Abstract—The unit-norm constraint optimization for joint shortening of channel and echo impulse response is presented in this paper. The optimization is performed in the mean-square sense (MSE) and compared to unit-tap constraint joint shortening optimization. The proposed method could serve in performing joint echo cancellation and equalization for data transmission over Category cables for the next IEEE standard on short-range ultra high-speed copper interconnections used in data servers, high-performance computing centers, local area networks, etc. The main objective of the presented algorithm is to extensively reduce the power and implementation complexities of a 40GBASE-T system.

Index Terms—Echo cancellation, equalization, shortening impulse response.

I. INTRODUCTION

THE NEW IEEE objectives on ultra high-speed data transmission over copper cables have been recently announced. IEEE is pursuing the standardization of data transmission over copper wire beyond 10Gbps, namely 40 Gbps and 100 Gbps, for short range interconnections used in data servers, high-performance computing centers, local area networks, etc. Elsewhere, we have investigated the technical feasibility of data transmission at the rate of 40Gbps over Category-7A cables up to 50 m [1]. The main challenges in designing this system seem to be echo cancellation and designing forward error correcting code. Echo cancellation imposes very harsh constraints on the speed and precision of mixed-signal circuitry. Very high-bandwidth transmission makes the design of broadband hybrid circuits extremely difficult, sometimes even impractical, to achieve fair isolation between transmitter and receiver. The result of such imperfect isolation is very long echo impulse responses which have to be cancelled digitally. Unfortunately, long digital echo cancellers suffer from convergence issues and high quantization noise of fixed-point implementation [2].

One solution to this problem is the shortening impulse response (SIR) technique as proposed for xDSL applications. Channel shortening can be investigated with various objectives in mind; depending on the criterion adopted different methods for shortening the channel will be employed. Minimum mean square error (MMSE), maximum shortening signal-to-noise-ratio (MSSNR), minimum ISI, maximum bit rate, and others are some of the criteria extensively reviewed in the literature [3]–[8].

The most common approach in designing the shortening impulse response filter is the MMSE shortening which was first proposed by Falconer and Magee [6] in the context of maximum likelihood receiver design. They succeeded in substantially decreasing the computational complexity of the Viterbi algorithm by shortening the channel impulse response, hence reducing the system memory. Later on, the idea was applied to multicarrier modulation, essentially to reduce the overhead of added cyclic prefix for long channel responses. Melssa et al. [3], [4] tackled the problem by maximizing the shortening SNR and extended the idea to jointly equalize the channel and shorten the echo impulse response.

The main contribution of this paper is presenting the shortening impulse response technique to jointly equalize the channel and shorten the echo impulse response. We propose the MMSE joint shortening technique with a unit-norm constraint. In MMSE shortening filter design, it is well known that unit-norm constraint optimizations, generally, outperform the unit-tap constraints optimization [9].

The paper is organized as follows. In Section II, the theory of echo cancellation is discussed in some detail and the requirement of very long echo cancellers for 40GBASE-T application, to achieve proper cancellation levels, is pointed out. In Section III, we first briefly overview the unit-tap constraint joint channel and echo impulse response shortening. Simplifications to minimization algorithm due to the structure of underlying matrices are proposed. Then, we present our proposed unit-norm MMSE shortening filter design followed by the maximum shortening signal-to-noise-ratio technique. Performance evaluations through simulation results are presented in Section IV. We finish with our summary and conclusions.

Throughout the paper, bold face letters (\( \mathbf{x}, \mathbf{y} \)) denote column-vectors with real components (\( x_1, y_1 \)). Capital letters (A, B) denote matrices with real entries (\( a_{ij}, b_{ij} \)). Furthermore, we shall adopt MATLAB notations to present vectors, matrices, and concatenated variables.

