF does not depend on the TEQ but it does effect the subchannel SNR terms, $\mathrm{SNR}_k(\mathbf{w})$. DNF D_k distribution is Gaussian by Central Limit Theorem with $\mathcal{N}(0, \sigma^2_{\mathrm{DNF}}(k))$ where σ^2_{DNF} is the noise variance [4]. Using these terms, we write the received data point in the k^{th} subchannel as

$$Y^k_{\mathrm{R}}(\mathbf{w}) = \mathbf{q}^{\mathrm{H}}_k \mathbf{U}^{\Delta}_{\mathrm{ISI}} \mathbf{H}\mathbf{w} + \mathbf{q}^{\mathrm{H}}_k \mathbf{G}_{\mathrm{AWGN}}\mathbf{w} + \mathbf{q}^{\mathrm{H}}_k \mathbf{G}_{\mathrm{NEXT}}\mathbf{w} + \mathbf{q}^{\mathrm{H}}_k \mathbf{G}_{\mathrm{ADC}}\mathbf{w} + D_k \tag{10}$$

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

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. We can design the TEQ to process the received samples to optimize these goals. Ideally, after the TEQ, one would prefer to have mimicked a circular convolution of the signal and the channel impulse response using the cyclic prefix. The desired circular convolution of the i^{th} symbol and the CIR in the k^{th} subchannel, after the TEQ and FFT, can be written as

$$Y^k_{\mathrm{D}}(\mathbf{w}) = \mathbf{q}^{\mathrm{H}}_k \left[\mathbf{U}^{\Delta}_i\right]_{\mathrm{circ}} \mathbf{H}\mathbf{w},\ k \in \{0, \cdots, N/2 - 1\} \tag{11}$$

[4] Assuming that the samples of $\mathbf{n_w}$ are i.i.d. according to $\mathcal{U}(0, \sigma^2_{nw})$ and independent of $\overline{\mathbf{y}}$, then $\sigma^2_{\mathrm{DNF}}(k)$ can be approximated as $\sigma^2_{\mathrm{DNF}}(k) = N\sigma^2_{nw}(\sigma^2_s|H(k)|^2 + \sigma^2_{\mathrm{ADC}})$, where $H(k) = \mathbf{q}^{\mathrm{H}}_k \mathbf{h}$ and $\sigma^2_{nw} = 2^{-2r}/12$ for a fixed-point mapping with r bits.

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

$$Y_{\mathrm{R}}^{k}(\mathbf{w}) = Y_{\mathrm{D}}^{k}(\mathbf{w}) + \underbrace{\left(Y_{\mathrm{R}}^{k}(\mathbf{w}) - Y_{\mathrm{D}}^{k}(\mathbf{w})\right)}_{\text{ISI, ICI, AWGN, NEXT, DNF}} \tag{12}$$

We then write $\mathrm{SNR}^{\mathrm{Model}}{}_{k}(\mathbf{w})$ for all k as

$$\mathrm{SNR}_{k}^{\mathrm{Model}}(\mathbf{w}) = \frac{E\left\{\left[Y_{\mathrm{D}}^{k}(\mathbf{w})\right]^{H} Y_{\mathrm{D}}^{k}(\mathbf{w})\right\}}{E\left\{\left[Y_{\mathrm{R}}^{k}(\mathbf{w}) - Y_{\mathrm{D}}^{k}(\mathbf{w})\right]^{\mathrm{H}}\left[Y_{\mathrm{R}}^{k}(\mathbf{w}) - Y_{\mathrm{D}}^{k}(\mathbf{w})\right]\right\}} \tag{13}$$

where $E[\cdot]$ is the statistical expectation operator and $[\cdot]^{\mathrm{Model}}$ stands for model.

We have defined the proposed SNR model 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 desired data. If there were no noise due to ISI and ICI, then the denominator would be reduced to the contributions of the AWGN, NEXT, ADC quantization noise and DNF.

In Appendix E we derive $\tilde{\mathbf{A}}_k$,

$$\tilde{\mathbf{A}}_k = \sigma_s^2 \mathbf{H}^{\mathrm{T}} \mathbf{Q}_k^{\mathrm{circ}} \left[\mathbf{Q}_k^{\mathrm{circ}}\right]^{\mathrm{H}} \mathbf{H} \tag{14}$$

Similarly, in Appendix F we derive $\tilde{\mathbf{B}}_k$,

$$\begin{aligned}
\tilde{\mathbf{B}}_k &= 2\sigma_s^2 \left(\mathbf{H}_u^{\mathrm{T}} \mathbf{V}_k \mathbf{V}_k^{\mathrm{H}} \mathbf{H}_u + \mathbf{H}_{\mathrm{b}}^{\mathrm{T}} \mathbf{W}_k \mathbf{W}_k^{\mathrm{H}} \mathbf{H}_{\mathrm{b}}\right) \\
&+ \sigma_{\mathrm{AWGN}}^2 \mathbf{Q}_k^{\mathrm{noise}} \left[\mathbf{Q}_k^{\mathrm{noise}}\right]^{\mathrm{H}} \\
&+ \mathbf{Q}_k^{\mathrm{noise}} \mathbf{\Sigma}_{\mathrm{NEXT}} \left[\mathbf{Q}_k^{\mathrm{noise}}\right]^{\mathrm{H}} \\
&+ \sigma_{\mathrm{ADC}}^2 \mathbf{Q}_k^{\mathrm{noise}} \left[\mathbf{Q}_k^{\mathrm{noise}}\right]^{\mathrm{H}} \\
&+ \frac{\sigma_{\mathrm{DNF}}^2}{\mathbf{w}^{\mathrm{T}}\mathbf{w}} \mathbf{I}
\end{aligned} \tag{15}$$

where $\mathbf{I}$ is the $M \times M$ identity matrix. $\tilde{\mathbf{A}}_k$ and $\tilde{\mathbf{B}}_k$ are Hermitian symmetric. Now, the model of the SNR (13) becomes

$$\mathrm{SNR}_{k}^{\mathrm{Model}}(\mathbf{w}) = \frac{\mathbf{w}^{\mathrm{T}} \tilde{\mathbf{A}}_k \mathbf{w}}{\mathbf{w}^{\mathrm{T}} \tilde{\mathbf{B}}_k \mathbf{w}} \tag{16}$$

$\mathrm{SNR}_{k}^{\mathrm{Model}}(\mathbf{w})$ is a ratio of quadratic functions of $\mathbf{w}$. A similar quadratic form has been derived in [10]. The proposed SNR model becomes equivalent to the SNR that could be measured at the output of the FFT in an ADSL system when the ISI and ICI have been removed from the received signal.

Section III-A establishes the relationship between the achievable data rate and the time domain equalizer coefficients using the proposed SNR model. Section III-B reflects on the impact of the transmission delay on the time domain equalizer coefficients and the achievable data rate.

A. Achievable data rate in terms of the proposed SNR model

The proposed SNR model is now substituted into the bit rate equation in order to extract the dependance of the bit rate on TEQ taps. By substituting $\mathrm{SNR}^{\mathrm{Model}}{}_{k}(\mathbf{w})$, our SNR model expression (16) into (5) we obtain

$$b_{\mathrm{DMT}}(\mathbf{w}, \mathcal{I}) = \sum_{k \in \mathcal{I}} \log_2 \left(\frac{\mathbf{w}^{\mathrm{T}}(\Gamma_k \tilde{\mathbf{B}}_k + \tilde{\mathbf{A}}_k)\mathbf{w}}{\mathbf{w}^{\mathrm{T}}(\Gamma_k \tilde{\mathbf{B}}_k)\mathbf{w}}\right) \tag{17}$$

where we include the dependence of the achievable DMT bit rate 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,

$$b_{\mathrm{DMT}}(\mathbf{w}, \mathcal{I}) = \sum_{k \in \mathcal{I}} \log_2 \left(\frac{\mathbf{w}^{\mathrm{T}} \mathbf{A}_k \mathbf{w}}{\mathbf{w}^{\mathrm{T}} \mathbf{B}_k \mathbf{w}} \right) \tag{18}$$

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 defined by

$$\mathbf{w}_{\mathrm{opt}} = \arg \max_{\mathbf{w}} \{ b_{\mathrm{DMT}}(\mathbf{w}, \mathcal{I}) | \mathbf{w} \in \mathcal{S} \} \tag{19}$$

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 (18) is scale invariant (i.e., $b_{\mathrm{DMT}}(\mathbf{w}, \mathcal{I}) = b_{\mathrm{DMT}}(\alpha \mathbf{w}, \mathcal{I}), \forall \alpha \neq 0$) [5] and noticing from (9) that choosing $\mathbf{w}^{\mathrm{T}}\mathbf{w} = \|\mathbf{w}\|^2 = 1$ does not enhance the ADC noise $\mathbf{n}_{\mathrm{y}}$, we adopt the unit norm constraint set $\mathcal{S} = \{\mathbf{w} : \mathbf{w}^{\mathrm{T}}\mathbf{w} = \|\mathbf{w}\|^2 = 1\}$. Notice that, $\tilde{\mathbf{B}}_k$ becomes independent of $\mathbf{w}$ over this constraint set since the last term in (15) becomes $\sigma^2_{\mathrm{DNF}}\mathbf{I}$.

Now we return to the issue of whether the maximizer of (18) will be the maximizer of

$$b^{\mathrm{int}}_{\mathrm{DMT}}(\mathbf{w}, \mathcal{I}) = \sum_{k \in \mathcal{I}} \left\lfloor \log_2 \left(\frac{\mathbf{w}^{\mathrm{T}} \mathbf{A}_k \mathbf{w}}{\mathbf{w}^{\mathrm{T}} \mathbf{B}_k \mathbf{w}} \right) \right\rfloor \tag{20}$$

Figs. 3 and 4 plot the values of both (18) and (20) 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 (18) and (20) 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 (18) maximizes (20) in these two cases; however, we cannot guarantee that this will be the case in general.

B. Impact of the transmission delay on time domain equalizer design

Dependence of the bit rate (SNR) on the transmission delay is not clearly visible from (18); however, transmission delay is an important design parameter. The transmission delay marks the amount of time that it takes the signal to reach the receiver from the time it has been transmitted. An 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 (7) and then used throughout in the definition of the proposed SNR measure. One can think of Δ as a parameter that selects the beginning of a symbol from a stream of samples at the output of the TEQ. Notice that once Δ is defined, all of the matrices that compose the SNR are defined, as well as (13). Thus, defining Δ prior to TEQ design also defines (18) and its optimal solution. A different Δ may result in a higher bit rate given by (18). 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 each optimization problem. This paper does not propose a better or less costly method. Developing a heuristic for the choice of optimal Δ remains an open research problem.

[5] 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.

