NEXT Cancellers Using FDLMS Filters with Improved Convergence Rate

R. C. Nongpiur, D. J. Shpak, and A. Antoniou
Dept. of Electrical and Computer Engineering
University of Victoria
Victoria, BC, Canada V8W 3P6
{rnongpiu, dale, aantoniou}@ece.uvic.ca

Abstract

A low-complexity technique for improving the convergence rate in a near-end crosstalk (NEXT) cancellation system is proposed. The cancellation system minimizes the computational load by first detecting the lines causing the major NEXT and then assigning FDLMS adaptive filters to cancel the NEXT. By initially making the step-size of each adaptive filter proportional to the magnitude of the NEXT it is supposed to cancel, the convergence rate is significantly improved at almost no extra computational cost, especially during the initial stage of adaptation. A high convergence rate during the initial stages is more critical as it reduces the latency time in xDSL systems, thereby resulting in a higher data throughput.

1 Introduction

As xDSL technology is improved, the data-rate that can be achieved through twisted pairs is rising. One major impairment that hinders the simultaneous transmission of signals in a bundle is near-end crosstalk (NEXT). Severe crosstalk not only limits the maximum data rate of any individual line but it can also degrade an existing service if a new service is added to the bundle.

The common technique for the reduction of NEXT is to use adaptive filters to attenuate the NEXT signals coming from the adjacent twisted pairs as much as possible. However, this technique can be very expensive if the bundle contains a large number of twisted pairs. A better option is to first detect the crosstalk present on a line and then add an adaptive filter to cancel each of the crosstalk sources detected [1].

As the sampling rate for the filter increases, the filter length required for NEXT cancellation also increases. By using the frequency-domain LMS algorithm [2][3] instead of the time-domain LMS algorithm, the amount of computation can be greatly reduced especially when the length of the adaptive canceller is large. We can use a fast method to cross-correlate [4] the received signal with the reference signal to roughly estimate the impulse response of the NEXT coupling function [1]. This can then be used to determine the magnitude of the step size for the adaptation algorithm in the initial stage of adaptation. It will be seen that this technique results in a significant increase in the convergence rate especially during the initial stage of adaptation. Having a faster initial convergence rate is more important as it allows the equalizer in the xDSL system to adapt at a much faster rate and thus lower the latency time for the system which results in a higher data throughput. The later stages of adaptation when the error has been reduced by more than 3 dB are not as critical since the data throughput in these stages is almost close to the optimum.

The paper is organized as follows. Section 2 describes the NEXT cancellation system which first detects the major NEXT sources before assigning FDLMS adaptive filters to cancel them. In Section 3 the convergence rate of this NEXT cancellation system is studied and a way to improve the convergence rate is proposed. Simulation results are presented in section 4.

2 NEXT Cancellation

The echo-cancelled signal received on a single twisted pair in a bundle of an xDSL system can be expressed as

$$\hat{y}(n) = r(n) + \eta_b(n) + \sum_i h_i(n) * d_i(n) \qquad (1)$$

where $\hat{y}(n)$ is the noisy observation, $r(n)$ is the received message signal, $\eta_b(n)$ is additive background noise, $d_i(n)$ is the signal transmitted on pair i, and $h_i(n)$ is the time-domain crosstalk coupling function between pair i and the considered pair.

For the purpose of crosstalk detection, we assume that the echo at the hybrid is already cancelled by

0-7803-8104-1/03/$17.00 ©2003 IEEE

the echo canceller before the crosstalk detection is performed. In most cases, the impulse response of the channel can be readily estimated [5]. Hence the received signal can be removed, leaving only the crosstalk signals. With the received signal and the transmitted signal echo removed, the signal $y(n)$ can be expressed as

$$y(n) = \eta_b(n) + \sum_i h_i(n) * d_i(n) \qquad (2)$$

Since the magnitude of the NEXT coming from the other twisted-pairs is variable, the block length of each adaptive filter can be adjusted according to the magnitude of the NEXT signal. In cases where the magnitude of the NEXT from a particular line is smaller than a prescribed threshold, the NEXT canceller for that line can be removed. In order to estimate the magnitude of each NEXT signal, the NEXT cancellers are coupled with the NEXT detectors [1] as shown in Fig. 1. The NEXT detectors use a fast cross-correlation of the input and received signal to estimate the magnitude of the NEXT coupling function.

