A MINIMUM MEAN-SQUARED ERROR INTERPRETATION OF RESIDUAL ISI CHANNEL SHORTENING FOR DISCRETE MULTITONE TRANSCEIVERS

D. Daly, C. Heneghan & A.D. Fagan

DSP Research Group
Department of Electronic & Electrical Engineering
University College Dublin, Belfield, Dublin 4, Ireland

donnacha@ee.ucd.ie conor.heneghan@ucd.ie tony.fagan@ucd.ie

ABSTRACT

Melsa *et al.* [1] presented a channel shortening technique for Discrete Multitone transceivers that reduces Intersymbol Interference (ISI) by forcing the effective channel's impulse response to lie within a window of v+1 consecutive samples. Arslan *et al.* [2] claim that although this method is intuitive, no previous study has been made on its optimality. They comment on its optimality by simulation. In this paper it is demonstrated that Melsa's approach is in fact theoretically equivalent to a minimum mean-squared error (MMSE) solution to the channel-shortening problem. As a corollary to this we are afforded an insight into MMSE channel shortening as originally proposed by Falconer and Magee [3]. Previously, it has not been intuitive as to why the Desired Impulse Response (DIR) should be made adaptive in this approach. Our result demonstrates that allowing DIR adaptation achieves a minimisation of the effective impulse response energy outside the desired window of v samples.

1. INTRODUCTION

ISI–free transmission can be achieved in a Discrete Multitone (DMT) system, by prepending a cyclic prefix of size v onto each block of N transmitted samples. This holds provided $v > M$ where M is the index of the discrete-time equivalent channel impulse response $h(k)$, beyond which the response can be considered insignificant. In order to minimise the bit-rate reduction factor $^{v}/_{N+v}$ caused by cyclic prefix insertion, it is desired that the effective channel impulse response length M be less than some nominal value. Viterbi decoding in the receiver, will also have its complexity reduced by transmission over a shortened channel [3]. To this end, an equaliser is commonly used at the receiver. The channel-shortening filter for DMT is generally referred to as a Time Domain Equaliser (TEQ).

In Section 2, we recap on the Residual ISI Minimisation (RISIM) TEQ presented by Mesla *et al.* In Section 3 a Minimum Mean-Squared Error (MMSE) TEQ is presented, which, in Section 4 is demonstrated to be equivalent to RISIM. The effect of system noise on this equivalence is discussed in Section 5. Simulation results are presented in Section 6, in which we investigate channel shortening in a multitone, Very high speed Digital Subscriber Line (VDSL) environment.

2. RISIM CHANNEL SHORTENING

Consider DMT transmission over a discrete channel $h(k)$, modelled by an FIR filter of order M, followed by a TEQ with p taps $w(k)$ as illustrated in Figure 1.

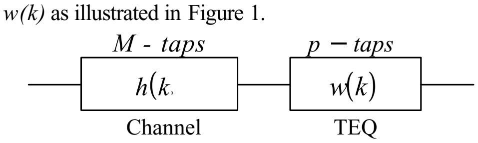

Figure 1 *Channel Shortening TEQ*

The effective signal path is given by the convolution of eqn. (1).

$$h_{\text{eff}}(k) = h(k) * w(k) \tag{1}$$

In order to achieve suitable channel shortening, it is desirable that most of the energy of $h_{\text{eff}}(k)$ will fall within a window of v taps. Referring to Figure 2, a measure of residual ISI ρ caused by imperfect choice of $w(k)$ is the ratio of energy outside the window of v samples to the energy within this window. In Melsa's intuitive approach to channel shortening [1], eigen-analysis is used to minimise ρ.

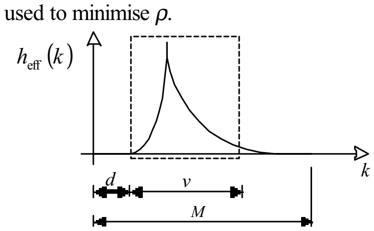

Figure 2 *Effective Channel Impulse Response*

Using vector notation eqn. (1) can be rewritten as

$$\mathbf{h}_{\text{eff}} = \mathbf{H}\mathbf{w} \tag{2}$$

where the $(M+p-1) \times p$ channel convolution matrix $\mathbf{H}$ has entries

$$\mathbf{H}(m,n) = h(m-n). \tag{3}$$

We start by partitioning the vector $\mathbf{h}_{\text{eff}}$ into the vx1 vector $\mathbf{h}_{\text{win}}$ (containing the samples of $\mathbf{h}_{\text{eff}}$ from within the window) and the (M-v)x1 vector $\mathbf{h}_{\text{wall}}$ (containing the samples of $\mathbf{h}_{\text{eff}}$ from outside the window). The residual ISI ρ can now be written

