Discrete Multitone Echo Cancelation

Minnie Ho, *Member, IEEE*, John M. Cioffi, *Senior Member, IEEE*, and John A. C. Bingham, *Fellow, IEEE*

***Abstract*— Multicarrier transmission methods have long been known to optimize the performance of data transceivers on bandlimited communication channels. One form of multicarrier transmission, known as discrete multitone modulation (DMT), is particularly attractive for its ability to be implemented using efficient digital signal processing techniques. Given a basic DMT system, it is possible to increase the aggregate data rate with full-duplex transmission using echo cancelation. However, DMT echo cancelation at first appears difficult because of the computational complexity required in a straightforward implementation to cancel the cross-echoes produced by each carrier into every carrier.**

This paper presents high-speed echo cancelation techniques for full-duplex data transmission using DMT systems. The techniques estimate the echo with a method of fast convolution that combines a complex multiply per tone with a short convolution in the time domain. In addition, the frequency-domain update of the echo parameters consists of one complex multiply per tap. As a result, these techniques can achieve much lower complexity than that required by traditional single-carrier technique, such as the least mean square (LMS) algorithm. This approach has been implemented for asymmetric digital subscriber line (ADSL) applications.

I. Introduction

In multicarrier modulation, the transmission channel is partitioned into a bank of orthogonal, memoryless subchannels. Data is transmitted through each subchannel independently of the other subchannels, allowing the flexible allocation of bits and power to each subchannel. Such methods have long been known to optimize the performance of data transceivers on bandlimited Gaussian channels with memory [1]. With the development of efficient multicarrier transceivers, variations of these methods have become more popular in recent years, resulting in two standards: discrete multitone modulation (DMT) for the asymmetric digital subscriber line (ADSL) standard in the United States [2] and coded orthogonal frequency division multiplexing (Eureka-147) for the Digital Audio Broadcast standard in Europe [3]. Multicarrier modulated signals appear Gaussian (or "noise-like") and have wide dynamic ranges of nearly continuous values, unlike those of more traditional single-carrier modulated signals. This leads to fundamentally optimum performance, but presents a number of interesting challenges to efficient transceiver design. One such challenge involves the design of a low-complexity echo canceler.

Echo cancelation separates two signals traveling simultaneously in opposite directions in the same channel over the same frequency bands. This precludes the need for multiplexing the two directions of transmission, thereby increasing the total data rate, and also precludes the need for two physically distinct channels, thereby lowering costs. Instead, separation is performed by estimating and canceling the echo produced by the leakage of the transmitted signal into the collocated receiver. Echo cancelers typically exploit structure in the transmitted signal to reduce computational complexity. At first glance, the multicarrier signal has no such structure to exploit, and this leads to the erroneous conclusion that echo cancelation is hard or impossible with multicarrier systems.

This conclusion was partially dispelled in a recent study [4] that presented one structure for echo cancelation for an efficient form of multicarrier modulation known as DMT. This structure is efficient for low-speed (sampling rates below 10 kHz) voiceband data modem transmission and exploits a dual time and frequency-domain architecture where canceler settings are updated in the frequency domain but echos are synthesized completely in the time-domain. The time-domain echo synthesis is used to circumvent the cross-echoes produced by each tone in a DMT signal into every tone. However, at higher sampling rates, the implementation cost of a finite impulse response (FIR) time-domain filter that synthesizes the echo becomes prohibitive.

This paper presents a new DMT echo cancelation structure that avoids most of the time-domain operation and cancels the echo in the frequency domain directly. Time-domain processing is used only to make the echo appear periodic so that cross-echoing does not occur. This time-domain processing costs significantly less than the processing required for full synthesis of the time-domain echo, leading to a much reduced complexity compared to the DMT echo canceler in [5].

II. A DMT Structure for Echo Cancelation

In a DMT transceiver, modulation and demodulation are performed by an inverse fast Fourier transform (IFFT) and a fast Fourier transform (FFT), respectively [6], [7]. After modulation, the last L samples of each real-valued data vector are copied and prefixed to the vector, where L is the length of the impulse response of the channel. The cyclic prefix operation, along with the IFFT modulation, transforms the channel into a bank of parallel, independent subchannels, or contiguous frequency bins. Given a total power constraint on the channel, data rates close to capacity can be achieved

Paper approved by J. H. Winters, the Editor for Equalization of the IEEE Communications Society. Manuscript received August 25, 1994; revised July 4, 1995. M. Ho was supported by a gift from Apple and by an AT&T Bell Laboratories Grant. This paper was presented at the International Conference on Communications, Geneva, Switzerland, May 1993.

M. Ho is with Radix Technologies, Mountain View, CA 94043 USA.

J. M. Cioffi is with the Information Systems Laboratory, Stanford University, Stanford, CA 94305 USA.

J. A. C. Bingham is with Amati Communications Corporation, Mountain View, CA 94301 USA.

Publisher Item Identifier S 0090-6778(96)05512-2.

0090–6778/96$05.00 © 1996 IEEE

by optimizing the allocation of bits and power among the independent subchannels using water-pour techniques [6], [8]. Furthermore, the application of trellis coding and shaping across the subchannels achieves data rates even closer to capacity [9].

Echo is created when the cyclically-prefixed data vectors are reflected back through the hybrid circuit, a four-to-two wire interface that comprises a transformer and a balanced load to match the average impedance of the various lines, into the near-end receiver. A typical analog hybrid generally provides less than 20 dB of average return loss from the transmitter to the near-end receiver [10], nowhere near the 70 dB required for echo cancelation in data transmission.

This paper models the echo as the linear convolution of a long FIR, $\boldsymbol{h}$, with the sequence of cyclically prefixed data symbols, $\boldsymbol{x}_i$. For simplicity, the FIR $\boldsymbol{h}$ is assumed to be of length $M \leq N$ samples, where N is the number of samples per symbol in a practical DMT system. The techniques are easily generalized for the case when $M > N$. The $N \times 1$ vector $\boldsymbol{H}$ denotes the N-point FFT of $\boldsymbol{h}$, and $\boldsymbol{h}$ will denote either the N-point zero-padded FIR or the M-point unpadded FIR, depending on the context.

The N-point block of echo channel output samples is: $\boldsymbol{y}_i = \boldsymbol{T}_{i,i-1}\boldsymbol{h} + \boldsymbol{n}_i$, where $\boldsymbol{n}_i$ represents distorted data from the far-end transmitter plus white Gaussian noise and the $N \times N$ Toeplitz matrix $\boldsymbol{T}_{i,i-1}$ is composed of samples $x_i(j)$ from the input data symbols $\boldsymbol{x}_i$ and $\boldsymbol{x}_{i-1}$, see equation at the bottom of the page.