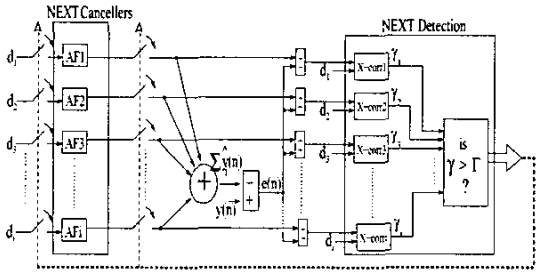

Figure 1: NEXT adaptive canceller.

Taking the ideal case where each transmitted signal d_i is uncorrelated and has an energy ε, i.e., $R_{d_i}(k) = \varepsilon\delta(k)$, we have

$$R_{yd_i}(l) = \varepsilon h_i(l) \qquad (3)$$

This shows that if the signals d_i are uncorrelated, the crosscorrelation is proportional to the impulse response of the crosstalk coupling.

The adaptation is started by using the NEXT detectors to estimate the magnitude of the NEXT in each line. To decide whether the crosstalk signal requires cancellation by an adaptive filter, a threshold value Γ can be set [1]. If

$$\gamma_i = \frac{\max|R_{yd_i}(l)|}{\sum_{l=0}^{L}|R_{yd_i}(l)|} > \Gamma \qquad (4)$$

then the NEXT signal is considered large enough to require cancellation.

3 Improved Convergence Rate

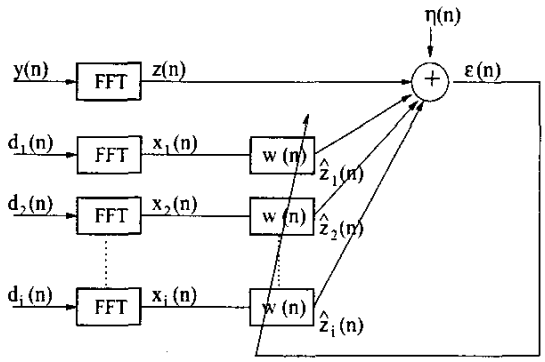

Figure 2: Adaptive filtering in the frequency domain for each frequency bin.

The model of the frequency domain adaptive filter algorithm for cancelling the NEXT is shown in Fig. 2. The model is an extension to the model used by Bershad et al. [6], so as to incorporate multiple adaptive filters. The input signal to each FFT block is assumed to be wide-sense stationary making the disjoint spectral output uncorrelated. The input is also assumed to be a joint Gaussian random process, so that the disjoint bins provide statistically independent outputs. Thus each of the spectral taps will be operating on independent data [7].

The input signals to the various FFT blocks are assumed to be statistically independent from one another. The weight update equation for adaptive filter i in each frequency bin is given by

$$w_i(n+1) = w_i(n) + \mu_i\epsilon(n)x_i^*(n) \qquad (5)$$

where $w_i(n)$ is the complex scalar weight on the nth iteration of adaptive filter i, $x_i(n)$ is the ith frequency-domain input data sequence, μ_i is the ith feedback coefficient, and

$\epsilon(n)$ = error $= z(n) - \sum_i \hat{z}_i(n) + \eta(n)$
$= \sum_i z_i(n) - \sum_i \hat{z}_i(n) + \eta(n)$
where $z_i(n)$ is the ith reference waveform, $\hat{z}_i(n)$ is the output from filter $i = w_i(n)x_i(n)$