$$\rho = \frac{\mathbf{h}_{\text{wall}}^T \mathbf{h}_{\text{wall}}}{\mathbf{h}_{\text{win}}^T \mathbf{h}_{\text{win}}} = \frac{\mathbf{w}^T \mathbf{H}_{\text{wall}}^T \mathbf{H}_{\text{wall}} \mathbf{w}}{\mathbf{w}^T \mathbf{H}_{\text{win}}^T \mathbf{H}_{\text{win}} \mathbf{w}} = \frac{\mathbf{w}^T \mathbf{A} \mathbf{w}}{\mathbf{w}^T \mathbf{B} \mathbf{w}} \tag{4}$$

where the (v x p) and (M+p-v-1) x p matrices $\mathbf{H}_{\text{win}}$ and $\mathbf{H}_{\text{wall}}$ are extracted from $\mathbf{H}$ as:

The authors gratefully acknowledge the support of PMC-Sierra and Enterprise Ireland

$$\mathbf{H}_{\text{win}} = \begin{bmatrix} h_d & h_{d-1} & \cdots & h_{d-p+1} \\ h_{d+1} & h_d & \cdots & h_{d-p+2} \\ \vdots & \vdots & \ddots & \vdots \\ h_{d+v-1} & h_{d+v-2} & \cdots & h_{d+v-p+1} \end{bmatrix} \tag{5}$$

$$\mathbf{H}_{\text{wall}} = \begin{bmatrix} h_0 & 0 & \cdots & 0 \\ h_1 & h_0 & \ddots & 0 \\ \vdots & & & \vdots \\ h_{d-1} & h_{d-2} & \cdots & h_{d-p} \\ \hdashline h_{d+v} & h_{d+v-1} & \cdots & h_{d+v-p+1} \\ \vdots & \ddots & & \\ h_{M-1} & h_{M-2} & \cdots & h_{M-p} \\ 0 & h_{M-1} & \cdots & h_{M-p-1} \\ 0 & 0 & \ddots & \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & h_{M-1} \end{bmatrix} \tag{6}$$

Accordingly, the entries of **B** and **A** are given by

$$\mathbf{B}(m,n) = \sum_{i=d}^{d+v-1} h(i-m)h(i-n) \tag{7}$$

$$\mathbf{A}(m,n) = \sum_{i=0}^{M-1} h(i)h(i+n-m) - \sum_{i=d}^{d+v-1} h(i-m)h(i-n) \tag{8}$$

where we have used the fact that

$$\mathbf{A} = \mathbf{H}_{\text{wall}}^T \mathbf{H}_{\text{wall}} = \left[\mathbf{H}^T\mathbf{H} - \mathbf{H}_{\text{win}}^T \mathbf{H}_{\text{win}}\right] \tag{9}$$

To minimise ρ, we minimise $\mathbf{w}^T\mathbf{A}\mathbf{w}$ while constraining

$$\mathbf{w}^T \mathbf{B}\mathbf{w} = 1 \tag{10}$$

to avoid unbounded scaling of the shortened response.

3. MMSE CHANNEL SHORTENING

This approach was originally proposed by Falconer and Magee [3] as an optimisation criterion for channel shortening in a maximum likelihood receiver, and has been drawn on extensively in any subsequent work in relation to DMT. In particular, references [4], [5] and [6] treat the TEQ of a multi-carrier receiver in great detail using this approach.