In the special case for which M the length of $\boldsymbol{h}$ is less than or equal to L the length of the cyclic prefix plus one, the linear convolution becomes a circular convolution, and the following circulant matrix operator $\mathcal{C}_i$ is equivalent to $\boldsymbol{T}_{i,i-1}$, in the sense that it generates the same N output samples

$$\mathcal{C}_i = \begin{bmatrix} x_i(1) & x_i(N) & \cdots & x_i(2) \\ \vdots & \ddots & \ddots & \vdots \\ x_i(N-1) & & & x_i(N) \\ x_i(N) & x_i(N-1) & \cdots & x_i(1) \end{bmatrix}. \quad (1)$$

Circular convolution is equivalent to an element-by-element multiply of the frequency-domain data symbol $\boldsymbol{X}_i$ with $\boldsymbol{H}$.

For each symbol, the echo canceler first estimates the echo by convolving the transmit data with $\boldsymbol{w}_i$, an estimate of $\boldsymbol{h}$. The estimated echo is then subtracted from the sampled received signal, and the resulting echo-canceled signal is used to adaptively update $\boldsymbol{w}_i$. The next two sections describe our DMT low-complexity techniques for estimating the echo and for computing the adaptive update.

A. Fast Convolution for Echo Estimation

For the special case in which $M \leq L + 1$, echo estimation can be performed with a simple operation of extremely low complexity. We first suppose the special case holds for the given application. The general case merely requires the addition of a second low-complexity operation.

Using the notation introduced in the previous section, the noiseless output of the echo channel is $\boldsymbol{y}_i = \mathcal{C}_i\boldsymbol{h}$. Taking the FFT of both sides gives $\boldsymbol{Y}_i = \text{diag}(\boldsymbol{X}_i)\boldsymbol{H}$, where $\text{diag}(\boldsymbol{X}_i)$ denotes the diagonal matrix whose diagonal is formed by the elements of the current data symbol, $\boldsymbol{X}_i$. This is exactly the basic idea behind DMT, since the channel has now been partitioned into a bank of parallel, orthogonal subchannels, except that here the channel is the echo channel instead of the transmission channel.

Therefore, to estimate the echo, the echo canceler simply performs a frequency-domain element-by-element multiplication, expressed as $\text{diag}(\boldsymbol{X}_i)\boldsymbol{W}_i$, where $\boldsymbol{W}_i$ is an adaptively updated estimate of $\boldsymbol{H}$. In the time-domain, this corresponds to the circular convolution of $\boldsymbol{x}_i$ with $\boldsymbol{w}_i$. In this way, the original long linear convolution reduces to an equivalent operation requiring one complex multiply per subchannel, so that the complexity of echo estimation is extremely low. Furthermore, the technique takes advantage of the structure of the DMT system, since $\boldsymbol{Y}_i$ and $\boldsymbol{X}_i$ are already provided by demodulation and modulation. Extra FFT's are not required, and the data is already blocked into frames.

The assumption of the special case rarely holds, and, in general, $M > (L + 1)$. In this case, the circulant matrix $\mathcal{C}_i$ is no longer valid, and the echo is $\boldsymbol{y}_i = \boldsymbol{T}_{i,i-1}\boldsymbol{h}$. Taking the FFT of both sides gives $\boldsymbol{Y}_i = \mathcal{F}_N\boldsymbol{T}_{i,i-1}\boldsymbol{h}$, where $\mathcal{F}_N$ is the matrix representing an N-point FFT. The echo channel can no longer be decomposed into a bank of parallel independent subchannels. Interference in a particular subchannel at time i arises not only from all the subchannels at time i, but also from every subchannel at time $i - 1$.

In the general case, the noiseless echo output is

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

The nonzero upper triangular Toeplitz portion of the matrix $\chi_{i,i-1} = \boldsymbol{T}_{i,i-1} - \mathcal{C}$ represents a short time-domain convolution operator that compensates for the difference between a

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

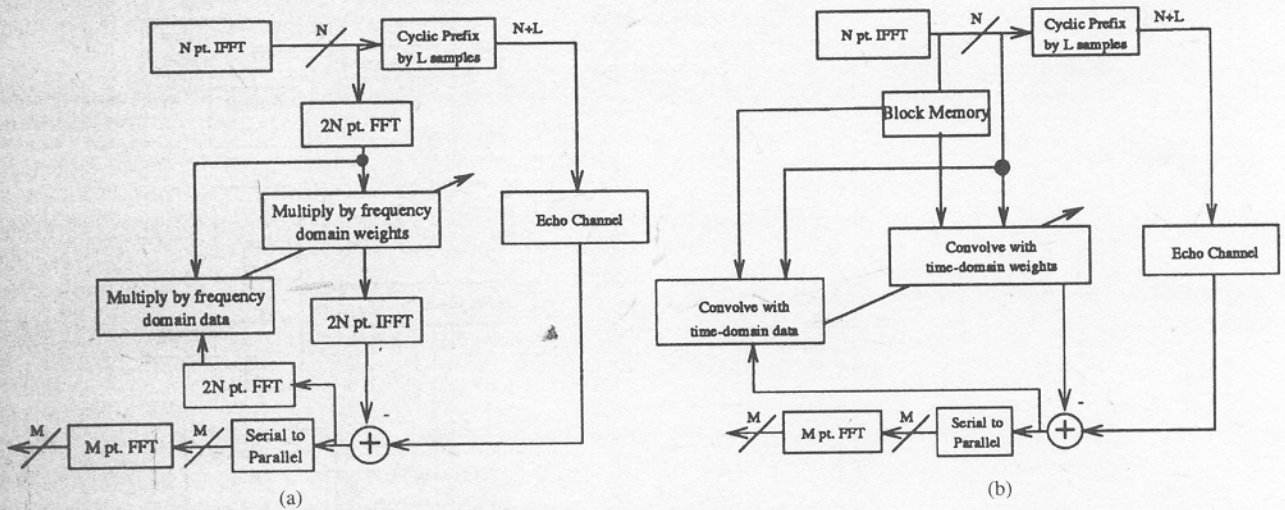


Fig. 1. (a) LMS echo cancelation. (b) FBLMS echo cancelation.

linear and a circular convolution of the input data with the echo parameters. Replacing $\boldsymbol{h}$ by $\boldsymbol{w}_i$ and $\boldsymbol{H}$ by $\boldsymbol{W}_i$, gives the two-part technique for fast echo estimation. The equation shows that the closer L is to M, the smaller the nonzero portion of $\chi_{i,i-1}$, and the fewer the computations needed to perform echo estimation.

Traditional techniques such as the least mean square (LMS) algorithm [11] or the frequency-domain version of LMS [12] are generally described with estimation performed solely in the time-domain or solely in the frequency domain. In the DMT techniques, the estimation is split between the two domains, lowering the complexity of each portion and thereby lowering the total complexity. Although this is a technique for fast convolution, it is not the same as the overlap–add technique [13] because we subtract the effects of the current time-domain vector for each block. Neither is it the same as overlap–save since we do not precompute the effects of the previous time-domain vector.