It is assumed that a statistically independent Gaussian additive noise $\eta(n)$ is added to the error $\epsilon(n)$ and that α_i is the complex correlation coefficient of the desired signal $z_i(n)$ when $E[|e(n)|^2]$ is minimized. The relevant statistics of $z_i(n)$ and $x_i(n)$ are

$$\begin{aligned} E[x_i(n)] &= E[z_i(n)] = 0 \\ E[|x_i(n)|^2] &= \sigma_x^2 \\ E[|z_i(n)|^2] &= |\alpha_i|^2\sigma_x^2 \end{aligned}$$

$$\begin{aligned} E[z_i(n)x_i^*(n)] &= \alpha_i\sigma_\eta^2 \\ E[|\eta(n)|^2] &= \sigma_\eta^2 \end{aligned} \tag{6}$$

Subtracting $w_{oi} = \alpha_i$ from both sides of (5) so that $w_i(n) - w_{oi} = \nu_i(n)$ yields

$$\begin{aligned} \nu_i(n+1) &= \nu_i(n) + \mu_i \Bigg[\sum_j z_j(n) + \eta(n) \\ &\quad - \sum_j w_j(n)x_j(n) \Bigg] x_i^*(n) \end{aligned} \tag{7}$$

Squaring the magnitude in (7), taking the expectation and then evaluating each of the terms, yields a first-order linear difference equation for $E[|\nu(n)|^2]$,

$$\begin{aligned} E[|\nu_i(n+1)|^2] &= (1 - 2\mu_i\sigma_x^2 + 2\sigma_x^4\mu_i^2)E[|\nu_i(n)|^2] \\ &\quad + \mu_i^2\sigma_x^4 \sum_{j,j\neq i} E[|\nu_j(n)|^2] \\ &\quad + \mu_i^2\sigma_x^2\sigma_\eta^2 \end{aligned} \tag{8}$$

Letting $\mu_{oi} = \mu_i\sigma_x^2$, (8) can be written as

$$\begin{aligned} E[|\nu_i(n+1)|^2] &= (1 - 2\mu_{oi} + 2\mu_{oi}^2)E[|\nu_i(n)|^2] \\ &\quad + \mu_{oi}^2 \sum_{j,j\neq i} E[|\nu_j(n)|^2] + \mu_{oi}^2(\frac{\sigma_\eta^2}{\sigma_x^2}) \end{aligned} \tag{9}$$

Now, (9) can be compactly written in matrix form

$$\mathbf{v}(n+1) = \mathbf{A}\mathbf{v}(n) + \mathbf{b} \tag{10}$$

where

$$\mathbf{v}(n) = \begin{bmatrix} E[|\nu_1(n)|^2] \\ E[|\nu_2(n)|^2] \\ \vdots \\ E[|\nu_N(n)|^2] \end{bmatrix} \tag{11}$$

$$\mathbf{A} = \begin{bmatrix} (1 - 2\mu_{o1} + 2\mu_{o1}^2) & \mu_{o1}^2 & \cdots & \mu_{o1}^2 \\ \mu_{o2}^2 & (1 - 2\mu_{o2} + 2\mu_{o2}^2) & \cdots & \mu_{o2}^2 \\ \vdots & \vdots & \cdots\cdots & \vdots \\ \mu_{oN}^2 & \mu_{oN}^2 & \cdots & (1 - 2\mu_{oN} + 2\mu_{oN}^2) \end{bmatrix} \tag{12}$$

and

$$\mathbf{b} = \frac{\sigma_\eta^2}{\sigma_x^2} \begin{bmatrix} \mu_{o1}^2 \\ \mu_{o2}^2 \\ \vdots \\ \mu_{oN}^2 \end{bmatrix} \tag{13}$$

The solution of the difference equation in (10) is given by

$$\mathbf{v}(n) = \mathbf{A}^n\mathbf{v}(0) + (\mathbf{A}^{n-1} + \mathbf{A}^{n-2} + \cdots + I)\mathbf{b} \tag{14}$$

or more compactly as

$$\mathbf{v}(n) = \mathbf{A}^n\mathbf{v}(0) + (I - \mathbf{A})^{-1}(I - \mathbf{A}^n)\mathbf{b} \tag{15}$$

assuming that the system is stable.