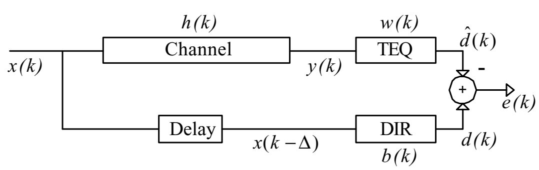

Figure 3 *MMSE Channel Shortening: the system is designed so the upper and lower paths match in a MMSE sense. h(k) is the original channel impulse response, which we wish to shorten. b(k) is the desired impulse response. Δ is simply a delay.*

Briefly, the theory is as follows. We generate an error signal by comparing the result of transmission of a training signal over each of the paths in Figure 3. The filter coefficients $w(k)$ and $b(k)$ are chosen to minimise the MSE value J of the error signal $e(k)$.

$$J = E\left[e^2(k)\right] = E\left[\left(d(k) - \hat{d}(k)\right)^2\right] \tag{11}$$

We have omitted system noise for the moment, which is discussed in Section 5. Expanding the relevant terms we get

$$J = E\left[\left(\sum_{l=0}^{v-1} b(l)x(k-\Delta-l) - \sum_{l=0}^{p-1} w(l)y(k-l)\right)^2\right] \tag{12}$$

At this juncture it is common to assume that the transmitted samples are uncorrelated, a fact represented by

$$E[x(i)x(j)] = \delta(i-j) \tag{13}$$

where $\delta(i)$ is the Kroeneker delta function. This allows us to write

$$J = \left[\begin{array}{l} \displaystyle\sum_{i=0}^{v-1} b(i)b(i) - 2\sum_{i=0}^{v-1}\sum_{m=0}^{p-1} b(i)w(m)h(i-m+\Delta) \\ \displaystyle + \sum_{n=0}^{p-1}\sum_{m=0}^{p-1}\sum_{j=0}^{M-1} w(m)w(n)h(j)h(j+n-m) \end{array}\right] \tag{14}$$

In vector notation this becomes

$$J = \mathbf{b}^T\mathbf{b} - 2\mathbf{b}^T\mathbf{R}_{\text{xy}}\mathbf{w} + \mathbf{w}^T\mathbf{R}_{\text{yy}}\mathbf{w} \tag{15}$$

where the v x p and p x p correlation matrices $\mathbf{R}_{\text{xy}}$ and $\mathbf{R}_{\text{yy}}$ respectively, have entries

$$\mathbf{R}_{\text{xy}}(m,n) = h(m-n+\Delta) \tag{16}$$

$$\mathbf{R}_{\text{yy}}(m,n) = \sum_{j=0}^{M-1} h(j)h(j+n-m) \tag{17}$$

To minimise the MSE we use partial differentiation.

$$\frac{\partial J}{\partial \mathbf{b}} = 2\mathbf{b} - 2\mathbf{R}_{\text{xy}}\mathbf{w} = 0 \tag{18}$$

$$\Rightarrow \mathbf{b}_{\text{opt}} = \mathbf{R}_{\text{xy}}\mathbf{w} \tag{19}$$

Back substitution of equation (19) into equation (15) yields

$$J = \mathbf{w}^T\left[\mathbf{R}_{\text{yy}} - \mathbf{R}_{\text{xy}}^T\mathbf{R}_{\text{xy}}\right]\mathbf{w} = \mathbf{w}^T\mathbf{R}_{\text{y|x}}\mathbf{w} \tag{20}$$

We have used the $\mathbf{R}_{\text{y|x}}$ notation to maintain consistency with the notation used in [2] and [7].

In minimising this we again apply the unit norm constraint, this time to avoid the trivial null solution. Now, however, the constraint is applied to the DIR coefficients $b(k)$ giving $\mathbf{b}^T\mathbf{b} = 1$, or from equation (19)

$$\mathbf{w}^T\mathbf{R}_{\text{xy}}^T\mathbf{R}_{\text{xy}}\mathbf{w} = 1 \tag{21}$$

4. EQUIVALENCE OF METHODS

We reiterate the two channel shortening methods under discussion:

RISIM: Minimise $\mathbf{w}^T\mathbf{A}\mathbf{w}$ while constraining $\mathbf{w}^T\mathbf{B}\mathbf{w} = 1$

MMSE: Minimise $\mathbf{w}^T\mathbf{R}_{\text{y|x}}\mathbf{w}$, constraining $\mathbf{w}^T\mathbf{R}^T_{\text{xy}}\mathbf{R}_{\text{xy}}\mathbf{w} = 1$