Once estimation is performed, the echo canceler subtracts the time-domain portion from the received signal prior to demodulation and subtracts the frequency-domain portion of the estimated echo from the demodulated received signal. In this way, the FFT that appears in (2) is exactly the same as the FFT used for demodulation; there is no need for an extra FFT. The ith vector of echo-canceled frequency-domain samples, $\boldsymbol{E}_i$, is

$$\boldsymbol{E}_i = \mathcal{F}_N(\boldsymbol{y}_i - \chi_{i,i-1}\boldsymbol{w}_i) - \operatorname{diag}(\boldsymbol{X}_i)\boldsymbol{W}_i \qquad (2)$$

where $\boldsymbol{y}_i$ is the ith received symbol vector prior to demodulation.

B. Frequency-Domain Adaptive Update

The DMT technique uses the following adaptive update: $\boldsymbol{W}_{i+1} = \boldsymbol{W}_i + \mu \operatorname{diag}(\boldsymbol{X}_i^*)\boldsymbol{E}_i$, where μ is the adaptive step-size and $*$ denotes complex conjugation. The time-domain echo parameters $\boldsymbol{w}_i$ are derived from $\boldsymbol{W}_i$ with an infrequently performed IFFT, so $\boldsymbol{w}_i$ is implicitly rather than explicitly updated. Also, assuming the maximum length of $\boldsymbol{h}$ to be $M < N$, we window $\boldsymbol{w}$ by setting the last $N - M$ elements to be zero.

To reflect these steps, (2) is

$$\boldsymbol{E}_i = \boldsymbol{Y}_i - (\mathcal{F}_N \chi_{i,i-1} \mathcal{F}_N^{-1} + \operatorname{diag}(\boldsymbol{X}_i))\boldsymbol{W}_i = \boldsymbol{Y}_i - \boldsymbol{A}_{i,i-1}\boldsymbol{W}_i$$

where $\boldsymbol{A}_{i,i-1} = \mathcal{F}_N \boldsymbol{T}_{i,i-1}$. The windowing of $\boldsymbol{w}$ is represented by setting the last $N-M$ columns of $\chi_{i,i-1}$ and $\boldsymbol{T}_{i,i-1}$ to zero. For simplicity, the same matrix notation is used to denote the windowed matrices.

When applying LMS methods, the update is formed by the gradient of the deterministic mean-squared error between the echo estimate and the actual echo. In the notation above, the update would be $\boldsymbol{W}_{i+1} = \boldsymbol{W}_i + \mu \boldsymbol{A}_{i,i-1}^* \boldsymbol{E}_i$. As a result, the DMT update is of considerably lower complexity, since including the full matrix $\boldsymbol{A}_{i,i-1}$ entails a convolution of the data with the error. This simpler update is not without its price; although the weights still converge, the rates of convergence are slower, as will be shown in Section VI. For ADSL applications, this trade-off is more than acceptable. For a pictorial comparison of the LMS, FBLMS, and DMT techniques, see Figs. 1 and 2.

III. Multirate DMT Echo Cancelation

In ADSL applications, higher data rates are required from the central office to the customer, with a lower rate reverse channel. Correspondingly, the typical transmit bandwidth at the remote-terminal or customer end (ADSL-RT) is a factor of k smaller than the receive bandwidth. If the symbol rates are equal in the two directions, then the received data symbols have k times more samples (subchannels) than the transmitted symbols. In this case, the basic techniques need only be slightly modified to handle the asymmetric multirate situation, and the following two sections describe the necessary modifications. We assume that k is an integer, and the symbol rate is the same for transmission as well as reception, so the transmit and receive carriers coincide.

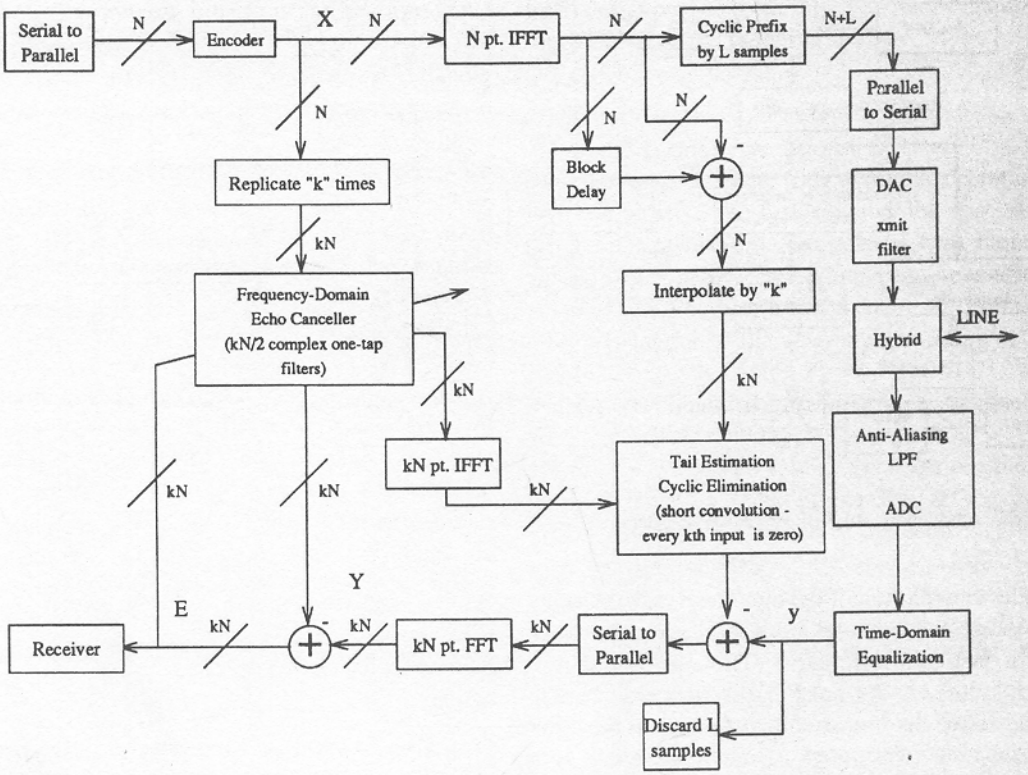


Fig. 2. ADSL-RT DMT echo cancelation.

A. Interpolated Echo Cancelation

When the transmit bandwidth is a factor of k smaller than the receive bandwidth, there is leakage of the higher frequency translates of the narrow-band transmit signal spectrum into the wide-band received signal, since the low-pass narrow-band reconstruction filter in the DAC is nonideal. Therefore, estimation of the echo requires the echo canceler to first interpolate the transmit data to the higher receive rate before performing the convolution.

In the frequency-domain, the data symbol is replicated k times, forming a block vector of k contiguous vectors. This operation can be represented as $\boldsymbol{R}\boldsymbol{X}_i$, where $\boldsymbol{R}$ denotes a block matrix of k vertically stacked $N \times N$ identity matrices. The time-domain data is interpolated by padding $k-1$ zeros between consecutive samples of $\boldsymbol{x}_i$. This time-domain operation is equivalent to that of aliasing in the frequency-domain and is the "interpolate by k" operation in Fig. 2. The upper triangular Toeplitz matrix $\chi_{i,i-1}$ is now composed of interpolated data.