3.1 Convergence rate

The mean-square error for the system is given as

$$E[|\epsilon(n)|^2] = \sigma_x^2 \sum_i E[|\nu_i(n)|^2] + \sigma_\eta^2 \tag{16}$$

In real-world echo cancellers, the typical value of the adaptation step-size μ is significantly smaller than the speed-optimal gain or the stability-limit gain [8]. The same is very much true for NEXT cancellers especially when the number of adaptive filters in the system is large. The value μ_{oi} is generally much smaller than the stability-limit gain which is always less than 1 and it would be safe to assume that μ_{oi} is significantly larger than μ_{oi}^2. Thus with this assumption matrix $\mathbf{A}$ can be simplified as

$$\mathbf{A} \approx \hat{\mathbf{A}} = \begin{bmatrix} (1 - 2\mu_{o1}) & 0 & \cdots & 0 \\ 0 & (1 - 2\mu_{o2}) & \cdots & 0 \\ \vdots & & \cdots & \vdots \\ 0 & \cdots & 0 & (1 - 2\mu_{oN}) \end{bmatrix} \tag{17}$$

Thus, (15) can be simplified as

$$\mathbf{v}(n) = \hat{\mathbf{A}}^n\mathbf{v}(0) + (I - \hat{\mathbf{A}})^{-1}(I - \hat{\mathbf{A}}^n)\mathbf{b} \tag{18}$$

where

$$\hat{\mathbf{A}}^n = \begin{bmatrix} (1 - 2\mu_{o1})^n & 0 & \cdots & 0 \\ 0 & (1 - 2\mu_{o2})^n & \cdots & 0 \\ \vdots & & \cdots & \vdots \\ 0 & \cdots & 0 & (1 - 2\mu_{oN})^n \end{bmatrix} \tag{19}$$

and

$$(I - \hat{\mathbf{A}})^{-1}(I - \hat{\mathbf{A}}^n) = \begin{bmatrix} \frac{1-(1-2\mu_{o1})^n}{2\mu_{o1}} & 0 & \cdots & 0 \\ 0 & \frac{1-(1-2\mu_{o2})^n}{2\mu_{o2}} & \cdots & 0 \\ \vdots & & \cdots & \vdots \\ 0 & \cdots & 0 & \frac{1-(1-2\mu_{oN})^n}{2\mu_{oN}} \end{bmatrix} \tag{20}$$

In NEXT cancellation, the NEXT coming from the various lines is variable and assuming the initial weight $w_i(0)$ of each adaptive canceller is zero, we have

$$\mathbf{v}(0) = \begin{bmatrix} E[|\nu_1(0)|^2] \\ E[|\nu_2(0)|^2] \\ \vdots \\ E[|\nu_N(0)|^2] \end{bmatrix} = \begin{bmatrix} |\alpha_1|^2 \\ |\alpha_2|^2 \\ \vdots \\ |\alpha_N|^2 \end{bmatrix} \tag{21}$$

Using (18) and (21) the mean-square error in (16) can be written as

$$E[|\epsilon(n)|^2] = \sigma_x^2 \sum_i \left\{ 2(1-2\mu_{oi})^n |\alpha_i|^2 + \mu_{oi}[1-(1-2\mu_{oi})^n] \frac{\sigma_\eta^2}{\sigma_x^2} \right\} + \sigma_\eta^2 \quad (22)$$

The mean-residual error (MRE) can be estimated by setting $n = \infty$ in (22) so that

$$\text{MRE} \approx E[|\epsilon(\infty)|^2] = \sigma_\eta^2 \sum_i \mu_{oi} + \sigma_\eta^2 \quad (23)$$

Simplifying (22) and neglecting the squared and higher terms of μ_{oi}, we get an approximate value of $E[|\epsilon(n)|^2]$ which is accurate for smaller integer values of n, and for μ_{oi} much smaller than 1, given as

$$E[|\epsilon(n)|^2] \approx \sigma_x^2 \sum_i |\alpha_i|^2 - 2\sigma_x^2 \sum_i n|\alpha_i|^2 \mu_{oi} + \sigma_\eta^2 \quad (24)$$

From (23) and (24) we see that for the same mean-residual error or $\sum_i \mu_{oi}$ = constant, $E[|\epsilon(n)|^2]$ will have a greater initial reduction per iteration if

$$\mu_{oi} = k|\alpha_i|^2 \quad (25)$$

where k is a constant, rather than making all of the μ_{oi} equal.