IV. Proposed TEQ Design: Optimal TEQ Filter Bank and single TEQ

A. TEQ Filter Bank: TEQFB

There are $N/2$ possible data carrying subchannels in DMT. We will now propose a method of finding the optimal $\mathbf{w}$ for every one of them. Maximizing the number of bits allocated in a single subchannel, $b_k(\mathbf{w})$, 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

$$\begin{aligned} b_k^{\text{opt}} &= \max_{\mathbf{w}_k:\|\mathbf{w}_k\|^2=1} \left(\log_2 \left(\frac{\mathbf{w}_k^{\mathrm{T}} \mathbf{A}_k \mathbf{w}_k}{\mathbf{w}_k^{\mathrm{T}} \mathbf{B}_k \mathbf{w}_k} \right) \right) \\ &= \log_2 \left(\max_{\mathbf{w}_k:\|\mathbf{w}_k\|^2=1} \left(\frac{\mathbf{w}_k^{\mathrm{T}} \mathbf{A}_k \mathbf{w}_k}{\mathbf{w}_k^{\mathrm{T}} \mathbf{B}_k \mathbf{w}_k} \right) \right) \end{aligned} \tag{21}$$

From [24], the maximization of a single ratio can be transformed into

$$p_k(\mathbf{w}_k, \lambda_k) = \max_{\mathbf{w}_k:\|\mathbf{w}_k\|^2=1} \left\{ \mathbf{w}_k^{\mathrm{T}} \mathbf{A}_k \mathbf{w}_k - \lambda_k \mathbf{w}_k^{\mathrm{T}} \mathbf{B}_k \mathbf{w}_k \right\} \tag{22}$$

where λ_k is a scalar. To solve (22), 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

$$\mathbf{A}_k^r \mathbf{w}_k = \lambda_k \mathbf{B}_k^r \mathbf{w}_k \tag{23}$$

Here $(\cdot)^r$ denotes the real part. This is the well-known generalized eigenvalue problem [25] and the solution is the generalized eigenvector $\mathbf{w}_k^{\text{opt}}$ corresponding to the largest generalized eigenvalue λ_k^{opt} of $(\mathbf{A}_k^r, \mathbf{B}_k^r)$:

$$\lambda_k^{\text{opt}} = \frac{\left(\mathbf{w}_k^{\text{opt}}\right)^{\mathrm{T}} \mathbf{A}_k^r \mathbf{w}_k^{\text{opt}}}{\left(\mathbf{w}_k^{\text{opt}}\right)^{\mathrm{T}} \mathbf{B}_k^r \mathbf{w}_k^{\text{opt}}} = \frac{\left(\mathbf{w}_k^{\text{opt}}\right)^{\mathrm{T}} \mathbf{A}_k \mathbf{w}_k^{\text{opt}}}{\left(\mathbf{w}_k^{\text{opt}}\right)^{\mathrm{T}} \mathbf{B}_k \mathbf{w}_k^{\text{opt}}} \tag{24}$$

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

$$b_{\text{DMT}}^{\text{opt}} = \sum_{k \in \mathcal{I}} \log_2 \left[\frac{(\mathbf{w}_k^{\text{opt}})^{\mathrm{T}} \mathbf{A}_k \mathbf{w}_k^{\text{opt}}}{(\mathbf{w}_k^{\text{opt}})^{\mathrm{T}} \mathbf{B}_k \mathbf{w}_k^{\text{opt}}} \right] \tag{25}$$

We could realize (25) as the optimal per-tone TEQ filter bank (TEQFB) with $|\mathcal{I}|$ TEQs. The FFT block could be implemented as a bank of Goertzel filters [17] with each one computing a single point DFT.

TEQFB provide the upper limit on what we can achieve in terms of maximizing bit allocation and is used for that purpose in this paper. TEQFB was not proposed as a practical approach to equalization but as a guide to the performance that can be achieved by a single TEQ FIR in an architecture where the FEQ is a separate block. 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 (21) 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:

$$\mathcal{J}_k = \{\mathbf{w} : \ \lfloor b_k(\mathbf{w}) \rfloor = \lfloor b_k(\mathbf{w}_k^{opt}) \rfloor, \ \mathbf{w} \in \mathcal{S}\} \tag{26}$$

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

$$\mathcal{J} = \{\mathbf{w} : \ \mathbf{w} \in \bigcap_{k \in \mathcal{I}} \mathcal{J}_k\} \tag{27}$$

$\mathcal{J}$ may contain many different single TEQs due to the integer-only loading and/or the behavior of b_{DMT}, or may be an empty set. The maximization of (25) should result in a member of $\mathcal{J}$ if one exists, or if $\mathcal{J}$ in a null set, it should find a single TEQ that performs as close as possible to TEQFB. So far, we are not aware of a method that lets us determine in advance if $\mathcal{J}$ is empty or not.

B. Single TEQ design for the subchannels of interest

Finding a single TEQ FIR whose performance is comparable to the performance of TEQFB would result in significant computational savings and simpler architecture. The problem we are considering is

$$\mathbf{w}_{\text{opt}} = \arg\max_{\mathbf{w}} \{b_{\text{DMT}}(\mathbf{w}, \mathcal{I})\} \tag{28}$$

The bit rate given by (18) 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

$$w_0 = \prod_{l=1}^{M-1} \sin(\theta_l) \tag{29}$$

$$w_m = \cos(\theta_m) \prod_{l=m+1}^{M-1} \sin(\theta_l), \ m = 1, 2, \cdots, M-1 \tag{30}$$

The unit hyper-sphere constraint allows us to represent the solution using efficient notation and it does not unduly constrain us as to leave possible theoretical solutions out of the scope of the problem. For example, for M=2, all of the possible $\mathbf{w}$'s lie on a unit circle. Changing the size of the circle would change the amplitude of each $\mathbf{w}$ but does not change their performance relative to to one another.

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

$$\begin{aligned} \mathcal{T} &= \{(\theta_1, \theta_2, \cdots, \theta_{M-1}) \\ \theta_m &\in [-\pi/2, \pi/2], \ m = 1, 2, \cdots, M-1\} \end{aligned} \tag{31}$$

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 Fig. 4.

We present a method that will find the global maximum provided that we have a "good" starting point. We also present a procedure for finding a good starting point. The method guarantees the closest local maximum to our starting point with quadratic convergence, without the oscillatory behavior associated with standard steepest ascent algorithms [26].

A method to maximize a sum-of-ratios (i.e. (18)) is an active research topic in the fractional programming community for which no definitive solution exists yet (see e.g. [24], [27]). We approach the problem by applying Almogy and Levin's method [28], which is based on the Dinkelbach approach [29], to find the root of the first derivative of (18) that corresponds to the closest local maximum of $b_{\mathrm{DMT}}(\mathbf{w}, \mathcal{I})$. For further discussion on our optimization approach leading to single TEQ design algorithm see Appendix G. 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 single TEQ design algorithm is shown in Fig. 5. The key steps in the proposed single TEQ Design algorithm (Fig. 5) 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.

C. Computational Complexity

1) Initialization Cost: The cost of creating matrices $\mathbf{A}_k$ and $\mathbf{B}_k$ for all $k \in \mathcal{I}$ taking into account that the matrices are symmetric, positive definite with a special structure (see the Appendices) is of the order $\mathcal{O}(|\mathcal{I}|M^2N)$ operations where M is the size of the TEQ FIR ranging from 2 to 32 in our simulations and $N = 512$ in G.DMT downstream. Here we assume that the channel and the variances of the signal and the noise are already known. The creation of these matrices is necessary for both the proposed TEQFB and single TEQ.

TEQFB algorithm requires up to $N_u = 250$ (when all DMT data-carrying subchannels are active) solutions of symmetric-definite problems where we are looking for the largest generalized eigenvalue-eigenvector pair. From [30, p. 465] the largest generalized eigenvalue-eigenvector pair can be found using an extension of the Rayleigh quotient iteration. The cost of this procedure is $iter_{\mathrm{R}}(M^3/3+12M^2+4M)$ where $iter_{\mathrm{R}}$ is the number of iterations used. Thus, the initialization cost of TEQFB is $N_u * iter_{\mathrm{R}}(M^3/3 + 12M^2 + 4M)$. The initialization cost of the proposed single TEQ algorithm is up to
$iter_{\mathrm{S}}\left[N_u(9M^2 + 12M) + iter_{\mathrm{R}}(M^3/3 + 12M^2 + 4M)\right]$ where $iter_{\mathrm{S}}$ means the number of iterations of the maximization procedure.

2) Data Transmission Complexity: The computational complexity and memory requirements during data transmission for the conventional, single TEQ architecture, TEQFB and PTE architectures are shown in Table I. Thus, TEQFB can have lower memory needs than PTE; however, it has significantly higher computational requirements during data transmission that make it too expensive for cost-effective embedded implementation today.

V. Simulation Results

The simulation results compare the performance of the proposed TEQ design methods with Min-ISI, MBR, MMSE-UEC, MMSE-UTC and LS PTE,initialized using the least-squares method. We use the eight standard downstream CSA loops [1] convolved with transmit and receive filters as the test CIR. The transmit and receive filters are modelled as first-order high-pass IIR filters, which are designed to separate ADSL from the 0-4 kHz voice band (a double zero is located at $z = 1$, while conjugate symmetric poles are located at $z = 0.9799 \pm \mathrm{j}0.0317$). All CIRs 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 use the FFT size $N = 512$ standard in downstream ADSL with $\Gamma = 9.8$ dB, and thus we do not add the effects of possible coding gain or margin [6]. All of the power terms used in the simulations are defined with respect to $100\ \Omega$ 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 [1]. The effects of DNF and ADC quantization are not simulated in this work. Delay Δ is taken as a free parameter in the optimization and is varied over values 0 to 40 with the best achieved data rate taken to be the best result. The upper limit of 40 was heuristically determined for the tested channels. Thus, once the optimal transmission delay Δ_{opt} is chosen, the simulation software proceeds with the initialization of the TEQ block filter coefficients according to the method in question (TEQFB, MBR, etc.)

TEQ FIR designed by Min-ISI, MBR, MMSE-UTC, MMSE-UEC methods is obtained using the Matlab DMT TEQ Design Toolbox [31]. PTE architecture presented in [13] was initialized using the least-squares method that does not require knowledge of the channel and fits the frequency-domain received symbols to the expected symbols through per-tone equalizers using normal equations [25, p. 106]. The same number of symbols (300) were used to let the LS PTE converge for all TEQ sizes M. The number of frames in the LS PTE training sequence was the same for all PTE equalizer lengths. The heuristic number of frames used for the LS PTE convergence arose from the simulations in which the training sequence length was varied from 300 to 1024 frames for a limited number of cases involving various CSA loops and PTE equalizer lengths resulting an improvement of less than 70 kbps.