Combining the two operations, we form the following echo-canceled frequency-domain received vectors[1]:

$$\boldsymbol{E}_i = \mathcal{F}_{kN}(\boldsymbol{y}_i - \chi_{i,i-1}\mathcal{F}_{kN}^{-1}\boldsymbol{W}_i) - \operatorname{diag}(\boldsymbol{R}\boldsymbol{X}_i)\boldsymbol{W}_i.$$

The $kN/2$ complex Hermitian-symmetric weights are updated using the interpolated data symbol, as follows:

$$\boldsymbol{W}_{i+1} = \boldsymbol{W}_i + \mu \operatorname{diag}(\boldsymbol{R}\boldsymbol{X}_i^*)\boldsymbol{E}_i.$$

[1] Note that we use FFT and IFFT's of size kN because the received vectors are of length $kN \times 1$.

B. Decimated Echo Cancelation

When the transmit bandwidth is a factor of k larger than the receive bandwidth, aliasing occurs when the relatively wide-band analog echo is resampled by the low-rate ADC, since the low-pass narrow-band anti-aliasing filter in the ADC is nonideal. Therefore, the echo canceler must add the effects of aliasing to the estimated echo just prior to canceling it from the narrow-band received signal.

In the frequency-domain, the output vectors are blocked into k contiguous vectors of N points each. These smaller vectors are then added together. This operation is labeled as the "block and add" operation of Fig. 3 and is represented by $\boldsymbol{R}' \operatorname{diag}(\boldsymbol{X}_i)\boldsymbol{W}_i$, where $\boldsymbol{R}'$ is a matrix of k horizontally stacked $N \times N$ identity matrices.

The time-domain portion of the estimated echo is decimated by computing every kth point of the convolution. This time-domain operation is equivalent to that of adding aliasing in the frequency-domain, and the matrix $\chi_{i,i-1}$ is decimated accordingly.

Combining the two operations, we form the following echo-canceled frequency-domain received vectors:

$$\boldsymbol{E}_i = \mathcal{F}_N(\boldsymbol{y}_i - \chi_{i,i-1}\mathcal{F}_{kN}^{-1}\boldsymbol{W}_i) - \boldsymbol{R}' \operatorname{diag}(\boldsymbol{X}_i)\boldsymbol{W}_i.$$

The frequency-domain adaptive update is

$$\boldsymbol{W}_{i+1} = \boldsymbol{W}_i + \mu \operatorname{diag}(\boldsymbol{X}_i^*)\boldsymbol{R}\boldsymbol{E}_i.$$

In this update, the frequency-domain error signal is interpolated to be of size $kN \times 1$, the same size as $\boldsymbol{W}_i$ and the data symbols $\boldsymbol{X}_i$.

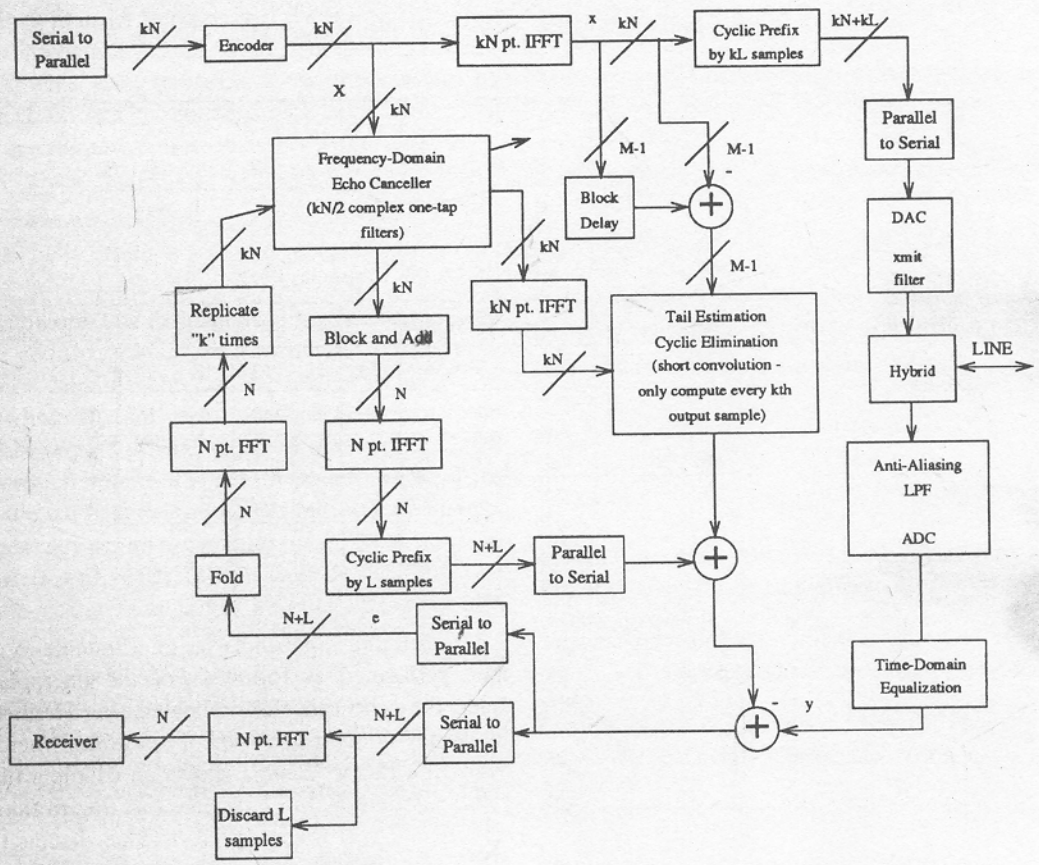


Fig. 3. ADSL-CO DMT echo cancelation.

IV. Frame-Asynchronous Cancelation

In frame-synchronous transmission, the transmitted and received symbols, or frames, are aligned. In general, only one transceiver can be frame-synchronous, since the frames at the other transceiver will be affected by the propagation delay of the transmission channel. In this event, we need to consider frame-asynchronous echo cancelation.

The frame-synchronous echo canceler can easily be modified to be frame-asynchronous as follows. To compensate for the misalignment of the vectors, we add an extra FFT and an extra IFFT (see Fig. 3). The IFFT transforms the frequency-domain portion of the estimated echo to the time-domain so that this portion can be canceled serially. The FFT converts the time-domain echo-canceled signal to the frequency-domain, since the echo-canceled demodulated symbols can no longer be used to compute the gradient.

We also need to expand the echo estimation to include computing the echo for the cyclic prefix samples as follows.