The first thing we note from equations (5) and (16) is that the matrices $\mathbf{H}_{\text{win}}$ and $\mathbf{R}_{\text{xy}}$ are in fact identical (by appropriate choice of Δ). Secondly, equations (8), (9) and (17) declare the equivalence of matrices $\mathbf{H}^T\mathbf{H}$ and $\mathbf{R}_{\text{yy}}$. These two facts can be used to deduce

$$\mathbf{B} = \mathbf{R}_{\text{xy}}^T\mathbf{R}_{\text{xy}} \tag{22}$$

$$\mathbf{A} = \mathbf{R}_{\text{y|x}} \tag{23}$$

This proves that the two channel shortening methods under discussion are identical. The solution to the minimisation problem is the same for each, and is taken as

$$\mathbf{w}_{\text{opt}} = \left(\sqrt{\mathbf{B}}^{T}\right)^{-1}\mathbf{q}_{\min} \quad (24)$$

where $\mathbf{q}_{\min}$ is the eigenvector corresponding to the minimum eigen-value $\lambda_{\min}$ of the matrix

$$\mathbf{C} = \left(\mathbf{Q}\sqrt{\Lambda}\right)^{-1}\mathbf{A}\left(\sqrt{\Lambda}\mathbf{Q}^{T}\right)^{-1} \quad (25)$$

The columns of $\mathbf{Q}$ consist of the orthonormal eigenvectors of $\mathbf{B}$, and Λ is a diagonal matrix with the entries of the eigenvalues. This solution relies on the positive definiteness of the matrix $\mathbf{B}$, which in turn depends on the channel convolution matrix $\mathbf{H}$ being of full rank [1]. It has been reported [1] that for most real channels the columns of $\mathbf{H}$ will indeed be linearly independent, and the matrix $\mathbf{B}$ will be positive definite. This has been verified by simulation for the twisted pair of the four VDSL test loops described in Section 6.

5. NOISE EFFECTS

We have made the claim that $\mathbf{A} = \mathbf{R}_{y|x}$ and $\mathbf{B} = \mathbf{R}^{T}_{xy}\mathbf{R}_{xy}$, which proves the equivalence of the two channel shortening methods under discussion. The question arises as to the validity of this claim when noise is present in the system.

Direct implementation of either method requires knowledge of the channel impulse response *h(k)*. The noise affecting the eigen-description matrices is thus a function of the channel identification procedure used at start-up. Since the matrices we are estimating are identical, we will get the same results in the presence of noise using either method, provided we use the same channel identification procedure in each case (clearly, this is true even if the added noise is coloured). So we conclude that we can achieve performance with the MMSE method, at least as good as that achievable using RISIM channel shortening.

A performance difference in the presence of noise may be discernable between the methods, however, if the MMSE method is implemented *adaptively* (which the RISIM method cannot be, in the form in which it is presented). In order to highlight this difference, consider the simple scenario, where added noise is white and gaussian distributed. On the one hand we determine the eigen-description matrices by channel identification at start-up and solve as before (RISIM); and on the other hand, we use a steepest descent algorithm to converge to the optimum eigen-solution (MMSE).

We can generate the RISIM eigen-description matrices using the simple channel identification procedure described in [5], whereby a periodic training sequence is used to generate an ensemble averaged frequency domain channel estimate, over L of its periods. If the system noise has variance σ_n^2, it is shown that the channel mean-square estimation-error is given by

$$E\left[\left(h_k - \hat{h}_k\right)^2\right] = \left(1 + \frac{1}{L}\right)\sigma_n^2 \quad (26)$$

(the hat symbol denotes an estimate), which converges asymptotically to σ_n^2 as we use more periods of the training sequence. This, along with equations (7) and (8), allows us to express the error variances for matrices $\mathbf{A}$ and $\mathbf{B}$ as

$$E\left[\left(\mathbf{A} - \hat{\mathbf{A}}\right)^2\right] = (M - v)\sigma_n^2\mathbf{I}_p \quad (27)$$

$$E\left[\left(\mathbf{B} - \hat{\mathbf{B}}\right)^2\right] = v\sigma_n^2\mathbf{I}_p \quad (28)$$

where $\mathbf{I}_p$ is the p x p identity matrix.

If, instead, we use adaptive MMSE channel shortening, without prior channel identification, uncorrelated white noise n_k of variance σ_n^2 added at the channel output will alter the MSE expression of equation (14) to read

$$J = \mathbf{b}^T\mathbf{b} - 2\mathbf{b}^T\mathbf{R}_{xy}\mathbf{w} + \mathbf{w}^T[\mathbf{R}_{yy} + \sigma_n^2\mathbf{I}_p]\mathbf{w} \quad (29)$$

Evidently, the resulting MSE expression will be similar to that of eqn. (20), except that $\mathbf{R}_{y|x}$ will now have the added noise term $\sigma_n^2\mathbf{I}_p$. This is less than the noise term added to $\mathbf{A}$ in (27), which indicates that we can achieve better channel shortening in the presence of white noise by an adaptive implementation of the MMSE method than we can with prior channel identification and a conventional RISIM TEQ.