We measure the SNR in our simulations as it would be measured in a real ADSL DMT system during modem initialization. During training, all subchannels are loaded with a randomly chosen two-bit constellation point at the transmitter (Fig. 1). The symbols are convolved with the channel impulse response (including the transmit filter, a CSA loop impulse response and a receive filter), passed through the TEQ block designed by one of the methods being evaluated. Signal is demodulated using the FFT block. The frequency domain equalizer removes the phase and magnitude distortion. QAM decoding compares the complex value received in a particular subchannel with the transmitted value so that SNR measurement can be derived from the power of the error averaged over 1000 symbols

[6] The bit loading tables in ADSL are 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.

for any particular subchannel k. The SNR measurement equation is

$$\mathrm{SNR}_k = 10\log_{10}\left(\frac{2}{\frac{1}{S}\sum_{i=0}^{S-1}\|X_k^i - Y_k^i\|^2}\right) \tag{32}$$

where $S = 1000$ symbols, X_k^i is the transmitted 2-bit constellation data in the subchannel k in the i^{th} symbol, Y_k^i is the received data in the subchannel k in the i^{th} symbol and 2 comes from the power of two-bit constellation point at 2-D constellation points $(\pm 1, \pm 1)$. The bit rates reported are calculated using the measured SNR given by (32) on subchannels 7-256, following the ADSL standard for 10^{-7} BER and up to 15-bit quadrature amplitude modulation constellations allowed per subchannel.

We calculate the accuracy of the SNR measurement using 1000 symbols to be approximately ± 0.5 dB, which means that our measured SNR is within 0.5 dB of the true SNR. This result was obtained using confidence intervals for variance estimation [32, p. 298]. The term "true" SNR refers to the SNR that would be measured at the output of the frequency domain equalizer as the number of symbols S goes to ∞ in (32). Given a 1000-symbol measurement, the reported bit rates are accurate up to ± 60 kbps. The measured SNR was calculated using double-precision arithmetic.

Proposed single TEQ and TEQFB , Min-ISI, MBR, MMSE-UTC, MMSE-UEC and LS PTE design methods are evaluated using this SNR measurement, 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. There is no assurance that the superior performance as defined using some of those measures indeed translates into superior bit rate as it is measured and applied in an actual ADSL DMT system initialization. That is why we chose to evaluate the performance of the proposed and previously published methods using SNR measured very closely to the way done in ADSL DMT systems and not using our own proposed SNR model. Our intent is to use an objective method of testing a TEQ design that relates the success of the TEQ design to the data rate (SNR) measurable on an ADSL DMT system.

In Fig. 2 we show the measured SNR and the proposed model SNR for the case of $M = 32$ and CSA loop 2. The measured SNR was obtained with the TEQFB in the signal path. The TEQFB was obtained by applying the described TEQFB design procedure. The model SNR was obtained by evaluating (13) for the calculated TEQFB. The proposed model SNR closely approximates the measured SNR.

Table II lists the highest data rate achieved with the proposed optimal TEQFB for the CIR including CSA loops 1-8 at the given signal and noise power levels. Table III shows the achieved data rates of the proposed single FIR TEQ, MBR, Min-ISI, MMSE-UEC, MMSE-UTC, and LS PTE for the channel impulse responses including CSA loops 1-8 as percentage (%) of the data rate of the proposed TEQFB. The row entry under $LoopAvg$ is calculated as

$$LoopAvg(CSA\ loop, Algorithm) = \frac{1}{31}\sum_{M=2}^{32}\frac{b_{\mathrm{Algorithm}}(\Delta_{\mathrm{opt}}, CSA\ loop, M)}{b_{\mathrm{DMT}}^{\mathrm{opt}}(CSA\ loop, M)} * 100\% \tag{33}$$

where $b_{\mathrm{Algorithm}}(*)$ is the highest bit rate achieved by the evaluated algorithm for the given TEQ size M; Δ_{opt} is found using line search, and $b_{\mathrm{DMT}}^{\mathrm{opt}}$ is the bit rate of TEQFB defined in (25) for the given TEQ size M. Experiments are performed for each CSA loop with M ranging from 2 to 32 for standard $N = 512$ and $\nu = 32$.

The last row entry is the average for the evaluated algorithm.

$$Average(method) = \frac{1}{8} \sum_{CSA\ loop=1}^{8} LoopAvg(CSA\ loop, Algorithm) \tag{34}$$

Our intent is to show how the evaluated algorithm may perform on "average" across loops and TEQ filter sizes because in a deployed ADSL DMT system, the algorithm will need to work across varying loop topologies and may use different TEQ sizes.

TEQFB simulation data rates in Table III are higher than the compared methods. The average difference in the achieved data rate between TEQFB and LS PTE is small, which can be expected given that the time domain filter bank architecture and the per-tone architecture perform equalization for each subchannel separately. Here, we note again that the initialization of TEQFB coefficients was performed with the full knowledge of the channel while the PTE used a training sequence and did not have the knowledge of the channel impulse response. The proposed single FIR TEQ design method achieves higher percentage of TEQFB data rate for each CIR than Min-ISI, MBR or MMSE-based methods. The proposed single FIR TEQ final average is almost 2% higher than either MBR or the Min-ISI and more than 15% higher than MMSE-UTC or MMSE-UEC. For a data rate of 11 Mbps, a 2% improvement amounts to 220 kbps. Overall, single FIR TEQ does achieve the data rate that the LS PTE achieves; however, the differences in the achieved data rate are small.

In Fig. 6 we show how the achieved data rate varies with the change in the number of TEQ taps M ranging from 2 to 32 in an example for the channel impulse response containing CSA loop 2. The bit rate grows significantly from 2 to 3 TEQ taps. The upward slope is present with the further increase in the number of TEQ taps M but it is significantly moderated. LS PTE follows TEQFB closely for M from 2 to 8, achieving a higher data rate than TEQFB for $M = 6$; however, the LS PTE data rate drops off for larger sizes of M for the largest gap of 140 kbps. The proposed single TEQ design performs closely to TEQFB and LS PTE for some sizes of M and achieves higher data rate than Min-ISI, or MBR for all sizes of M in this figure.

Fig. 7 shows how the bit rate changes as a function of transmission delay Δ for the proposed single TEQ design method. The CIR shown includes the CSA loop 1 for the TEQ length $M = \{3, 10, 30\}$, $N = 512$, and $\nu = 32$. 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" can aid TEQ design methods when it comes to the choice of Δ.

Fig. 8 compares the measured SNR achieved with the proposed single TEQ design, MBR and Min-ISI for CSA loop 4. 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.

VI. Conclusions

In this paper we proposed an SNR model 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 model 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 model is a ratio of quadratic functions in TEQ taps $\mathbf{w}$. We use the SNR model in the definition of the bit rate which we maximize as a function of the TEQ block.

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 simplify the FFT calculations to perform multicarrier demodulation. 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 single 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. Thus, single TEQ outperforms state of the art MBR and Min-ISI by 2%, and performs closely with LS PTE. 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 would put a number of nulls elsewhere in the spectrum.

Acknowledgments

The authors would like to thank Mr. Ketan J. Mandke and Ms. Esther I. Resendiz of The University of Texas (Austin, TX) for their Matlab implementations of part of the proposed TEQ design algorithms. The authors would also like to thank Mr. Geert Ysebaert and Mr. Koen Vanbleu of Katholieke Univeristeit Leuven (Leuven, Belgium), Dr. Lloyd D. Clark of Schlumberger (Houston, TX), Dr. Güner Arslan of Silicon Laboratories (Austin, TX), and Mr. Ming Ding of The University of Texas (Austin, TX) for their helpful suggestions on the paper.

APPENDIX

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

Let

$$\mathbf{U}_i^{\Delta} = \left[\begin{array}{cc} \left(\mathbf{U}_i^{\Delta}\right)_L & \left(\mathbf{U}_i^{\Delta}\right)_R \end{array} \right] \tag{35}$$

where

$$\left(\mathbf{U}_i^{\Delta}\right)_L = \begin{bmatrix} u_i^{\Delta} & u_i^{\Delta-1} & \cdots & u_i^{0} & u_i^{N-1} & \cdots & u_i^{N-\nu} \\ u_i^{\Delta+1} & u_i^{\Delta} & \cdots & \cdots & u_i^{0} & u_i^{N-1} & \cdots \\ \vdots & \vdots & & & & \vdots & \vdots \\ u_i^{N-1} & u_i^{N-2} & \cdots & \cdots & \cdots & u_i^{0} & u_i^{N-1} \\ 0 & u_i^{N-1} & u_i^{N-2} & \cdots & \cdots & u_i^{1} & u_i^{0} \\ \vdots & & & & & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & u_i^{N-1} \\ \vdots & & & & & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \end{bmatrix} \tag{36}$$

and

$$\left(\mathbf{U}_i^{\Delta}\right)_R = \begin{bmatrix} 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ u_i^{N-\nu} & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ u_i^{N-2} & \cdots & u_i^{N-\nu} & 0 & \cdots & 0 \\ u_i^{N-1} & \cdots & u_i^{N-\nu-1} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ u_i^{N-2} & \cdots & u_i^{0} & u_i^{N-1} & \cdots & u_i^{N-\nu} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & u_i^{N-1} & \cdots & u_i^{\Delta-M-1} \end{bmatrix} \tag{37}$$

$$\mathbf{U}_{i-1}^{\Delta} = \begin{bmatrix} 0 & \cdots & 0 & u_{i-1}^{N-1} & u_{i-1}^{N-2} & \cdots & u_{i-1}^{\nu+\Delta-M+1} \\ \vdots & & & & & & \vdots \\ 0 & \cdots & \cdots & 0 & u_{i-1}^{N-1} & \cdots & u_{i-1}^{N+\nu-M+1} \\ \vdots & & & & & & \vdots \\ 0 & \cdots & \cdots & 0 & \cdots & 0 & u_{i-1}^{N-1} \\ 0 & \cdots\cdots & 0 & 0 & \cdots & \cdots & 0 \\ \vdots & & & & & & \vdots \\ 0 & \cdots\cdots & 0 & 0 & \cdots & \cdots & 0 \end{bmatrix} \tag{38}$$

$$\mathbf{U}_{i+1}^{\Delta} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \cdots & 0 & 0 & 0 \\ \vdots & & & & & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ u_{i+1}^{N-\nu} & 0 & \cdots & \cdots & 0 & \cdots & \cdots \\ \vdots & & & & & & \vdots \\ u_{i+1}^{0} & u_{i+1}^{N-1} & \cdots & u_{i+1}^{N-\nu} & 0 & \cdots & 0 \\ \vdots & & & & & & \vdots \\ u_{i+1}^{\Delta-\nu-1} & \cdots & u_{i+1}^{0} & u_{i+1}^{N-1} & \cdots & u_{i+1}^{N-\nu} & 0 \end{bmatrix} \tag{39}$$

$$\left[\mathbf{U}_{i}^{\Delta}\right]_{\mathrm{circ}} = \begin{bmatrix} u_i^{\Delta} & \cdots & u_i^{0} & u_i^{N-1} & \cdots & u_i^{\Delta-M+2} \\ \vdots & & & & \vdots & \\ u_i^{N-1} & \cdots & u_i^{\Delta} & u_i^{\Delta-1} & \cdots & u_i^{N-M+1} \\ u_i^{0} & \cdots & u_i^{\Delta+1} & u_i^{\Delta} & \cdots & u_i^{N-M+2} \\ \vdots & & & & & \vdots \\ u_i^{\Delta-1} & \cdots & u_i^{0} & u_i^{N-1} & \cdots & u_i^{\Delta-M-1} \end{bmatrix} \tag{40}$$