The frequency-domain portion of echo estimation goes through the added IFFT operation previously discussed, and a cyclic prefix is appended. In the time-domain portion of the echo estimation, the convolution is expanded to include the cyclic prefix. With these modifications, the received data symbols are decoupled from the transmitted symbols. Younce *et al.* [14] propose an alternative method for frame-asynchronous operation that involves knowledge of amount of slippage between the two data symbol streams. However, the overhead involved in tracking the amount of slippage is high and more than offsets the lower computational complexity that is achieved.

We apply these modifications to the echo canceler structures described in the preceding sections. As an example, we choose the decimated echo canceler to be a frame asynchronous echo canceler. First, we transform the block frequency-domain portion of the estimated echo into time-domain vectors and add a cyclic prefix. The output is $\boldsymbol{P}\mathcal{F}_N^{-1}\boldsymbol{R}'\ \mathrm{diag}(\boldsymbol{X}_i)\boldsymbol{W}_i$. In this equation, $\boldsymbol{P} = [\boldsymbol{e}_{N-L}, \boldsymbol{e}_{N-L+1}, \cdots \boldsymbol{e}_N, \boldsymbol{I}_N]'$ represents the prefix operation, and $\boldsymbol{e}_i$ are the ith basis vectors of size $N \times 1$.

Next, we expand the nonzero portion of the matrix $\chi_{i,i-1}$ by kL columns and L rows. Finally, the echo-canceled $N+L \times 1$ length time-domain vectors are

$$\boldsymbol{e}_i = \boldsymbol{y}_i - \boldsymbol{P}\mathcal{F}_N^{-1}\boldsymbol{R}'\ \mathrm{diag}(\boldsymbol{X}_i)\boldsymbol{W}_i - \chi_{i,i-1}\mathcal{F}_{kN}^{-1}\boldsymbol{W}_i.$$

As before, the adaptive update is

$$\boldsymbol{W}_{i+1} = \boldsymbol{W}_i + \mu\ \mathrm{diag}(\boldsymbol{X}_i^*)\boldsymbol{R}\mathcal{F}_N\boldsymbol{P}'\boldsymbol{e}_i.$$

For ADSL applications, choosing the decimated (ADSL-CO) echo canceler to be the frame-asynchronous end adds less total complexity, since the extra IFFT/FFT pair is of smaller size at the CO end. Also, the line impedance seen from the central office is, in general, much simpler to balance since there are more bridged taps near the customer than the central office. Therefore, fewer time-domain taps are required to model the echo impulse response, so that even if the

TABLE I
EXAMPLE INITIALIZATION SEQUENCE FOR THE INTERPOLATED EC

Symbol	Sample Number															
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	R_0	R_1	R_2	R_1	R_0	R_1	R_2	R_1	R_0	R_1	R_2	R_1	R_0	R_1	R_2	R_1

TABLE II
EXAMPLE INITIALIZATION SEQUENCE FOR THE DECIMATED EC

Symbol	Sample Number															
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	R_0	R_1	R_2	0	0	0	0	0	0	0	0	0	0	0	R_2	R_1
1	0	0	I_2	I_3	I_4	0	0	0	0	0	0	0	$-I_4$	$-I_3$	$-I_2$	0
2	0	0	0	0	R_4	R_5	R_6	0	0	0	R_6	R_5	R_4	0	0	0
3	0	0	0	0	0	0	I_6	I_7	R_8	$-I_7$	$-I_6$	0	0	0	0	0

rates were symmetric, a frame-asynchronous CO echo canceler would reduce the complexity seen by the customer.

V. INITIALIZATION

Convergence of stochastic gradient-approximation algorithms is slow, and fast initialization methods for the echo cancelers are needed to reduce start-up time. During start-up, the echo canceler operates in half-duplex mode (the double-talker is silenced), so that the received symbol primarily consists of echo interference. Our DMT method repeatedly transmits data symbols without cyclic prefixing, forming a periodic training sequence. The frequency-domain received sample is divided by the corresponding frequency-domain data sample to get the frequency-domain echo canceler tap settings: $\hat{W}(k) = Y_i(k)/X_i(k)$. The performance of this fast-start algorithm is similar to that of the periodogram in spectral analysis (see [15] and [16]).

If the kth frequency-domain sample of the echo is $Y_i(k) = H(k)X_i(k) + N_i(k)$, where the noise is zero-mean, then the above method produces an unbiased estimate if M (length of $\boldsymbol{H}$) $< N$ (length of $\boldsymbol{X}$). Otherwise, suppose $\boldsymbol{w}$ denotes the estimated time-domain impulse reponse. Then, the bias is

$$|E[w(k)] - h(k)| \leq \begin{cases} \sum_{m=1}^{M/N} |h(k+mN)|, & k < N \\ |h(k)|, & k > N \end{cases}$$

where $h(k)$ are samples of $\boldsymbol{h}$. Since typical hybrid impulse responses have a exponentially decaying tail with a large time constant, this bias is negligible.

Although the estimate is approximately unbiased, the variance is nonzero. If the noise is zero-mean white Gaussian noise with variance σ^2, the variance of the estimate is $E[W_k - H_k]^2 = \sigma^2/|X_k|^2$. We can reduce the variance by maximizing $|X_k|^2$ subject to a power constraint on the tone. Also, averaging the estimate over m blocks also reduces the variance by a factor of m.

The training algorithms for multirate echo cancelation are then performed as follows. For the interpolated echo canceler, the echo canceler replicates each Hermetian-symmetric frequency-domain data symbol mk times, where m is the noise averaging factor. Table I shows an example of the replicated symbol, where R_i or (I_i) means that the ith sample is assigned a particular real (or imaginary) value. In this table, there are 16 samples in each received symbol and four samples in each transmitted symbol. Each frequency-domain output sample is multiplied by the reciprocal of the input symbol to obtain an estimate of each tap. The use of strictly real (imaginary) input symbols reduces the complexity of multiplication.

For the decimated echo canceler, resampling at the lower rate causes aliasing to occur. As a result, the k sections of the frequency-domain echo canceler taps must be initialized one at a time. This is performed by sending k frequency-domain training symbols, whose nonzero elements are Hermitian-symmetric and correspond to the section of frequency-domain taps to be initialized. Table II gives an example of a training sequence for the decimated canceler. In this table, there are 16 samples in each transmitted symbol and four samples in each received symbol. Symbol 0 is composed of three distinct real values R_i, with the second sample being equal to the 16th, and so forth. Imaginary values are denoted I_i. The frequency-domain hybrid output is then multiplied by the reciprocal of the aliased input symbol (formed with the "block-and-add" operation of the decimated echo canceler), to get the corresponding section of frequency-domain echo canceler taps. Real and imaginary symbols are alternated in order to separately estimate the real and imaginary parts of the taps that overlap as a result of aliasing. In practice, we have access to only half (the "positive" frequencies) of the tables.