II. ECHO CANCELLATION

The echo impulse response is denoted by \( \mathbf{p} \), which consists of \( \nu_{\mathbf{p}}+1 \) taps. The near-end (echo) data sequence \( \{ z_k \} \) is real, zero-mean, and has autocorrelation matrix \( \mathbf{R}_{zz} \). The noise sequence \( \nu_k \) is assumed real, zero-mean, independent of the input and echo sequences, and has an autocorrelation matrix \( \mathbf{R}_{\nu
\nu} \). At time
we define the tap-weight and the tap-voltage vectors as
\[
\hat{p} = [\hat{p}_0, \ldots, \hat{p}_{v_p}]^T,
\]
\[
z_{k:k-v_p} = [z_{k-v_p}, \ldots, z_{k-1}]^T.
\]
(1)

The constant echo channel impulse response is
\[
p = [p(\tau), p(T+\tau), \ldots, p(v_pT+\tau)]^T.
\]
(2)

where \(\tau\) is the sampling epoch chosen by the receiver. Note that, the transmitter pulse shaping is included in \(p(t)\). The sample at time \(kT\) of the received signal is
\[
r_k = z^T p + v_k
\]
(3)

where \(r_k\) is the desired far-end signal plus Gaussian noise at time \(kT\). The optimum tap-weight vector \(\hat{p}_{\text{opt}}\) minimizes the mean-square error \(J = E\{e_k^2\}\). We define the error signal at time \(kT\) as
\[
e_k = r_k - \hat{r}_k
\]
\[
= \sum_{l=0}^{v_p} p(l)z_{k-l} + v_k - \sum_{l=0}^{v_p} \hat{p}(l)z_{k-l}
\]
\[
= \sum_{l=0}^{v_p} (p(l) - \hat{p}(l))z_{k-l} + v_k + \sum_{l=v_p+1}^{v_p} p(l)z_{k-l}
\]
\[
= (p(z_{v_p} - \hat{p})^T z_{k:k-v_p} + u_k.
\]
(4)

The mean square error is then summarized to
\[
J = (p(z_{v_p} - \hat{p})^T R_{zz} (p(z_{v_p} - \hat{p})) + \sigma_n^2
\]

where
\[
\sigma_n^2 = \sigma_n^2 + p(z_{v_p+1:2v_p} R'_{zz} p(z_{v_p+1:2v_p})
\]
\[
R'_{zz} = E\{z_{v_p+1:2v_p} z_{v_p+1:2v_p}^T\}.
\]
(6)

The global minimizer of this quadratic form is \(\hat{p}_{\text{opt}} = p(z_{v_p})\), and consequently \(J_{\text{min}} = \sigma_n^2\). Thus, it only remains to obtain an estimate of \(p(z_{v_p})\) from the noisy received samples. The well known least-squares or least-mean-squares (LMS) estimations can obtain this estimate during a training sequence transmission. The solution to the LS estimate is [10]
\[
p_{\text{LS}} = (Z'Z)^{-1} Z'r_{v_p+1:N}
\]
(7)

where the Toeplitz matrix \(Z\) of dimension \((N - \hat{p}_p) \times (\hat{p}_p + 1)\) contains the last \(N - \hat{p}_p\) symbols of a block of \(N\) transmitted training symbols. If the additive Gaussian noise is not white, the LS estimate becomes
\[
p_{\text{LS}} = (Z' R''_{zz} Z)^{-1} R''_{zz} Z' r_{v_p+1:N}.
\]
(8)

The echo return loss enhancement is therefore defined as
\[
\text{ERLE} = 10 \log \frac{E\{r_k^2\}}{E\{e_k^2\}}.
\]

The echo impulse response of a typical CAT-7A cable is shown in Fig. 1. For this impulse response, the ERLE versus \(\hat{p}_p\) is calculated for 40 GBASE-T application over 50 and 100 m, illustrated in Fig. 2. Even long estimators with 1000 taps can not satisfy the cancellation level requirement of 40 GBASE-T, which is estimated to be about 60–65 dB [1].

A veracious inspection of the echo impulse response reveals that most of the energy is concentrated at the beginning of the impulse response and the rest is distributed over a large time span. However, most of these taps contain low energy and can be considered zero as long as their total effect is bellow some margin level. Therefore, a zero-tap detection process can find the negligible taps and omits the corresponding multipliers in convolution operation. This reduces the implementation complexity of the corresponding circuit in terms of gate counts, and consequently consumes less power. The zero-tap detection process is applied to a 1000-tap echo cancellor obtained by (7) at different threshold levels. Fig. 3 illustrates the resulting ERLE versus sparsity of the corresponding FIR filter, i.e., the number of zero-tap coefficients. Fig. 3 reveals that a 6-dB back-off from optimal point can reduce the complexity by about 20% (200 taps out of 1000).