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

$$\mathbf{H} = \begin{bmatrix} h_0 & 0 & 0 & \cdots & 0 \\ h_1 & h_0 & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ h_{M-1} & \cdots & & \cdots & h_0 \\ \vdots & & & & \vdots \\ h_{N-1} & \cdots & & \cdots & h_{N-1-M} \\ 0 & h_{N-1} & \cdots & \cdots & h_{N-M} \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & h_{N-1} \end{bmatrix} \tag{41}$$

Define

$$\mathbf{H} = \begin{bmatrix} (\mathbf{H}_{\mathrm{u}})_{\Delta\times M} \\ (\mathbf{H}_{\mathrm{c}})_{\nu\times M} \\ (\mathbf{H}_{\mathrm{b}})_{(N-\nu-\Delta+M-1)\times M} \end{bmatrix} \tag{42}$$

C. Definition of noise-dependent matrices: $\mathbf{G}_{\mathrm{AWGN}}$,$\mathbf{G}_{\mathrm{NEXT}}$ and $\mathbf{G}_{\mathrm{ADC}}$; all $(N+M-1)\times M$ matrices

Let the NEXT, AWGN and ADC quantization noise vectors be

$$\mathbf{n}^{\mathrm{xx}} = \{n^{\mathrm{xx}}_{-M+1}, n^{\mathrm{xx}}_{-M+2}, \cdots, n^{\mathrm{xx}}_{0}, \cdots, n^{\mathrm{xx}}_{N-1}\}^{\mathrm{T}} \tag{43}$$

where the subscript $(.)_{\mathrm{xx}}$ is equal to NEXT, AWGN or ADC. Then the $(N+M-1)\times M$ AWGN, NEXT and ADC convolution matrices with the TEQ, $\mathbf{G}_{\mathrm{AWGN}}$, $\mathbf{G}_{\mathrm{NEXT}}$ and $\mathbf{G}_{\mathrm{ADC}}$ are

$$\mathbf{G}_{\mathrm{xx}} = \begin{bmatrix} n^{\mathrm{xx}}_{0} & n^{\mathrm{xx}}_{-1} & \cdots & n^{\mathrm{xx}}_{-M+1} \\ n^{\mathrm{xx}}_{1} & n^{\mathrm{xx}}_{0} & \cdots & n^{\mathrm{xx}}_{-M+2} \\ \vdots & & & \vdots \\ n^{\mathrm{xx}}_{M-1} & \cdots & \cdots & n^{\mathrm{xx}}_{0} \\ \vdots & & & \vdots \\ n^{\mathrm{xx}}_{N-1} & \cdots & & n^{\mathrm{xx}}_{N-1-M} \\ \vdots & & & \vdots \\ n^{\mathrm{xx}}_{N+M} & n^{\mathrm{xx}}_{N+M-1} & \cdots & n^{\mathrm{xx}}_{N-1} \end{bmatrix} \tag{44}$$

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

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

$$\mathbf{Q}_k^{\mathrm{noise}} = \begin{bmatrix} 0 & 0 & \cdots & 0 & q_k^0 & \cdots & q_k^{N-1} \\ 0 & 0 & \cdots & q_k^0 & \cdots & q_k^{N-1} & 0 \\ \vdots & & & & & & \vdots \\ q_k^0 & \cdots & q_k^{N-1} & 0 & \cdots & 0 & 0 \end{bmatrix} \tag{45}$$

where $q_k^{(\cdot)}$ are members of the vector $\mathbf{q}_k$ (8) and $\mathbf{Q}_k^{\mathrm{noise}}$ is $M\times(N+M-1)$ matrix.

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

$$\mathbf{V}_k = \begin{bmatrix} q_k^{N-\Delta} & q_k^{N-\Delta+1} & \cdots & q_k^{N-2} & q_k^{N-1} \\ q_k^{N-\Delta+1} & q_k^{N-\Delta} & \cdots & q_k^{N-1} & 0 \\ \vdots & & & \vdots & \vdots \\ q_k^{N-1} & 0 & \cdots & 0 & 0 \end{bmatrix}_{\Delta\times\Delta} \tag{46}$$

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

$$\mathbf{W}_k = \begin{bmatrix} 0 & 0 & \cdots & 0 & q_k^0 \\ 0 & 0 & \cdots & q_k^0 & q_k^1 \\ \vdots & & & \vdots & \vdots \\ q_k^0 & q_k^1 & \cdots & \cdots & q_k^{N-\nu-\Delta+M-2} \end{bmatrix} \tag{47}$$

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

$$\mathbf{Q}_k^{\mathrm{circ}} = \begin{bmatrix} q_k^{ND+1} & \cdots & \cdots & q_k^{N1} & q_0 & \cdots & q_k^{ND} \\ q_k^{ND+2} & \cdots & q_k^{N1} & q_0 & \cdots & \cdots & q_k^{ND+1} \\ \vdots & & & & & & \vdots \\ q_k^0 & q_k^1 & \cdots & \cdots & \cdots & \cdots & q_k^{N1} \\ \vdots & & & & & & \vdots \\ q_k^{MD-3} & q_k^{MD-2} & \cdots & q_0 & \cdots & \cdots & q_k^{N-MD-2} \end{bmatrix} \tag{48}$$

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

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

$$\begin{aligned} E\left[\left(Y_{\mathrm{D}}^k\right)^{\mathrm{H}} Y_{\mathrm{D}}^k\right] &= E\left[\mathbf{w}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\left[\mathbf{U}_i^{\Delta}\right]_{\mathrm{circ}}^{\mathrm{T}}\mathbf{q}_k\mathbf{q}_k^{\mathrm{H}}\left[\mathbf{U}_i^{\Delta}\right]_{\mathrm{circ}}\mathbf{H}\mathbf{w}\right] \\ &= \mathbf{w}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{Q}_k^{\mathrm{circ}} E\left[\mathbf{u}_i\mathbf{u}_i^{\mathrm{T}}\right]\left[\mathbf{Q}_k^{\mathrm{circ}}\right]^{\mathrm{H}}\mathbf{H}\mathbf{w} \\ &= \sigma_s^2\mathbf{w}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{Q}_k^{\mathrm{circ}}\left[\mathbf{Q}_k^{\mathrm{circ}}\right]^{\mathrm{H}}\mathbf{H}\mathbf{w} \\ &= \mathbf{w}^{\mathrm{T}}\tilde{\mathbf{A}}_k\mathbf{w} \end{aligned} \tag{49}$$

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.

F. 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.

$$\begin{aligned} Y_{\mathrm{R}}^k - Y_{\mathrm{D}}^k &= \mathbf{q}_k^{\mathrm{H}}\left(\mathbf{U}_{i-1}^{\Delta} + \mathbf{U}_{i+1}^{\Delta} + \underbrace{\mathbf{U}_i^{\Delta} - [\mathbf{U}_i^{\Delta}]_{\mathrm{circ}}}_{\mathbf{P}}\right)\mathbf{H}\mathbf{w} \\ &+ \mathbf{q}_k^{\mathrm{H}}\mathbf{G}_{\mathrm{AWGN}}\mathrm{w} \\ &+ \mathbf{q}_k^{\mathrm{H}}\mathbf{G}_{\mathrm{NEXT}}\mathbf{w} \\ &+ \mathbf{q}_k^{\mathrm{H}}\mathbf{G}_{\mathrm{ADC}}\mathbf{w} + D_k \end{aligned} \tag{50}$$

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

$$\mathbf{P} = \left[\mathbf{U}_{i,\mathrm{L}}^{\Delta}\right] + \left[\mathbf{U}_{i,\mathrm{U}}^{\Delta}\right] \tag{51}$$

where

$$\left[\mathbf{U}_{i,\mathrm{L}}^{\Delta}\right] = \begin{bmatrix} 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & & & & & & \vdots \\ u_i^0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ u_i^1 & u_i^0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ u_i^{\Delta-1} & u_i^{\Delta-2} & \cdots & u_i^0 & 0 & \cdots & 0 \end{bmatrix} \tag{52}$$

and

$$\left[\mathbf{U}_{i,\mathrm{U}}^{\Delta}\right] = \begin{bmatrix} 0 & \cdots & 0 & u_i^{NV1} & u_i^{NV1-1} & \cdots & u_i^{DM+2} \\ 0 & \cdots & 0 & 0 & u_i^{NV1} & \cdots & u_i^{DM+3} \\ \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & 0 & u_i^{NV1} \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \tag{53}$$

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 as

$$Z_k = Y_{\mathrm{R}}^k - Y_{\mathrm{D}}^k \tag{54}$$

Then the power in Z_k is

$$\begin{aligned} EP &= E\left[Z_k^{\mathrm{H}} Z_k\right] \qquad (55)\\ &= \mathbf{w}^{\mathrm{T}} E\left[\mathbf{S}\right]\mathbf{w} + \mathbf{w}^{\mathrm{T}} E\left[\underbrace{\mathbf{G}_{\mathrm{AWGN}}^{\mathrm{T}}\mathbf{q}_k\mathbf{q}_k^{\mathrm{H}}\mathbf{G}_{\mathrm{AWGN}}}_{\mathbf{EP1}}\right]\mathbf{w} + \mathbf{w}^{\mathrm{T}} E\left[\underbrace{\mathbf{G}_{\mathrm{NEXT}}^{\mathrm{T}}\mathbf{q}_k\mathbf{q}_k^{\mathrm{H}}\mathbf{G}_{\mathrm{NEXT}}}_{\mathbf{EP2}}\right]\mathbf{w} \\ &+ \mathbf{w}^{\mathrm{T}} E\left[\underbrace{\mathbf{G}_{\mathrm{NEXT}}^{\mathrm{T}}\mathbf{q}_k\mathbf{q}_k^{\mathrm{H}}\mathbf{G}_{\mathrm{ADC}}}_{\mathbf{EP3}}\right]\mathbf{w} + E\left[D_k^{\mathrm{H}} D_k\right] \end{aligned}$$