VI. PERFORMANCE AND ADSL

This section first proves convergence in the mean for the DMT algorithms, then estimates the rates of convergence and compares them with those of LMS (FBLMS). The DMT techniques are shown to be unbiased estimators that can require

50% more iterations in order to converge. These computations are supported by simulation results. Next, computational and hardware requirements are calculated using the number of multiplies per iteration and the number of bytes required for storage at each iteration. We find that the DMT algorithms can require 93% fewer multiplies, so that in seconds, these techniques converge more quickly.

A. Convergence

We assume, as before, that the echo channel can be modeled by an FIR with additive white Gaussian noise (AWGN), so that the frequency-domain echo can be expressed as $\boldsymbol{Y}_i = \boldsymbol{A}_{i,i-1}\boldsymbol{H} + \boldsymbol{N}_i$, where $\boldsymbol{N}_i$ is a vector of white Gaussian noise. Secondly, the data vectors $\boldsymbol{X}_i$ are assumed to be uncorrelated with each other, which implies that the ith weight vector $\boldsymbol{W}_i$ is independent of the ith data vector $\boldsymbol{X}_i$, using the data update. Thirdly, the frequency-domain elements of $\boldsymbol{X}_i$ are assumed to be uncorrelated with each other. This implies that the time-domain data vector (excluding the guard band) is also independent from one time sample to the next. Finally, we assume equal energy on all the tones, such that if $X_i(n)$ is an element of $\boldsymbol{X}_i$, then $E[X_i^*(n)X_i(n)] = c$.

The LMS (FBLMS) algorithm can be applied to any of the echo cancelers (symmetric, multirate, or frame-asynchronous) with the same basic update $\boldsymbol{W}_{i+1} = \boldsymbol{W}_i + \mu\boldsymbol{A}_{i,i-1}^*\boldsymbol{E}_i$, where $\boldsymbol{E}_i = \boldsymbol{Y}_i - \boldsymbol{A}_{i,i-1}\boldsymbol{W}_i$. The differences arise in the choice of $\boldsymbol{A}_{i,i-1}$ and the representation domain (time or frequency). For example, for the ADSL-RT echo canceler, $\boldsymbol{A}_{i,i-1} = \mathcal{F}_{kN}\boldsymbol{T}_{i,i-1}\mathcal{F}_{kN}^{-1}$, where the Toeplitz matrix $\boldsymbol{T}_{i,i-1}$ comprises the interpolated data. For the ADSL-CO frame-asynchronous echo canceler, $\boldsymbol{A}_{i,i-1} = \tilde{\boldsymbol{T}}_{i,i-1}\mathcal{F}_{kN}^{-1}$, where the Toeplitz matrix $\tilde{\boldsymbol{T}}_{i,i-1}$ is a decimated matrix. In addition, $\boldsymbol{E}_i$ and $\boldsymbol{Y}_i$ are replaced by the (longer) time-domain vectors $\boldsymbol{e}_i$ and $\boldsymbol{y}_i$, respectively.

Let us first look at the ADSL-RT echo canceler. The LMS (FBLMS) adaptive update equation is

$$\boldsymbol{W}_{i+1} = (\boldsymbol{I} - \mu\boldsymbol{A}_{i,i-1}^*\boldsymbol{A}_{i,i-1})\boldsymbol{W}_i + \mu\boldsymbol{A}_{i,i-1}\boldsymbol{Y}_i. \tag{3}$$

Suppose $\boldsymbol{\Delta}_i$ denotes $\boldsymbol{W}_i - \boldsymbol{H}$. Furthermore, let $\boldsymbol{A}_i$ be exactly equal to $\boldsymbol{A}_{i,i-1}$ with the samples of the $(i-1)$th time-domain data vector in $\boldsymbol{T}_{i,i-1}$ zeroed. Then (3) is

$$E[\boldsymbol{\Delta}_{i+1}] = E[\boldsymbol{I} - \mu\boldsymbol{A}_i^*\boldsymbol{A}_i]E[\boldsymbol{\Delta}_i]. \tag{4}$$

Equation (4) is easily derived using the premises and some arithmetic. In addition, we use the independence assumption often used in heuristic arguments for convergence of the LMS algorithm [17] to separate the auto-correlation matrix from the weight tap vector. Suppose that $\tilde{\boldsymbol{\Delta}}_i$ denotes $\boldsymbol{w}_i - \boldsymbol{h}$. Then, (3) can be simplified to be

$$E[\tilde{\boldsymbol{\Delta}}_{i+1}] = (\boldsymbol{I} - \mu c\lambda\boldsymbol{I})E[\tilde{\boldsymbol{\Delta}}_i].$$

The autocovariance matrix of the data is simply the identity matrix multiplied by a constant, $\mu c\lambda$.

We conclude that the modes are decoupled. Since all the eigenvalues are the same and equal to $c\lambda$, each mode has the same convergence rate, with a time constant $\tau = 1/2\mu c\lambda$. For the ADSL-RT canceler, the eigenvalue λ has the value N/N. Performing the same computations for the frame-asynchronous ADSL-CO canceler gives the same final equation, with $\lambda = N + L/kN$. The ADSL-CO echo canceler has a larger time constant and, therefore, converges more slowly.

For the DMT algorithms, the different types of echo cancelers can also be represented with one update $\boldsymbol{W}_{i+1} = \boldsymbol{W}_i + \mu\boldsymbol{B}_i^*\boldsymbol{E}_i$, where $\boldsymbol{E}_i = \boldsymbol{Y}_i - \boldsymbol{A}_{i,i-1}\boldsymbol{W}_i$. For the ADSL-RT echo canceler, $\boldsymbol{A}_{i,i-1}$ is the same as in the ADSL-RT case for LMS (FBLMS), while $\boldsymbol{B}_i = \mathrm{diag}(\boldsymbol{R}\boldsymbol{X}_i^*)$. For the frame-asynchronous ADSL-CO echo canceler, the notation is again the same as in the ADSL-CO case for LMS (FBLMS) with the addition of $\boldsymbol{B}_i = \mathrm{diag}(\boldsymbol{X}_i^*)\boldsymbol{R}\mathcal{F}_N\boldsymbol{P}'$.

These examples show that the adaptive update employs a one-tap complex LMS filter on each echo tap W_i. In effect, the echo subchannels are assumed to be independent, so the taps can be decoupled. As previously mentioned, this only happens if $M < L$. In general, the subchannels are correlated, and performance suffers for having assumed independence. We now investigate this effect.

The update is

$$\boldsymbol{W}_{i+1} = (\boldsymbol{I} - \mu\boldsymbol{B}_i^*\boldsymbol{A}_{i,i-1})\boldsymbol{W}_i + \mu\boldsymbol{A}_{i,i-1}\boldsymbol{Y}_i. \tag{5}$$

Suppose that $\boldsymbol{\Lambda}$ denotes a diagonal matrix. Then, for the ADSL-RT canceler, (5) is

$$E[\boldsymbol{\Delta}_{i+1}] = E[\boldsymbol{I} - \mu\boldsymbol{B}_i^*\boldsymbol{A}_i]E[\boldsymbol{\Delta}_i] = (\boldsymbol{I} - \mu c\mathcal{F}_{kN}\boldsymbol{\Lambda}\mathcal{F}_{kN}^{-1})E[\boldsymbol{\Delta}_i]$$

or

$$E[\tilde{\boldsymbol{\Delta}}_{i+1}] = (\boldsymbol{I} - \mu c\boldsymbol{\Lambda})E[\tilde{\boldsymbol{\Delta}}_i].$$

In this case, the independence assumption is not required. The equation for the ADSL-CO canceler has a similar form.