The argument is incomplete, however. To achieve the MSE performance of equation (29) by adaptive methods would require an adaptive algorithm with zero excess MSE [8, pp.395]. Allowing an asymptotic analysis, whereby the adaptive algorithm has an infinitesimal step size and sufficient time to train, we should approach the MSE solution of equation (29). In this case our adaptive MMSE implementation will out-perform the non-adaptive RISIM TEQ. In real-life however, we will have an excess MSE term due to a finite step-size, and finite implementation time for the TEQ training algorithm. Furthermore, noise will in general be coloured due to crosstalk effects, and narrowband interference, adding cross-correlation terms to the MSE expression of equation (29). In general, therefore, we cannot expect the proposed MMSE method to outperform RISIM TEQ, although it may do in particularly benign noise environments.

We *can* guarantee the non-adaptive MMSE method will perform at least as well as the RISIM method in the presence of noise (white or coloured), since they are identical if both are implemented using the same channel identification procedure. This is the implementation described in the simulation results of Section 6.

6. SIMULATION RESULTS

Very high speed Digital Subscriber Loop (VDSL) refers to emerging ANSI and ETSI standards for high-speed (up to 55Mbaud unidirectionally) data transmission over existing twisted pair copper cabling in the telephone network.

We simulate channel-shortening equalisation for four of the ANSI standard VDSL test loops [9] shown in Figure 4. We attempt to shorten the channel to a length of 64 samples at a sampling rate of 20 MHz, as suggested in [10], using a 40-tap TEQ. An exhaustive search is used to find the optimum window delay at the receiver. System noise has been modelled as a combination of white gaussian noise (-140dBm flat power spectrum) and twenty ADSL NEXT interferers, added at the receiver. The crosstalk model used is that recommended in [9].

Noisy estimates of the VDSL channel impulse responses are made using a training sequence based channel identification procedure as described in Section 5. Eigen-system description matrices are generated accordingly, and the resulting TEQ settings are calculated by the constrained minimisation technique described earlier. Figures 5 and 6 show the channel impulse responses for test loops VDSL-4 and VDSL-7 both before and after equalisation (plots have been normalised for clarity). Vertical lines delimit the desired response windows.