Where

$$\begin{aligned} \mathbf{S} &= \underbrace{\mathbf{H}^{\mathrm{T}}\left[\mathbf{U}_{i-1}^{\Delta}\right]^{\mathrm{H}}\mathbf{q}_k\mathbf{q}_k^{\mathrm{H}}\mathbf{U}_{i-1}^{\Delta}\mathbf{H}}_{\mathbf{S1}} \\ &+ \underbrace{\mathbf{H}^{\mathrm{T}}\left[\mathbf{U}_{i+1}^{\Delta}\right]^{\mathrm{H}}\mathbf{q}_k\mathbf{q}_k^{\mathrm{H}}\mathbf{U}_{i+1}^{\Delta}\mathbf{H}}_{\mathbf{S2}} \\ &+ \underbrace{\mathbf{H}^{\mathrm{T}}\mathbf{P}^{\mathrm{H}}\mathbf{q}_k\mathbf{q}_k^{\mathrm{H}}\mathbf{P}\mathbf{H}}_{\mathbf{S3}} \end{aligned} \tag{56}$$

Notice that now we can transform the parts of (56) as follows

$$\left[\mathbf{U}_{i+1}^{\Delta}\right]^{\mathrm{H}}\mathbf{q}_k = \left[\begin{array}{ccc}\mathbf{V}_k^{\mathrm{b}} & \mathbf{0} & \mathbf{V}_k^{\mathrm{u}}\\ \mathbf{0} & \mathbf{0} & \mathbf{0}\end{array}\right]_{(N+M-1)\times N}\mathbf{u}_{i+1} \tag{57}$$

according to definitions in (46). Similarly,

$$\left[\mathbf{U}_{i-1}^{\Delta}\right]^{\mathrm{H}}\mathbf{q}_k = \left[\begin{array}{cc}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{W}_k\end{array}\right]_{(N+M-1)\times N}\mathbf{u}_{i-1} \tag{58}$$

where $\mathbf{W}_k$ has been defined in (47). Then

$$\left[\mathbf{U}_{i,\mathrm{L}}^{\Delta}\right]^{\mathrm{H}}\mathbf{q}_k = \left[\begin{array}{cc}\mathbf{V}_k & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{array}\right]_{(N+M-1)\times N}\mathbf{u}_{i} \tag{59}$$

and

$$\left[\mathbf{U}_{i,\mathrm{U}}^{\Delta}\right]^{\mathrm{H}}\mathbf{q}_k = \left[\begin{array}{ccc}\mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{W}_k & \mathbf{0}\end{array}\right]_{(N+M-1)\times N}\mathbf{u}_{i} \tag{60}$$

Now we can go forward with the derivation in

$$\begin{aligned}E\left[\mathbf{S2}\right] &= \mathbf{H}^{\mathrm{T}}\left[\begin{array}{ccc}\mathbf{V}_k^{\mathrm{b}} & \mathbf{0} & \mathbf{V}_k^{\mathrm{u}}\\ \mathbf{0} & \mathbf{0} & \mathbf{0}\end{array}\right]E[\mathbf{u}_{i+1}\mathbf{u}_{i+1}^{\mathrm{T}}]\left[\begin{array}{cc}\left[\mathbf{V}_k^{\mathrm{b}}\right]^{\mathrm{H}} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\\ \left[\mathbf{V}_k^{\mathrm{u}}\right]^{\mathrm{H}} & \mathbf{0}\end{array}\right]\mathbf{H}\\ &= \sigma_s^2\mathbf{H}^{\mathrm{T}}\left[\begin{array}{cc}\mathbf{V}_k^{\mathrm{b}}\left[\mathbf{V}_k^{\mathrm{b}}\right]^{\mathrm{H}}+\mathbf{V}_k^{\mathrm{u}}\left[\mathbf{V}_k^{\mathrm{u}}\right]^{\mathrm{H}} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{array}\right]\mathbf{H}\\ &= \sigma_s^2\mathbf{H}_{\mathrm{u}}^{\mathrm{T}}\mathbf{V}_k\mathbf{V}_k^{\mathrm{H}}\mathbf{H}_{\mathrm{u}}\end{aligned} \tag{61}$$

where $\mathbf{H}_u$ was defined in (42). Similarly,

$$\begin{aligned}E\left[\mathbf{S1}\right] &= \mathbf{H}^{\mathrm{T}}\left[\begin{array}{cc}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{W}_k\end{array}\right]E[\mathbf{u}_{i-1}\mathbf{u}_{i-1}^{\mathrm{T}}]\left[\begin{array}{cc}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{W}_k^{\mathrm{H}}\end{array}\right]\mathbf{H}\\ &= \sigma_s^2\mathbf{H}^{\mathrm{T}}\left[\begin{array}{cc}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{W}_k\end{array}\right]\left[\begin{array}{cc}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{W}_k^{\mathrm{H}}\end{array}\right]\mathbf{H}\\ &= \sigma_s^2\mathbf{H}_{\mathrm{b}}^{\mathrm{T}}\mathbf{W}_k\mathbf{W}_k^{\mathrm{H}}\mathbf{H}_{\mathrm{b}}\end{aligned} \tag{62}$$

where $\mathbf{H}_{\mathrm{b}}$ was defined in (42).

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 (61) and (62). The contribution of AWGN, NEXT and ADC noise are respectively given by

$$\begin{aligned}E\left[\mathbf{EP1}\right] &= \mathbf{Q}_k^{\mathrm{noise}}E[\mathbf{n}^{\mathrm{AWGN}}\left(\mathbf{n}^{\mathrm{AWGN}}\right)^{\mathrm{T}}]\left[\mathbf{Q}_k^{\mathrm{noise}}\right]^{\mathrm{H}}\\ &= \sigma_{\mathrm{AWGN}}^2\mathbf{Q}_k^{\mathrm{noise}}\left[\mathbf{Q}_k^{\mathrm{noise}}\right]^{\mathrm{H}}\end{aligned} \tag{63}$$

$$\begin{aligned} E\left[\mathbf{EP2}\right] &= \mathbf{Q}_k^{\mathrm{noise}} E[\mathbf{n}^{\mathrm{NEXT}} \left(\mathbf{n}^{\mathrm{NEXT}}\right)^{\mathrm{T}}] \left[\mathbf{Q}_k^{\mathrm{noise}}\right]^{\mathrm{H}} \\ &= \mathbf{Q}_k^{\mathrm{noise}} \Sigma_{\mathrm{NEXT}} \left[\mathbf{Q}_k^{\mathrm{noise}}\right]^{\mathrm{H}} \end{aligned} \tag{64}$$

$$\begin{aligned} E\left[\mathbf{EP3}\right] &= \mathbf{Q}_k^{\mathrm{noise}} E[\mathbf{n}^{\mathrm{ADC}} \left(\mathbf{n}^{\mathrm{ADC}}\right)^{\mathrm{T}}] \left[\mathbf{Q}_k^{\mathrm{noise}}\right]^{\mathrm{H}} \\ &= \sigma_{\mathrm{ADC}}^2 \mathbf{Q}_k^{\mathrm{noise}} \left[\mathbf{Q}_k^{\mathrm{noise}}\right]^{\mathrm{H}} \end{aligned} \tag{65}$$

Then

$$\begin{aligned} E\left[Z_k^{\mathrm{H}} Z_k\right] &= \mathbf{w}^{\mathrm{T}} 2\sigma_s^2 \left(\mathbf{H}_{\mathrm{u}}^{\mathrm{T}} \mathbf{V}_k \mathbf{V}_k^{\mathrm{H}} \mathbf{H}_{\mathrm{u}} + \mathbf{H}_{\mathrm{b}}^{\mathrm{T}} \mathbf{W}_k \mathbf{W}_k^{\mathrm{H}} \mathbf{H}_{\mathrm{b}}\right) \mathbf{w} \\ &+ \sigma_{\mathrm{AWGN}}^2 \mathbf{w}^{\mathrm{T}} \mathbf{Q}_k^{\mathrm{noise}} \left[\mathbf{Q}_k^{\mathrm{noise}}\right]^{\mathrm{H}} \mathbf{w} \\ &+ \mathbf{w}^{\mathrm{T}} \mathbf{Q}_k^{\mathrm{noise}} \Sigma_{\mathrm{NEXT}} \left[\mathbf{Q}_k^{\mathrm{noise}}\right]^{\mathrm{H}} \mathbf{w} \\ &+ \sigma_{\mathrm{ADC}}^2 \mathbf{w}^{\mathrm{T}} \mathbf{Q}_k^{\mathrm{noise}} \left[\mathbf{Q}_k^{\mathrm{noise}}\right]^{\mathrm{H}} \mathbf{w} \\ &+ \mathbf{w}^{\mathrm{T}} \frac{\sigma_{\mathrm{DNF}}^2}{\mathbf{w}^{\mathrm{T}}\mathbf{w}} \mathbf{I} \mathbf{w} \\ &= \mathbf{w}^{\mathrm{T}} \tilde{\mathbf{B}}_k \mathbf{w} \end{aligned} \tag{66}$$

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 $\mathbf{w}$.

G. Discussion of the maximization method leading to single TEQ design algorithm

Almogy and Levin intended to maximize a sum-of-ratios problem by solving for roots of a parametric problem as in

$$b = \max_{x \in R} \sum_{i=1}^{n} \frac{f_i(x)}{g_i(x)} \tag{67}$$

which they transformed into the parametric problem

$$H_n(\mathbf{q}) = \max_{x \in R} \sum_{i=1}^{n} [f_i(x) - q_i g_i(x)] \tag{68}$$

where, $\mathbf{q} = [q_1, q_2, \cdots, q_n]^{\mathrm{T}}$ and where by analogy to the single-ratio maximization Dinkelbach approach (which is the same as Lagrangian multipliers [26]), they defined

$$q_i = \frac{f_i(x_p)}{g_i(x_p)} \tag{69}$$

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 $H_1(q_1)$ where q_1 is maximized and the point x maximizing the current iteration is the same as x_p. $H_1(q_1)$ is a convex, non-increasing function of q_1 with a single root. In that problem q_1 increases with every iteration, while $H_1(q_1)$ decreases. By analogy to Dinkelbach method, Almogy and Levin solve $H_n(\mathbf{q}) = 0$ to find

the optimal solution of the sum-of-ratios problem. This approach is erroneous as shown by Falk and Palocsay [33] and validated in our experiments in using Almogy and Levin's iteration to maximize a version of (18). That said, function $H_n(\mathbf{q})$ is still a convex, non-increasing function of $\mathbf{q}$ with a single crossing [24]. In the sum-of-ratios problem, finding a zero of $H_n(\mathbf{q})$ does not produce b. However, Almogy and Levin's iteration does find the roots of the nonlinear $H_n(\mathbf{q})$ efficiently; hence, we use their idea with modifications specific to our problem, to find an root of the gradient of (18) that corresponds to the closest local maximum to the set (27). Our experimental results presented in the Section V support that. The gradient of (18) is

$$\frac{db_{\mathrm{DMT}}(\mathbf{w},\mathcal{I})}{d\mathbf{w}} = \frac{2}{\ln 2}\sum_{k\in\mathcal{I}} r_k(\mathbf{w})\left[\mathbf{A}_k^r - \lambda_k(\mathbf{w})\mathbf{B}_k^r\right]\mathbf{w} \tag{70}$$

where

$$r_k(\mathbf{w}) = \frac{1}{\mathbf{w}^{\mathrm{T}}\mathbf{A}_k\mathbf{w}} \text{ and } \lambda_k(\mathbf{w}) = \frac{\mathbf{w}^{\mathrm{T}}\mathbf{A}_k\mathbf{w}}{\mathbf{w}^{\mathrm{T}}\mathbf{B}_k\mathbf{w}} \tag{71}$$