The process of zero-tap detection can be modified by a more rigorous and accurate method stated as follows. One can obtain the least-square estimate (7) by solving the following norm optimization [10]:
\[
\min \|Zp - r_{v_p+1}\|_2^2
\]

Fig. 1. Echo impulse response of 100 m CAT-7A with 1400-MHz bandwidth sampled at Nyquist rate.

Fig. 2. ERLE versus number of least-square estimated echo canceller.
We define the class of $k$-*sparsity* containing vectors $\mathbf{x} \in \mathbb{B}^n$ with at least $k$ coordinates zero. One can find a sparse solution of the minimization problem by imposing the $k$-sparsity constraint, i.e.,

$$\begin{align*}
\text{minimize} & \quad ||\mathbf{Zp} - \mathbf{r}_{\eta j+1}||_2^2 \\
\text{subject to} & \quad p_j = 0, \quad j = j_1, j_2, \ldots, j_k \\
& \quad 1 \leq j \leq n.
\end{align*}$$

Unfortunately, we cannot predetermine what sparsity pattern, i.e., the subset $\{j_1, j_2, \ldots, j_k\}$, gives the best estimation. An exhaustive search towards the optimum sparsity patterns requires examining $n!/(k!(n-k)!)$ different combinations, which is quite impractical even for moderate size filter lengths. A heuristic approach for solving this problem is discussed in [14].

### III. CHANNEL SHORTENING FILTER DESIGN

#### A. MMSE

We consider the joint MMSE channel and echo impulse response shortening scenario depicted in Fig. 4. We follow the notation of Al-Dhahir in [11]. The problem and formulation presented in [11] are briefly stated as follows. The channel impulse response is denoted by $\mathbf{b}$, which consists of $v_h+1$ taps. The echo impulse response is as what was described earlier in Section II. Our objective is to shorten both channel and echo impulse responses, through a linear equalization, to the FIR filters $b_k$ and $c_k$, which consist of $v_h + 1$ and $v_c + 1$ taps, respectively. Also, we assume $v_h \gg v_c$ and $v_p \gg v_c$, otherwise the shortening does not make any sense. The far-end sequence data $\{x_k\}$ is also assumed real, zero-mean, with autocorrelation matrix $R_{xx}$.

The input-output relationship is given by

$$\mathbf{y}_{k+N-1:k} = \mathbb{H} \mathbf{x}_{k+N-1:k-1} + \mathbb{P} \mathbf{z}_{k+N-1:k-1} + \mathbf{n}_{k+N-1:k}. \quad (12)$$

The error sequence subject to mean square minimization is

$$e_k = \mathbb{E}^t \mathbf{x}_{k+N-1:k-1} + \mathbb{E}^t \mathbf{z}_{k+N-1:k-1} - \mathbb{E}^t \mathbf{y}_{k+N-1:k-1} = \left[ \begin{array}{c}
\mathbb{E}^t \mathbf{b} \\
\mathbb{E}^t \mathbb{c}
\end{array} \right] \left[ \begin{array}{c}
\mathbb{E}^t \mathbf{x}_{k+N-1:k-1} \\
\mathbb{E}^t \mathbb{z}_{k+N-1:k-1}
\end{array} \right] - \mathbb{E}^t \mathbf{y}_{k+N-1:k-1} \quad (13)$$

where $\mathbb{E}^t \mathbf{b}$ is a concatenation of $\Delta_b$ leading zeros with $\mathbf{b}$, followed by $s_1$ tail zeros, where $s_1 = N + v_h - v_b - \Delta_b$. $\mathbb{E}^t \mathbb{c}$ is defined in a same way, i.e., concatenation of $\Delta_c$ leading zeros with $\mathbf{c}$, followed by $s_2 = N + v_p - v_c - \Delta_c$ tail zeros. The optimal shortening filter is derived by applying the orthogonality principle, i.e., $\mathbb{E} \left\{ e_k \mathbf{y}_{k+N-1:k} \right\} = 0$. This results in an optimum shortening filter as

$$\mathbf{w}_{opt} = \left[ \begin{array}{c}
\mathbb{E}^t \mathbb{b} \\
\mathbb{E}^t \mathbb{c}
\end{array} \right] \left[ \begin{array}{c}
R_{xyy} \\
R_{zyy}
\end{array} \right] R_{yy}^{-1}. \quad (14)$$