4 Simulation Results

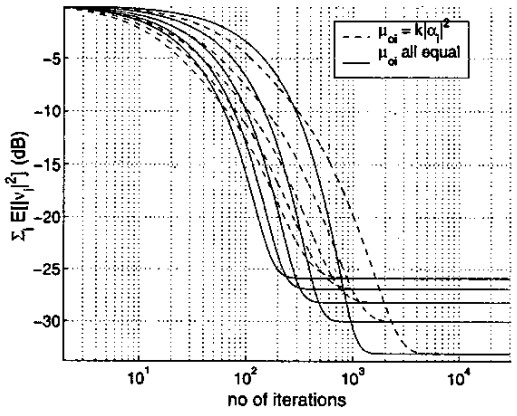

Figure 3: Plot of $\sum_i E[|\nu_i(n)|^2]$ for μ_{oi} all equal and μ_{oi} proportional to $|\alpha_i|^2$.

By evaluating $\sum_i E[|\nu_i(n)|^2]$, which is the variable term in (16), we can get a clearer picture of the convergence of $E[|\epsilon(n)|^2]$. Using (14), $\sum_i E[|\nu_i(n)|^2]$ is evaluated for two cases - the first when the μ_{oi} are all equal and the second when the μ_{oi} are made proportional to $|\alpha_i|^2$ as in (25). Fig. 3 shows the plots generated for the two cases with different residual values. The different residual values are generated by varying the proportion of μ_{oi}. In evaluating the plots, σ_x^2 is set to one, and σ_η^2 is set to 0.1. The initial weight error vector $\mathbf{v}(0)$ is a 6-element vector with variable elements given by

$$\mathbf{v}(0) = [0.9 \;\; 0.8 \;\; 0.2 \;\; 0.1 \;\; 0.3 \;\; 0.2]^T \quad (26)$$

From the plot it can be seen that varying μ_{oi} as in (25), only helps in the initial portion of the convergence until the error has been reduced by around 5 dB. Below 5 dB the advantage of varying μ_{oi} in proportion to $|\alpha_i|^2$ is reduced. Hence, an effective strategy for increasing the convergence rate would be to use (25) to speed up the initial convergence, and once the error has been reduced by around 5 dB, the μ_{oi} can then be made all equal until the system converges.

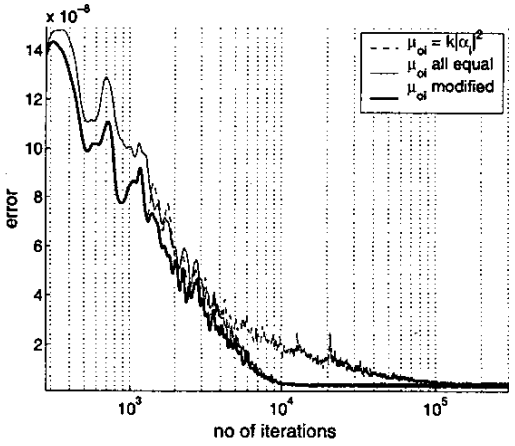

Figure 4: Plot of $\sum_i E[|\nu_i(n)|^2]$ for μ_{oi} all equal and μ_{oi} proportional to $|\alpha_i|^2$.