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

$$\mathbf{C}_k(\mathbf{w}) = r_k(\mathbf{w})\left[\mathbf{A}_k^r - \lambda_k(\mathbf{w})\mathbf{B}_k^r\right] \text{ and } \mathbf{C}(\mathbf{w},\mathcal{I}) = \sum_{k\in\mathcal{I}}\mathbf{C}_k(\mathbf{w}) \tag{72}$$

This leads to an equation similar to (68) in which $r_k(\mathbf{w})$, $\lambda_k(\mathbf{w})$ and $\mathbf{C}_k(\mathbf{w})$ are projected according to (71) and (72) during the iterative procedure that finds the optimal root of (70). Let

$$\begin{aligned} H(\lambda) &= \max_{\mathbf{w}\in\mathcal{S}} \mathbf{w}^{\mathrm{T}}\mathbf{C}(\mathbf{w},\mathcal{I})\mathbf{w} \\ &= \max_{\mathbf{w}\in\mathcal{S}} \sum_{k\in\mathcal{I}} r_k\left[\mathbf{w}^{\mathrm{T}}\mathbf{A}_k^r\mathbf{w} - \lambda_k\mathbf{w}^{\mathrm{T}}\mathbf{B}_k^r\mathbf{w}\right] \end{aligned} \tag{73}$$

where $\lambda = [\cdots,\lambda_k,\cdots]^{\mathrm{T}}, k\in\mathcal{I}$. Now we see the similarities between (68) and (73). Our method uses Almogy and Levin's method as basis, but modified with specific requirements of our problem to find a zero of (73) corresponding to the closest maximum of (18). In the process, $\lambda_k(\mathbf{w})$ will always increase, thereby increasing (18).

REFERENCES

[1] ANSI, "Network and customer installation interfaces: Asymmetric digital subscriber line (ADSL) metallic interface," Tech. Rep. T1E1.413, American National Standards Institute, 1998.

[2] ITU-T, "Asymmetrical digital subscriber line (ADSL) transceivers," Tech. Rep. G.992.1, Int. Telecomm. Union, 1999.

[3] ITU-T, "Splitterless asymmetric digital subscriber line (ADSL) transceivers," Tech. Rep. G.992.2, Int. Telecomm. Union, 1999.

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

[5] VDSLalliance, "SDMT VDSL draft standard proposal," Tech. Rep. T1E1.4/97-332, American National Standards Institute, 1997.

[6] ETSI, "Transmission and multiplexing (TM); access transmission systems on metallic access cables; very high speed digital subscriber line (VDSL); part 2: Transceiver specification," Tech. Rep. TS 101 270-2, European Telecomm. Standards Institute, 1998.

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

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

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

[10] G. Arslan, B. L. Evans, and S. Kiaei, "Equalization for discrete multitone transceivers to maximize bit rate," *IEEE Trans. on Signal Proc.*, vol. 49, pp. 3123–3135, Dec. 2001.

[11] B. Farhang-Boroujeny and M. Ding, "An eigen-approach to design of near-optimum time domain equalizer for DMT tranceivers," in *Proc. IEEE Int. Conf. Comm.*, pp. 937–941, 1999.

[12] P. J. W. Melsa, R. C. Younce, and C. E. Rohrs, "Impulse response shortening for discrete multitone transceivers," *IEEE Trans. on Comm.*, vol. 44, pp. 1662–1672, Dec. 1996.

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

[14] G. Ysebaert, K. Vanbleu, G. Cuypers, and M. Moonen, "ADSL per-tone/per-group optimal equalizer and windowing design," in *Proc. IEEE Asilomar Conf. on Signals, Systems and Computers*, pp. 383–387, Nov. 2002.

[15] M. Milosevic, L. F. C. Pessoa, B. L. Evans, and R. Baldick, "DMT bit rate maximization with optimal time domain equalizer filter bank architecture," in *Proc. IEEE Asilomar Conf. on Signals, Systems and Computers*, (Pacific Grove, CA), pp. 377–382, Nov. 2002.

[16] M. Milosevic, L. F. C. Pessoa, and B. L. Evans, "Simultaneous multichannel time domain equalizer design based on the maximum composite shortening SNR," in *Proc. IEEE Asilomar Conf. on Signals, Systems and Computers*, (Pacific Grove, CA), pp. 1895–1899, Nov. 2002.

[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 Trans. on Comm.*, vol. 43, pp. 2595–2604, Oct. 1995.

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

[22] H. Stark and J. Woods, *Probability, Random Processes, and Estimation Theory for Engineers*, pp. 213–218. Prentice Hall, 2nd ed., 1994.

[23] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Kluwer Academic Publishers, 1992.

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

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

[26] D. P. Bertsekas, *Nonlinear Programming*, pp. 54–75, 253. Athena Scientific, 1995.

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

[28] Y. Almogy and O. Levin, "A class of fractional programming problems," in *Operations Research*, vol. 19, pp. 57–67, 1971.

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

[30] G. H. Golub and C. F. V. Loan, *Matrix Computations*. John Hoppkins, 3rd ed., 1996.

[31] G. Arslan, M. Ding, B. Lu, M. Milosevic, Z. Shen, and B. L. Evans, "Matlab DMT TEQ design toolbox," Tech. Rep. v3.2beta available at http://www.ece.utexas.edu/∼ bevans/projects/adsl/dmtteq/, Embedded Signal Processing Lab, The University of Texas at Austin, Mar. 2003.

[32] J. L. Devore, *Probability and Statistics for Engineering and the Sciences*, pp. 298–299. Brooks/Cole Publishing Company, 4 ed., 1995.

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

LIST OF FIGURES

LIST OF TABLES

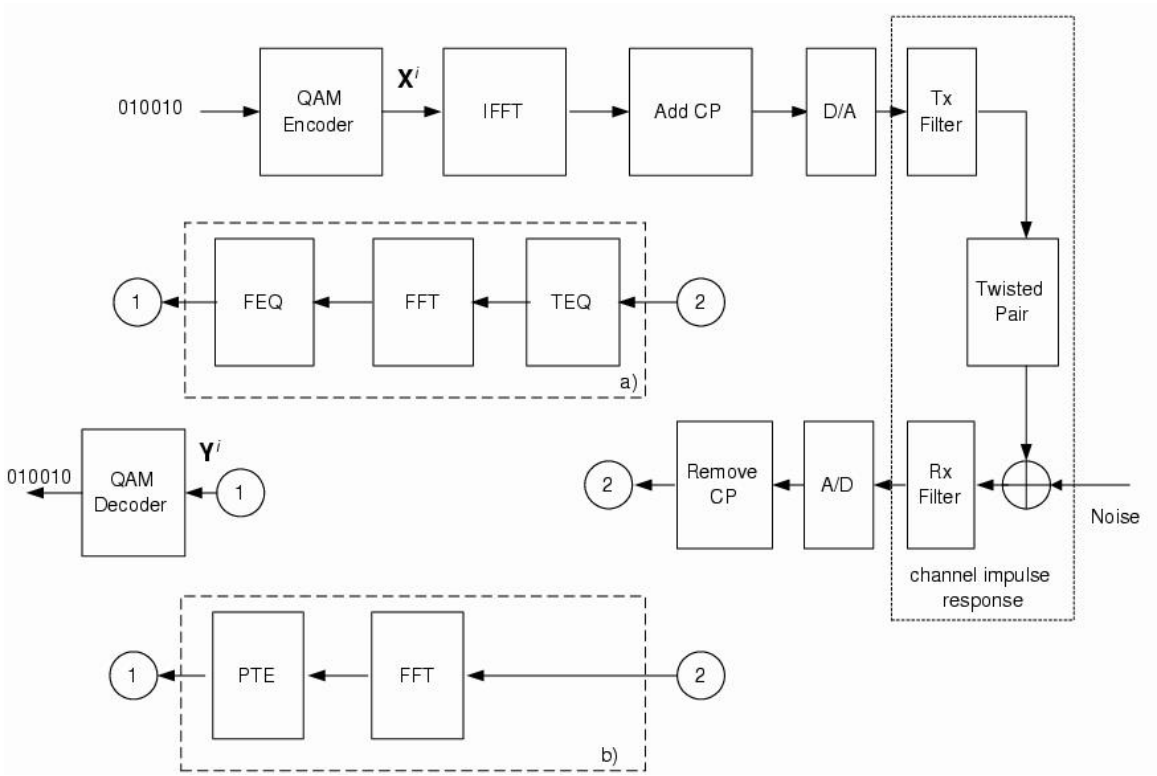

Fig. 1. Simplified block diagram of an ADSL DMT System. The receiver diagram shows two options for the equalization subsystem: a) time-domain TEQ (single or a filter bank) architecture and b) frequency-domain per-tone architecture. Input i^{th} data symbol X^i results in the estimated data symbol Y^i.

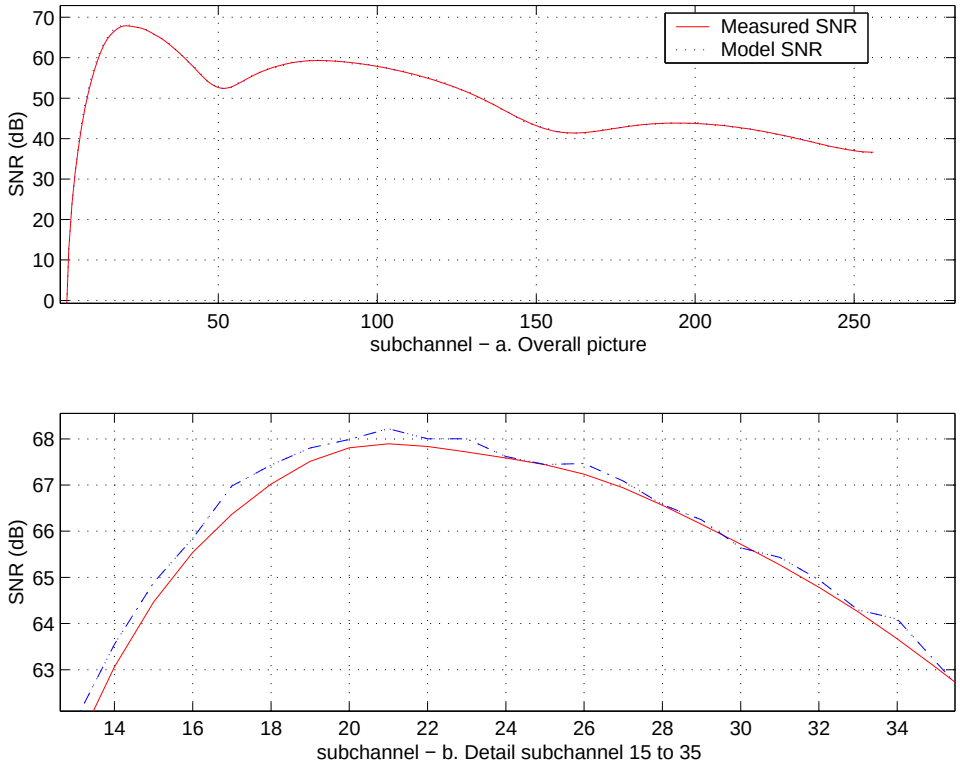

Fig. 2. Comparison of the measured SNR and the proposed SNR model for $M = 32$, CSA loop 2. The figure shows the closeness of the measured SNR and the proposed SNR model when most of the ISI has been removed.