The mean square error (MSE) is given by

$$\text{MSE} = \mathbb{E} \left\{ e_k^2 \right\} = [\mathbf{b} \mathbf{c}]^t \mathbf{R}^{-1} \mathbf{g} \mathbf{R}^t \mathbf{g} \quad (15)$$

where

$$\begin{align*}
\mathbf{R} = & \mathbf{q}^{-1} \left[ \begin{array}{cc}
R_{yy} & -R_{zyy} R_{yy}^{-1} R_{zyz}
\end{array} \right] R_{zyy} R_{yy}^{-1} R_{zyz} \\
R_{zy} = & \mathbf{R}_{xxy} R_{xyy} R_{yy}^{-1} R_{zyx} R_{zly} = \mathbf{R}_{xzy} \mathbf{R}_{zyy}^{-1} \mathbf{R}_{zyx} \mathbf{R}_{zly}\end{align*}$$

and

$$\mathbf{q} = \left[ \begin{array}{cc}
0, v_h + v_c + 2 \times \Delta_b, 0, v_h + v_c + 2 \times (s_1 + \Delta_c)
\end{array} \right] \mathbf{C}^{-1} \left[ \begin{array}{cc}
0, v_h + v_c + 2 \times \Delta_c
\end{array} \right].$$

The optimum solution to this quadratic optimization using brute-force search is extensively formulated in [11]. To avoid the trivial all-zero solution, the unit-tap constraint was suggested and imposed on this optimization, i.e., one of the coefficient taps of $\mathbf{b}$ and $\mathbf{c}$ are set to one. Although this calculation can be done once, it requires large matrix multiplications and inversions. The sparse structure of the matrix $\mathbf{q}$ suggests further investigations to reduce possible redundancies. If we rewrite the matrix $\mathbf{S}$ as the following block matrix:

$$\mathbf{S} = \left[ \begin{array}{ccccc}
\mathbf{S}_{11} & \mathbf{S}_{12} & \cdots & \mathbf{S}_{15} \\
\mathbf{S}_{21} & \mathbf{S}_{22} & \cdots & \mathbf{S}_{25} \\
\vdots & \vdots & \ddots & \vdots \\
\mathbf{S}_{51} & \mathbf{S}_{52} & \cdots & \mathbf{S}_{55}
\end{array} \right].$$

The sizes of submatrices $\mathbf{S}_{ij}$ are as follows:

$$\begin{align*}
\mathbf{S}_{1j} : & \quad \Delta_b \times \Delta_b \\
\mathbf{S}_{2j} : & \quad (v_h + 1) \times (v_h + 1) \\
\mathbf{S}_{3j} : & \quad (s_1 + \Delta_c) \times (s_1 + \Delta_c) \\
\mathbf{S}_{4j} : & \quad (v_c + 1) \times (v_c + 1) \\
\mathbf{S}_{5j} : & \quad s_2 \times s_2 \quad 1 \leq j \leq 5.
\end{align*}$$

![Fig. 3. ERLE versus sparsity of corresponding FIR echo canceller. The canceller obtained by least-square estimation has a total of 1000 taps.](image-url)
Then, after simple block-matrix multiplications, the matrix $q^T R q$ is reduced to

\[
q^T S q = \begin{bmatrix} B^T C \end{bmatrix} \begin{bmatrix} S_{22} & S_{24} \\ S_{42} & S_{44} \end{bmatrix} \begin{bmatrix} B \\ C \end{bmatrix} \triangleq [B^T C] A \begin{bmatrix} B \\ C \end{bmatrix}.
\]  