For the RT echo canceler, the diagonal of $\boldsymbol{\Lambda}$ is the following vector composed of M nonzero elements

$$\left[\underbrace{1,\cdots 1}_{(L+1)k},\underbrace{\frac{N-1}{N},\cdots,\frac{N-1}{N}}_{k},\underbrace{\frac{N-2}{N},\cdots,\frac{N-2}{N}}_{k},\cdots\right].$$

For the CO echo canceler, the diagonal of $\boldsymbol{\Lambda}$ is the following vector of M nonzero elements:

$$\left[\frac{N+L}{kN},\underbrace{2\frac{N+L-1}{kN},\cdots\frac{N+L-1}{kN}}_{k},\underbrace{2\frac{N+L-2}{kN},\cdots\frac{N+L-2}{kN}}_{k},\cdots\right].$$

Equation (6) shows that the modes can still be decoupled, but there exists an eigenvalue spread, and we have a sum of first-order decaying exponentials of different time constants. For derivation details of all results in this section, we refer the reader to [18].

The results above are verified by simulations. In these simulations, each tone transmits a 4-QAM signal constellation of $\pm 1 \pm j$. The transmit blocklength N is chosen to be 512,

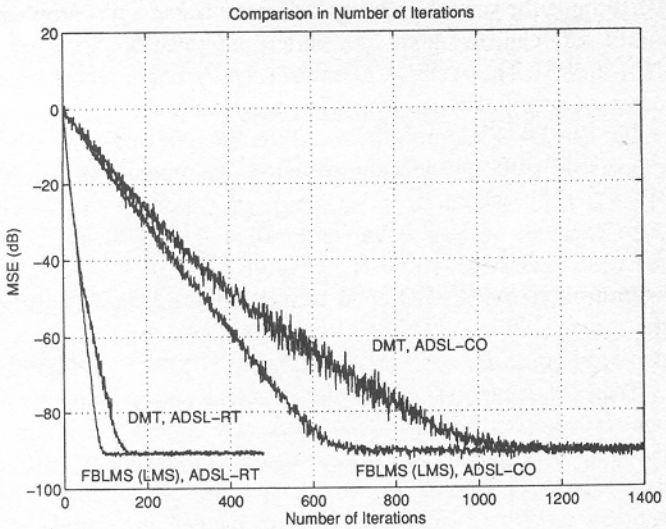


Fig. 4. Comparison of convergence in iterations.

M is 221, and the receive blocklength is 64, so that $k = 8$. The step-size $\mu = 2^{-4}$ and the weights are initialized to zero.

Fig. 4 shows the four example convergence curves for each of the four algorithms. The time constant τ for the ADSL-CO canceler using FBLMS is approximately 31, very close to the theoretical value of 30. For the ADSL-RT canceler using FBLMS, we compute τ to be approximately equal to four, and this matches the simulation. For the DMT algorithms, the initial weight vector partially cancels the effect of the eigenvalue spread. Most ADSL channels are low-pass, so the weights near Nyquist associated with the slow eigenvalues will already be fairly close to the optimum values.

Both computation and simulation results show that the DMT algorithms converge more slowly (in number of iterations) than FBLMS, at a given step size. These examples show that the DMT algorithms can require 50% more data samples than FBLMS.

It is possible to initially increase the convergence rates of the DMT algorithms by the use of training sequences. In steady-state operation, we can increase the step-size. In this way, we equate the time constant of the slowest mode of the DMT algorithms with that of FBLMS. However, a trade-off exists, since an increase in step-size leads to an increase in the amount of misdjustment noise inherent in a stochastic gradient-approximation technique. It is also possible to use a variable step-size per tone in order to precondition the input data. This can also compensate for "slow" tones with little or no energy assigned to them.

Although the DMT algorithms converge at a slower rate for a given step-size, these comparisons are misleading, since they do not account for the computational cost of the algorithm at each iteration. We quantify this in the next subsection.

B. Computational Complexity and Memory Requirements

The computational complexity of the algorithms is approximated by the total number of multiplies per iteration. Table III lists the total complexity for the interpolated (ADSL-RT) and decimated (ADSL-CO) echo cancelers. In this table, M is the number of time-domain taps, N is the smaller blocklength, k is the upsampling factor, the infrequently-performed IFFT that transforms $\boldsymbol{W}_i$ to $\boldsymbol{w}_i$ is averaged over m samples, each complex multiply equals four real multiplies, and the time-domain convolution complexity is approximated. The ADSL-CO echo canceler is assumed to be frame-asynchronous, so an extra N-point FFT/IFFT pair is added.

TABLE III
COMPLEXITY COMPARISON, MULTIPLICATIONS/SYMBOL

Application	*Echo Emulation*	*Adaptive Update*
ADSL (CO)	$2kN + \frac{(M-1)^2}{2k} + 2N \log N$	$2kN + \frac{2kN \log kN}{m} + 2N \log N$
ADSL (RT)	$2kN + \frac{(M-L-1)^2}{2k}$	$2kN + \frac{2kN \log kN}{m}$

TABLE IV
COMPLEXITY COMPARISON, MULTIPLICATIONS/SYMBOL

	FBLMS	LMS	DMT
Adaptive Update	$2(4N \log_2 2N) + 2N$	MN	$2N + \frac{2N \log N}{m}$
Echo Emulation	$4N \log_2 2N + 2N$	MN	$2N + \frac{(M-L-1)^2}{2}$
MIPS (804 kHz)	100	206	17

The complexity of the time-domain convolution can further be reduced by up to a factor of two by controlling the alignment between transmit and receive data symbol streams [19]. Table IV compares the complexity of the LMS, FBLMS, and DMT techniques using an example symmetric-rate scenario. In this table, $L = 32, N = 512, M = 128, m = 24$.

Complexity is dependent on the system parameters. For instance, a longer guard band, L, will decrease the complexity of the time-domain convolution for the algorithms in comparison with time-domain LMS. However, increasing L will decrease the available bandwidth of the system, so a trade-off exists. With this caveat in mind, the example in Table IV shows that the use of more traditional methods for DMT echo cancelation is impractical.

These tables also show that the lower complexity of the DMT adaptive update more than compensates for the slower rate of convergence. Fig. 5 plots convergence against total sequential multiplies and shows a two to three-fold improvement over LMS. Although not shown, the gains over FBLMS are even more dramatic because of the FFT size. For applications that require high data rates with minimum hardware and for which the channel is relatively stationary, such a trade-off between per-iteration convergence and per-iteration complexity is a highly acceptable one.