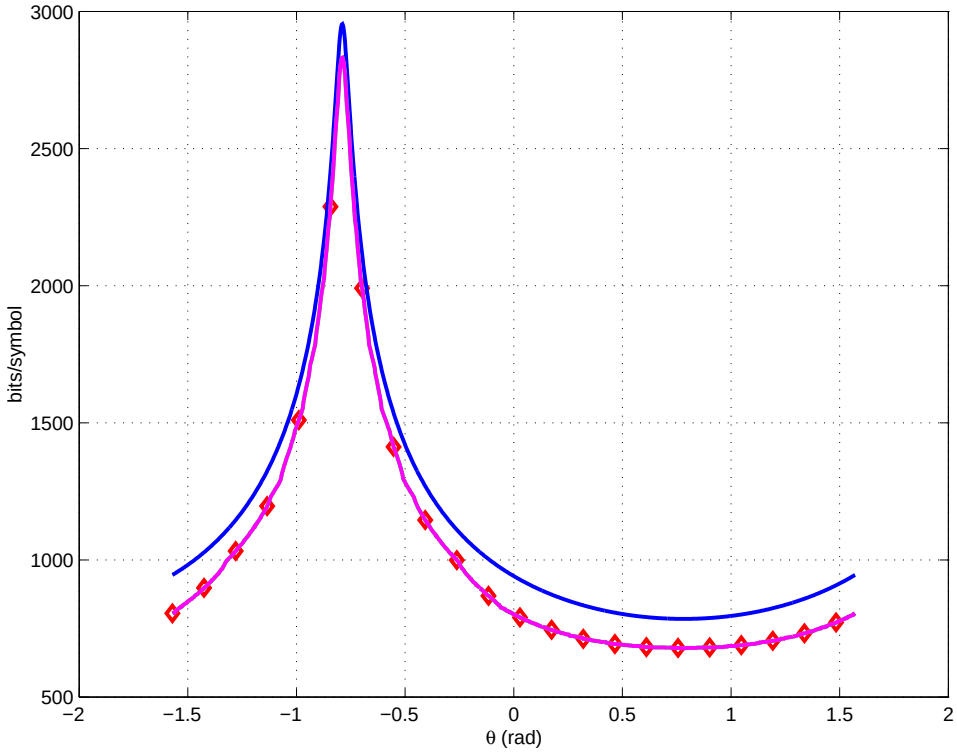

Fig. 3. Evaluation of Fractional Bit Rate b_{DMT} (solid) and Integer Bit Rate $b_{\mathrm{DMT}}^{\mathrm{int}}$ (Solid with Diamonds) for All Values of Time Domain Equalizer $\mathbf{w} = [\sin(\theta), \cos(\theta)]^{\mathrm{T}}, \theta = \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ of Length $M = 2$ for Carrier Service Area Loop 3 with Transmission Delay Δ=15, FFT Size $N = 512$, Cyclic Prefix Length $\nu = 32$, Input Power = 0.2472 W, Additive White Gaussian Noise Power =-140 dBm/Hz, Near-end Crosstalk modelled as 49 ADSL Disturbers and θ Sampled with 1081 Points. Notice That $b_{\mathrm{DMT}}^{\mathrm{int}}$ and b_{DMT} Share Global Maximizer.

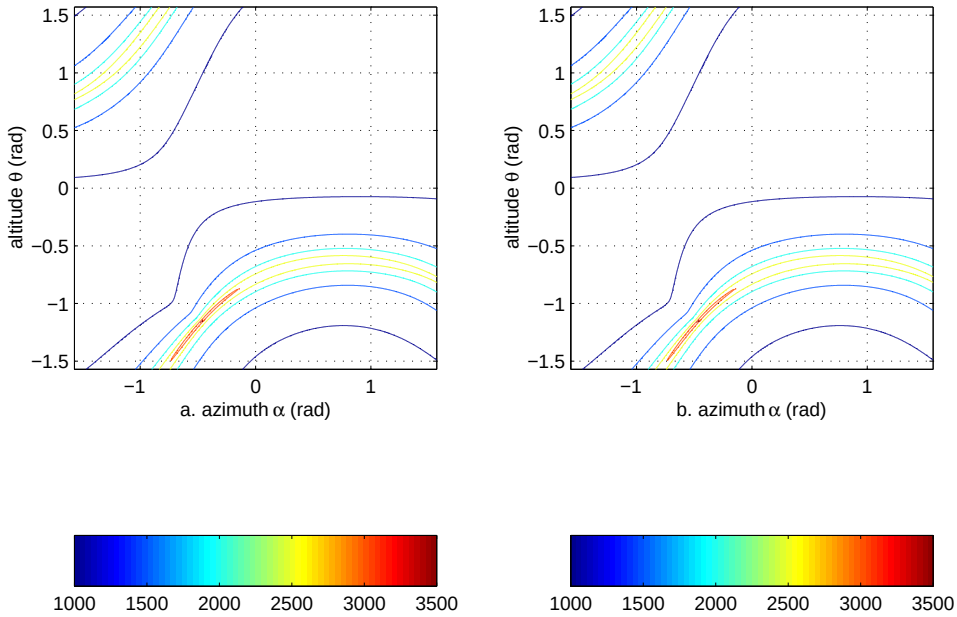

Fig. 4. The Evaluation of a. $b_{\mathrm{DMT}}^{\mathrm{int}}$ and b. b_{DMT} for All Values of $\mathbf{w} = [\sin(\alpha)\sin(\theta), \cos(\alpha)\sin(\theta), \cos(\theta)]^{\mathrm{T}}$ for Azimuth Angle $\alpha = \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ and Altitude Angle $\theta = \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ for Channel Impulse Response Containing Carrier Service Area Loop 3 with Time Domain Equalizer Length $M = 3$, with Transmission Delay Δ=15, FFT Size $N = 512$, Cyclic Prefix Length $\nu = 32$, Input Power = 0.2472 W, Additive White Gaussian Noise Power =-140 dBm/Hz, Near-end Crosstalk modelled as 49 ADSL Disturbers and θ Sampled with 1081 Points. The Maxima of Both b_{DMT} and b_{DMT}^{int} Occur for $\alpha = -0.4741$ rad and $\theta = -1.1606$ rad

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.

Fig. 5. Proposed single TEQ Design algorithm. For the initial point of the optimization algorithm, choose $\mathbf{w}_{\mathrm{opt}} = \mathbf{w} = \mathbf{w}_{k_{\mathrm{o}}}^{\mathrm{opt}}$. Define $k_{\mathrm{o}} = \arg\max_k \left\{b_{\mathrm{DMT}}(\mathbf{w}_k^{\mathrm{opt}}, \mathcal{I}), k \in \mathcal{I}\right\}$. Set iteration counter $i = 0$ and smoothing factor $\alpha = 0$.

TABLE I

DATA TRANSMISSION COMPUTATIONAL COMPLEXITY FOR SAMPLE RATE $f_s = 2.208$ MHz, SYMBOL RATE $f_{sym} = 4$ kHz, FFT SIZE $N = 512$, CYCLIC PREFIX LENGTH $\nu = 32$, TIME DOMAIN EQUALIZER FILTERS LENGTH $M = 2$ TO 32 AND ASSUMING $\frac{N}{2}$ DATA-CARRYING SUBCHANNELS.

single TEQ	**Real MACs**	**Words/Sym**
TEQ	Mf_s	$2M$
FFT	$2N \log_2 N f_{sym}$	$4N$
FEQ	$2Nf_{sym}$	$2N$
8-tap example	$59 * 10^6$	3088
TEQFB	**Real MACs**	**Words/Sym**
TEQ FB	$\frac{N}{2}Mf_s$	$M(1+\frac{N}{2})$
Goertzel FB	$(N^2+N)f_{sym}$	$4N$
FEQ	$2Nf_{sym}$	$2N$
8-tap example	$5577 * 10^6$	5128
PTE	**Real MACs**	**Words/Sym**
FFT	$2N \log_2 N f_{sym}$	$4N + 2\nu$
Sliding FFT	$2(M-1)Nf_{sym}$	N
Combiner	$2NMf_{sym}$	$(M+1)N$
8-tap example	$98 * 10^6$	7232

TABLE II

HIGHEST ACHIEVED BIT RATES FOR TIME DOMAIN EQUALIZER FILTER BANK ON STANDARD CARRIER SERVICE AREA LOOPS 1-8 WITH TRANSMIT AND RECEIVE HIGH-PASS FILTERS FOR FFT SIZE $N = 512$, CYCLIC PREFIX LENGTH $\nu = 32$ WITH INPUT POWER = 0.2472 W, ADDITIVE WHITE GAUSSIAN NOISE POWER =-140 DBM/HZ, NEAR-END CROSSTALK MODELLED AS 49 ADSL DISTURBERS AND OPTIMAL TRANSMISSION DELAY Δ. EACH TIME DOMAIN EQUALIZER FILTER BANK SUBCHANNEL TIME DOMAIN EQUALIZER IS OF LENGTH M. THERE ARE $N/2$ DATA-CARRYING SUBCHANNELS AND NO LIMITATIONS ON THE BIT ALLOCATION.

CIR containing	**Proposed TEQFB (Mbps)**	$\boldsymbol{\Delta}_{\mathrm{opt}}$	M
CSA loop 1	11.417	15	8
CSA loop 2	12.680	22	12
CSA loop 3	10.995	26	8
CSA loop 4	11.288	35	6
CSA loop 5	11.470	32	16
CSA loop 6	10.861	20	8
CSA loop 7	10.752	34	13
CSA loop 8	9.615	35	11

TABLE III

PERFORMANCE OF SIMULATED TIME DOMAIN EQUALIZER METHODS VS. TIME DOMAIN EQUALIZER FILTER BANK AVERAGED OVER TIME DOMAIN EQUALIZER LENGTH $M \in \{2, \cdots, 32\}$ FOR STANDARD CARRIER AREA SERVICE LOOPS 1-8 WITH TRANSMIT AND RECEIVE HIGH-PASS FILTERS FOR FFT SIZE $N = 512$, CYCLIC PREFIX LENGTH $\nu = 32$ WITH INPUT POWER = 0.2472 W, ADDITIVE WHITE GAUSSIAN NOISE POWER =-140 DBM/HZ, NEAR-END CROSSTALK MODELLED AS 49 ADSL DISTURBERS AND OPTIMAL TRANSMISSION DELAY Δ. THE DESIGN METHODS ARE 1-PROPOSED SINGLE TIME DOMAIN EQUALIZER, 2-MBR,3-MIN-ISI,4-LS PTE,5-MMSE-UEC,6-MMSE-UTC.