In adaptive filtering applications, $|\alpha_i|$ is generally not known until the adaptive filter has fully converged. Also, in an FDLMS adaptive system $|\alpha_i|$ will vary for each frequency bin depending upon the spectrum of the NEXT transfer function. However, a good estimate of the average value of $|\alpha_i|$ over the entire NEXT frequency range is the maximum value of $|R_{yd_i}(l)|^2$ defined in (3), that is

$$|\alpha_i|^2_{average} \approx c \left[\max_l \{|R_{yd_i}(l)|^2\} \right] \quad (27)$$

where $|\alpha_i|_{average}$ is the average value of $|\alpha_i|$ in each frequency bin, and c is some constant.

Since $R_{yd_i}(l)$ has already been calculated when estimating γ_i in (4), it can again be reused without incuring additional computation. Thus, for all the frequency bins in the ith adaptive filter, we have

$$\mu_{oi} = \kappa \left[\max_{l} \{|R_{yd_i}(l)|^2\} \right] \tag{28}$$

where κ is the appropriate constant chosen so that the mean-residual error in (23) is below a prescribed threshold.

Looking at the plot in Fig. 4, a comparison is made between the FDLMS cancellers where all μ_{oi} are made equal, and the FDLMS cancellers where μ_{oi} is varied according to (28). The plot labelled as 'μ_{oi} modified' corresponds to the case where μ_{oi} is varied according to (28) during the initial portion of the convergence and once the error has reduced by 5 dB, all μ_{oi} are made equal. The simulations were performed in an HDSL2 [9] environment using FDLMS adaptive filters to cancel NEXT coming from 6 different twisted pairs in a bundle. The relative magnitude of the 6 NEXTs used in the simulations are [0.0064 0.28 0.07 1 0.02 0.58] with the largest NEXT normalized to 1. It can be seen that the third plot shows an initial convergence which is the same as the first plot with '$\mu_{oi} = k|\alpha_i|^2$' but is much faster than for the second plot with 'μ_{oi} all equal'. As the error reduces below 5dB the convergence rate in the first plot reduces and becomes the lowest of the three plots. For the third plot, the μ_{oi} are made all equal after 5dB reduction in error. This causes the third plot to exhibit the fastest convergence rate until the system converges. In the figure, all the three plots eventually converge to the same mean-residual error.

5 Conclusions

A new technique for improving the convergence rate in a NEXT cancellation system was proposed. Using FDLMS adaptive filters, gives a lower computational complexity when the sampling rate is high and the filter length is long. Using the proposed modification to the FDLMS algorithm, the convergence rate is significantly improved especially during the initial stages where it is so critical, and at almost no extra computational cost.

Acknowledgements

The authors are grateful to PMC-Sierra, Micronet, NCE Program, and the Natural Sciences and Engineering Research Council of Canada for supporting this work.

References

[1] R. C. Nongpiur, D. J. Shpak, and A. Antoniou, "NEXT cancellation in xDSL systems," *36th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA., Nov. 3-6, 2002.*

[2] J. J. Shynk, "Frequency-domain and multirate adaptive filtering," *IEEE Signal Processing Magazine*, vol. 9, no. 1, pp. 14-37, 1997.

[3] S. Haykin, *Adaptive Filter Theory*, Third Edition, Prentice Hall 1996.

[4] T. J. Shan and T. Kailath, "Adaptive algorithms with an automatic gain control feature," *IEEE Transactions on Circuits and Systems*, vol. 35, no. 1, pp. 122-127, Jan. 1988.

[5] C. Zeng, C. Aldana, A. A. Salvekar, and J. M. Cioffi, "Crosstalk identification in xDSL systems," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 8, pp. 1488-1496, Aug. 2001.

[6] N. J. Bershad and P. L. Feintuch, "A normalized frequency domain LMS adaptive algorithm," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-34, no. 3, Jun. 1986.

[7] N. J. Bershad and P. L. Feintuch, "Analysis of the frequency domain adaptive filter," *Proceedings of the IEEE*, vol. 67, no. 12, Dec. 1979.

[8] D. L. Duttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancellers," *IEEE Transactions on Speech and Audio Signal Processing*, vol. 8, no. 5, Sep. 2000.

[9] "Draft for HDSL2 Standard," *T1E1.4/99-006R5.*