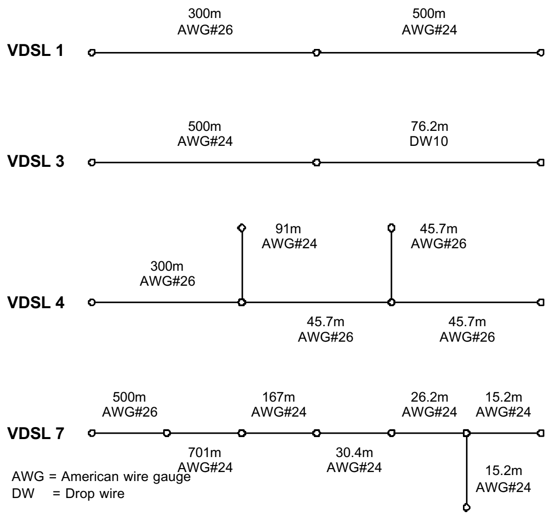

Figure 4 *VDSL Test Loops*

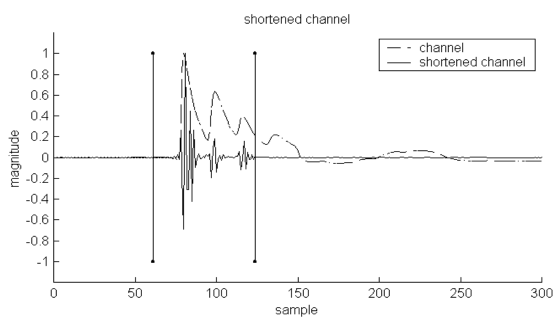

Figure 5 *Channel responses for test loop VDSL-4*

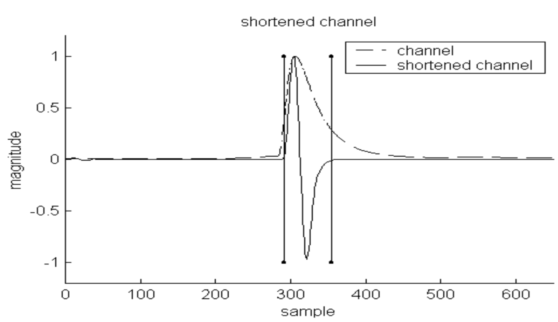

Figure 6 *Channel responses for test loop VDSL-7*

The channel shortening effects are immediately apparent. To give a more quantative description, residual ISI measurements have been made on four test loops, and are provided in Table 1. We reiterate that the measure of residual ISI ρ is the percentage of shortened impulse response energy outside the guard interval at the receiver. Note that only one set of results has been presented, since it was found that both methods of channel shortening presented here are in fact identical.

	Channel ISI	Residual ISI after Channel Shortening
Test Loop VDSL-1	11.71%	0.10%
Test Loop VDSL-3	12.87%	0.09%
Test Loop VDSL-4	36.50%	2.11%
Test Loop VDSL-7	24.35%	0.79%

Table 1 *ISI measurements in the VDSL system*

7. CONCLUSION

We have demonstrated the previously unreported fact that RISIM channel shortening as proposed by Melsa *et. al.* is in fact a minimum mean-squared error solution to the channel shortening problem. This proof gives a sound mathematical basis to the intuitive idea that minimising the impulse response energy outside a certain window will give good channel shortening. Furthermore, we are afforded an insight into MMSE channel shortening as originally proposed by Falconer and Magee [3]. Previously, it has not been intuitive as to why the Desired Impulse Response should be made adaptive in this approach. Our result demonstrates that allowing DIR adaptation achieves a minimisation of the effective impulse response energy outside the desired window of v samples. This interpretation also allows an adaptive implementation of Melsa's method without channel identification (although this implementation is not detailed as it has been studied extensively elsewhere [11], [6] in relation to MMSE channel shortening).

8. REFERENCES

[1] P.J.W. Melsa, R.C. Younce and C.E. Rohrs, "Impulse response shortening for discrete multi-tone transceivers", *IEEE Transactions on Communications*, vol. 44, pp. 1662-1672, Dec. 1996.

[2] G. Arslan, B.L. Evans and S. Kiaei, "Optimal channel shortening for discrete multi-tone transceivers", *IEEE International Conference on Acoustic Speech and Signal Processing*, (*ICASSP*), pp. 2965-2968, July 2000.

[3] D.D. Falconer and F.R. Magee "Adaptive channel memory truncation for maximum likelihood sequence estimation", *Bell Systems Technical Journal*, pp. 1541-1562, Nov.1973.

[4] M. V. Bladel and M. Moenceclaey, "Time-domain equalisation for multi-carrier communications", *IEEE Global Telecommunications Conference*, pp. 167-171, 1995.

[5] I. Lee, J. Chow and J.M. Cioffi, "Performance evaluation of a fast computation algorithm for the DMT in high speed subscriber loop", *IEEE Journal on Selected Areas in Communications*, Vol. 13, No. 9, pp. 1564-1570, Dec. 1995.

[6] J.S Chow and J.M Cioffi, "Equaliser training algorithms for multi-carrier communications systems", *IEEE International Conference on Communications*, (*ICC '93*), pp. 761-765, 1993.

[7] Y.H. You, J.H. Paik, H.K. Song, J.K. Kim, W.Y. Yang and Y.S. Cho, "Low-complexity channel shortening technique for DMT-based xDSL modems", *IEICE Trans. Commun.*, Vol.E82-B, No. 11, Nov 1999.

[8] S. Haykin, "Adaptive filter theory", 3rd ed., *Prentice-Hall*, 1996.

[9] VDSL Alliance, "Very-high-speed digital subscriber lines; System requirements; Draft Technical Report", T1E1.4/97-131R3, *American National Standards Institute*, 1997.

[10] VDSL Alliance, "SDMT VDSL Draft standrd proposal", T1E1.4/97-332, *American National Standards Institute*, 1997.

[11] T. Starr, J.M. Cioffi and P.J. Silverman, "Understanding digital subscriber line technology", *Prentice-Hall*, 1999