The DMT algorithms require slightly more memory than LMS and FBLMS, since we store two sets of echo taps: $\boldsymbol{W}_i$ and $\boldsymbol{w}_i$. There are $N/2$ complex frequency-domain taps and $M < N$ real time-domain taps. In addition, we store a portion of the previous data symbol, $\boldsymbol{x}_{i-1}$, although this would also be required by FBLMS. The stored value of $\boldsymbol{x}_{i-1}$ can be used to reduce processing latency, so that while $\boldsymbol{x}_{i-1}$ is being transmitted, $\boldsymbol{x}_i$ is being modulated.

Another multicarrier echo canceler [20] was developed based on Hirosaki's multicarrier system [21], where a bank of low-rate bandpass FIR filters separate the subchannels in

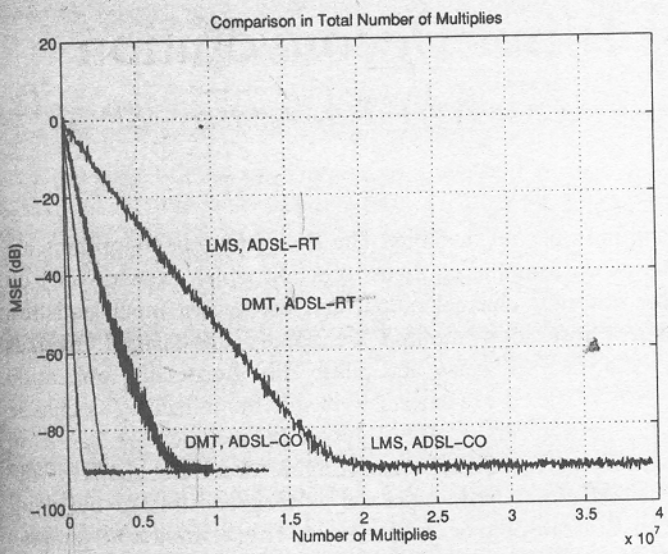


Fig. 5. Comparison of convergence in multiplies.

the frequency domain. Although this echo canceler structure is simpler than the one presented here, the complexity of modulation and demodulation more than offsets the gain in echo cancelation.

VII. Conclusion

The problem of echo cancelation for systems using DMT had been widely believed to be impossible to solve because of the high computational complexity required by traditional methods. However, we have proposed feasible echo-cancelation techniques that are of extremely low complexity when compared to alternative methods. These methods take advantage of the structure of the DMT signal to reduce computational complexity, by using a fast convolution technique that exploits both time and frequency domains for estimating the echo and a simple frequency-domain adaptive update. These techniques are easily generalized for multirate DMT echo cancelation and have been implemented for real-time ADSL applications, allowing for data rates approaching the capacity of the system.

Acknowledgment

The authors thank J. Chow of Amati Communications, M. Morf of Stanford, R. Younce of Tellabs, and the anonymous reviewers for their helpful comments.

References

[1] J. L. Holsinger, "Digital communication over fixed time-continuous channels with memory, with special applications to telephone channels," *MIT Res. Lab. Electron. Rep.*, vol. 430, 1964.

[2] ANSI "Network and customer installation interfaces: Asymmetric digital subscriber line (ADSL) metallic interface," in *Draft American National Standard for Telecommunications*, vol. T1E1.4/94-007R8, 1995.

[3] G. Plenge, "DAB—A new sound broadcasting system," *EBU Rev.*, vol. 246, pp. 1–26, Apr. 1991.

[4] J. Cioffi and J. A. C. Bingham, "A data-driven multitone echo canceller," *IEEE Trans. Commun.*, pp. 2853–2869, Oct 1994.

[5] J. M. Cioffi and J. A. C. Bingham, "A data-driven multitone echo canceller," in *GLOBECOM '91*, Phoenix, Az, Jan. 1991, pp. 57–61.

[6] A. Ruiz, J. M. Cioffi, and S. Kasturia, "Discrete multiple tone modulation with coset coding for the spectrally shaped channel," *IEEE Trans. Commun.*, vol. 40, pp. 1012–1029, June 1992.

[7] J. M. Cioffi, "A multicarrier primer," in *ANSI T1E1.4 Contributions*, Boca Raton, FL, pp. 91–157, Nov. 1991.

[8] P. S. Chow, "Bandwidth optimized digital transmission techniques for spectrally shaped channels with impulse noise," Ph.D. dissertation, Stanford University, Stanford, CA, 1993.

[9] T. N. Zogakis, "Coding and shaping for discrete multitone modulation over band-limited channels," Ph.D. dissertation, Stanford University, Stanford, CA, 1994.

[10] W. Chen, J. L. Dixon, and D. L. Waring, "High bit rate digital subscriber line echo cancelation," *IEEE J. Select. Areas Commun.*, vol. 9, pp. 848–860, Aug. 1991.

[11] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.

[12] B. Widrow, J. McCool, and M. Ball, "The complex LMS algorithm," *Proc. IEEE*, vol. 63, pp. 719–720, Apr. 1975.

[13] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[14] R. C. Younce, P. J. W. Melsa, and S. Kapoor, "Echo cancelation for asymmetrical digital subscriber lines," in *1994 Int. Conf. Commun.*, New Orleans, LA, May 1994.

[15] L. Ljung, *System Identification, Theory for the User*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[16] J. M. Cioffi, "A fast echo canceller initialization method for the CCITT V32 modem," in *GLOBECOM'87*, Tokyo, Japan, Nov. 1987.

[17] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1986.

[18] M. Ho, "Multicarrier echo cancelation and multichannel equalization," Ph.D. dissertation, Stanford University, Stanford, CA, 1995.

[19] R. Younce, "ANSI ADSL Standard contribution T1E1.4/93-267." Distributed in Oct., 1993.

[20] J. Yang, S. Roy, and N. H. Lewis, "Data driven echo cancelation for a multi-tone modulation system," *IEEE Trans. Commun.*, vol. 42, pp. 2133–2144, May 1994.

[21] S. Hirosaki, "An orthogonally multiplexed QAM system using the discrete fourier transform," *IEEE Trans. Commun.*, pp. 982–989, 1981.



Minnie Ho (S'87–M'93) received the B.S.E.E. degree in electrical engineering and computer science from Princeton University, Princeton, NJ, in 1988, and the M.S.E.E. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1990 and 1995, respectively.

She has been employed by Radix Technologies in Mountain View, CA, since 1994. Her interests are in data transmission, communication theory, and coding.

Dr. Ho was the recipient of the National Science Foundation Graduate Fellowship, from 1988–1991.

John M. Cioffi (S'77–M'78–SM'90), for a photograph and biography, see p. 64 of the January 1996 issue of this TRANSACTIONS.

John A. C. Bingham (M'59–SM'68–F'93), photograph and biography not available at the time of publication.