	Algorithm					
CIR of CSA loop	**1**	**2**	**3**	**4**	**5**	**6**
$LoopAvg(1, Algorithm)$	99.6%	97.3%	97.5%	99.5%	86.3%	84.4%
$LoopAvg(2, Algorithm)$	99.6%	97.0%	97.3%	99.5%	87.2%	85.8%
$LoopAvg(3, Algorithm)$	99.6%	97.8%	97.3%	99.6%	83.9%	83.0%
$LoopAvg(4, Algorithm)$	99.3%	98.1%	98.2%	99.1%	81.9%	81.5%
$LoopAvg(5, Algorithm)$	99.6%	97.7%	97.2%	99.5%	88.6%	88.9%
$LoopAvg(6, Algorithm)$	99.5%	97.7%	98.3%	99.4%	82.7%	79.8%
$LoopAvg(7, Algorithm)$	98.8%	96.3%	96.3%	99.6%	75.8%	78.4%
$LoopAvg(8, Algorithm)$	98.7%	97.4%	97.5%	99.2%	82.6%	83.6%
$Average(Algorithm)$	**99.3%**	**97.4%**	**97.5%**	**99.4%**	**83.6%**	**83.2%**

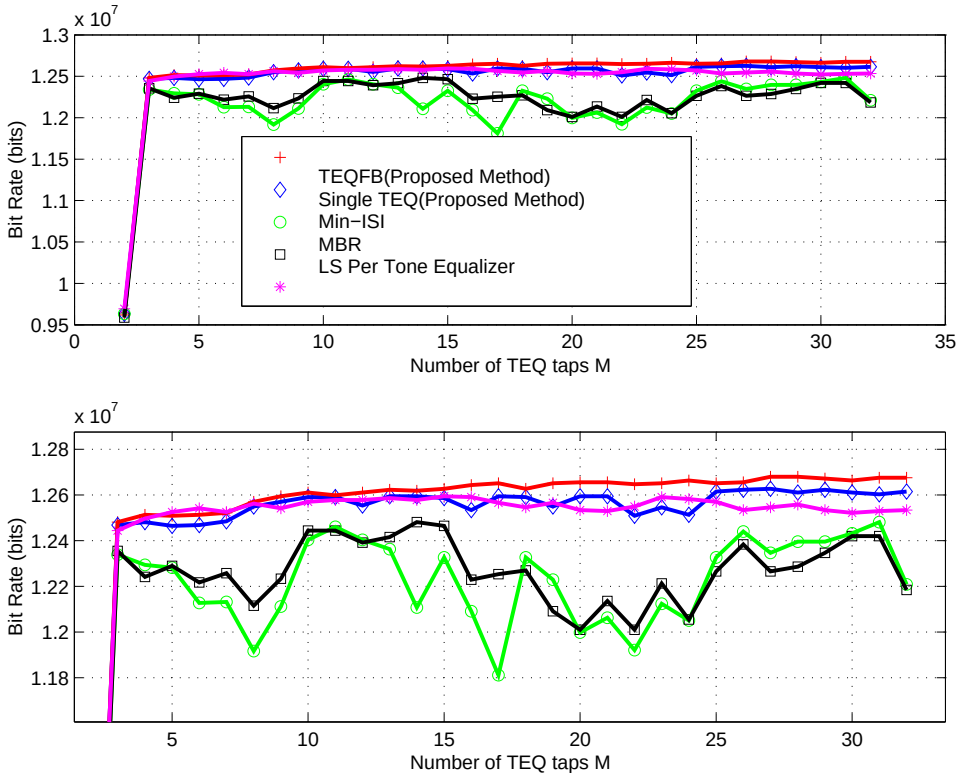

Fig. 6. Simulation Data Rates Achieved on Carrier Service Area Loop 2 vs. the Length of Time Domain Equalizer for FFT size $N = 512$, Cyclic Prefix Length $\nu = 32$ with Input Power = 0.2472 W, Additive White Gaussian Noise Power =-140 dBm/Hz, Near-end Crosstalk Modelled as 49 ADSL Disturbers and Optimal Transmission Delay Δ. a. Entire Graph b. Detail of Higher Data Rates. Reported Rates are up to ± 60 Kbps Away from True Values.

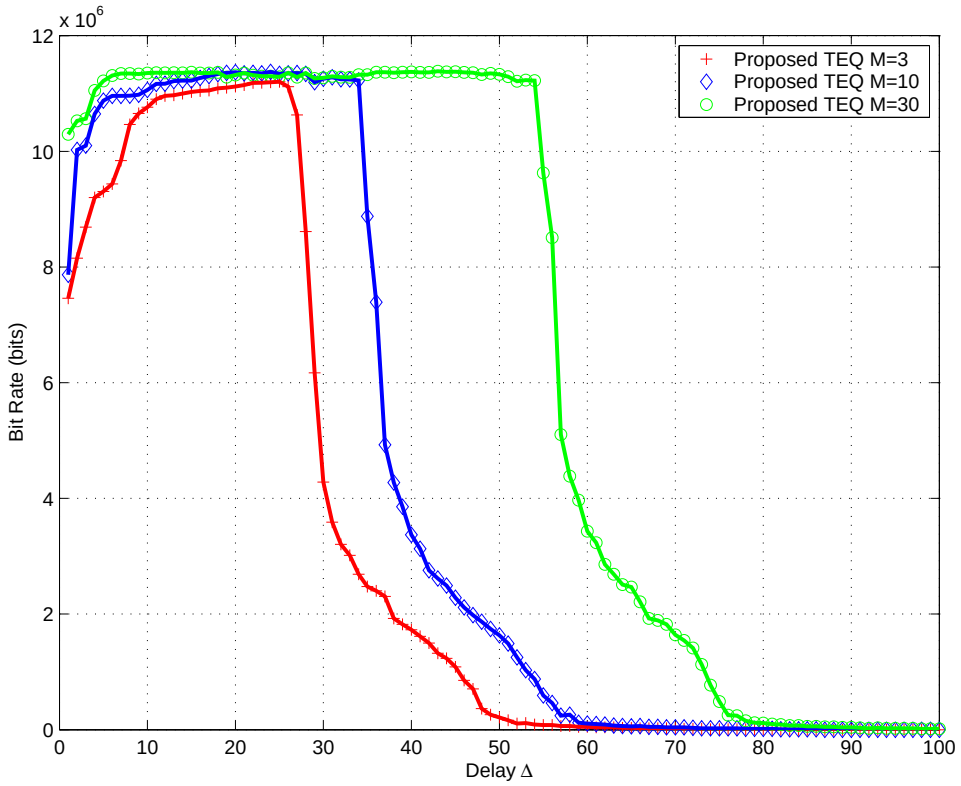

Fig. 7. Simulation Data Rates of the Proposed Single Time Domain Equalizer Design Achieved on Carrier Service Area Loop 1 vs. Transmission Delay Design Parameter for FFT size $N = 512$, Cyclic Prefix Length $\nu = 32$ with Input Power = 0.2472 W, Additive White Gaussian Noise Power =-140 dBm/Hz, Near-end Crosstalk Modelled as 49 ADSL Disturbers.

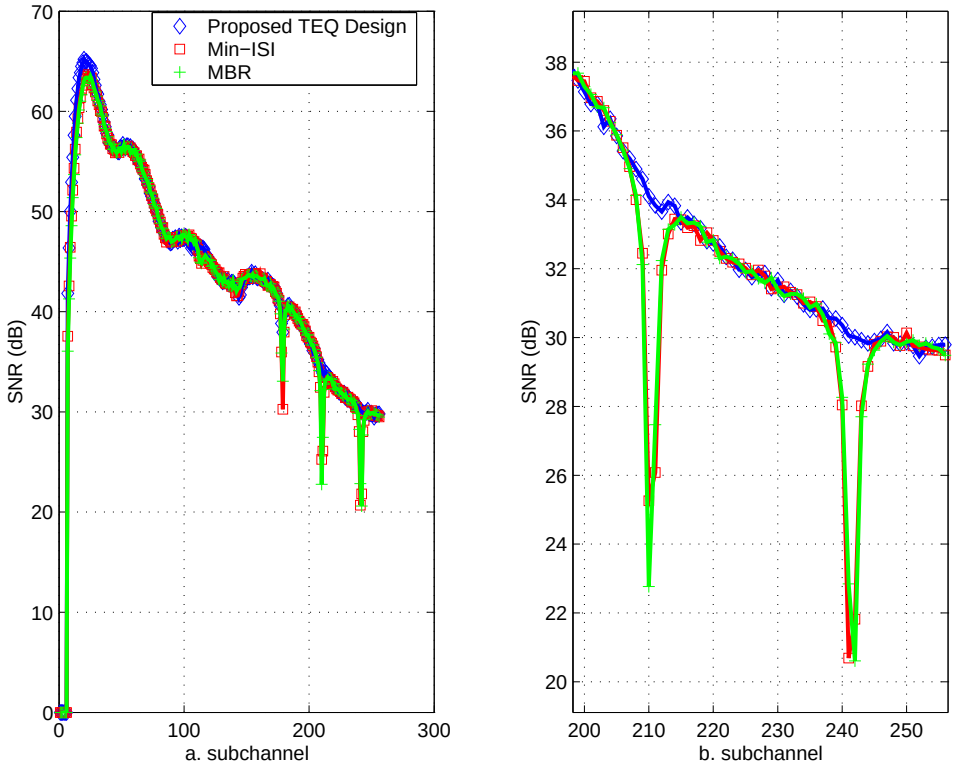

Fig. 8. Signal-to-noise Ratio of Proposed Single Time Domain Equalizer Design vs. the Maximum Bit Rate Method and Minimum-ISI Method for Carrier Area Service Loop 4 for Time Domain Equalizer Length $M = 18$, FFT size $N = 512$, Cyclic Prefix Length $\nu = 32$ with Input Power = 0.2472 W, Additive White Gaussian Noise Power =-140 dBm/Hz, Near-end Crosstalk Modelled as 49 ADSL Disturbers and Optimal Transmission Delay Δ; Maximum Bit Rate Method and Minimum-ISI Methods Place Nulls in the Passband Signal-to-noise Ratio while Proposed Single Time Domain Equalizer Design Does Not. Nulls Limit Data Rate. a. Signal-to-noise Ratio, b. Detail of Nulls