(20)

This can be further reduced to $[B^T C] = I(v_b + v_c + 2)$ noticing that\n
\[
[B^T C] = I(v_b + v_c + 2).
\]  

(21)

Therefore, the MSE is reduced to $g^T A g$ where the moderate size matrix $A$ will greatly reduce the computational complexity of the original minimization problem. However, it is known and reported several times in the literature that unit-norm constraint can achieve a better performance than unit-tap constraint in MMSE shortening filter design [9]. Here, we impose the unit-norm constraint on target channel response $b$, i.e.,

\[
\begin{align*}
\text{minimize} & \quad J_3(b, c) = [b^T c^T] A \begin{bmatrix} b \\ c \end{bmatrix} \\
\text{subject to} & \quad ||b||_2 = 1, \\
& \quad c_j = 1, \quad \text{for } 1 \leq j \leq m.
\end{align*}
\]

(22)

We should emphasize here that the matrix $A$ is not necessarily symmetric, but we conjecture that $A$ is a positive semidefinite (PSD) even though PSD is commonly defined for symmetric matrices [12], [13]. One can simply resolve this issue by replacing matrix $A$ with $0.5(A + A^T)$, which is obviously symmetric. Therefore, hereon, we shall assume $A$ is a PSD matrix.

If we expand the MSE in terms of properly sized block matrices $A_{11}$, $A_{12}$, and $A_{22}$

\[
J_1(b, c) = b^T A_{11} b + 2b^T A_{12} c + c^T A_{22} c.
\]  

(23)

One can work out this problem iteratively by solving it first with respect to $c$ assuming $b$ is fixed. The solution to this unit-tap minimization problem, $c^*(j)$, is presented in Appendix A for a general quadratic programming (QP). The minimizer of this quadratic form is then replaced in our original objective function $J_1$

\[
J_2(b) = J_3(b, c^*(j)) = b^T \left( A_{11} - A_{12} A_{22}^{-1} A_{12}^T \right) b + \left( 1 + b^T A_{12} A_{22}^{-1} A_{12}^T (s, j) \right)^2 A_{22}^{-1} (j, j).
\]  

(24)

The resulting objective function is another quadratic form that can be minimized with the unit-norm constraint. The solution to this problem is discussed in Appendix B. This process can be iterated with respect to index $j$ to find out the global minimizers $b^*$ and $c^*$. This is summarized in the Algorithm 1.

Algorithm 1: Unit-norm and unit-tap joint optimization

\[
J_{\text{min}} = +\infty
\]

for $j = 1 : m$

\[
\begin{align*}
A &= A_{11} - A_{12} A_{22}^{-1} A_{12}^T + \cdots \\
& \quad A_{12} A_{22}^{-1} A_{12}^T (s, j) / A_{22}^{-1} (j, j) \\
A_{22}^{-1} (j, j) &= \left[ A_{22}^{-1} (j, j) \right]^T A_{22}^{-1} (j, j) \\
\end{align*}
\]

\[
b = \left[ A_{22}^{-1} (j, j) \right]^T A_{12} A_{22}^{-1} (j, j)
\]

\[
\begin{align*}
\text{minimize} & \quad x^T A x + 2b^T x + 1/A_{22} (j, j) \\
\text{subject to} & \quad ||x||_2 = 1 \\
x(j) = x^* \\
y(j) = \cdots A_{12} A_{22}^{-1} A_{12}^T \left( s, j \right) / A_{22}^{-1} (j, j) - A_{22} x^* \\
J_j = J_x (x(j), y(j)) \quad \text{if } J_j < J_{\text{min}} \\
x_{\text{opt}} = x(j) \\
y_{\text{opt}} = y(j) \\
J_{\text{min}} = J_j
\]

end

end

We now show that the norm-constraint on $b$ is sufficient to avoid trivial solutions. In fact, having only norm-constraint on target channel impulse response, automatically results in a nonzero solution for target echo impulse response. The optimization problem can be stated as

\[
\begin{align*}
\text{minimize} & \quad J_3(b) = [b^T c^T] A \begin{bmatrix} b \\ c \end{bmatrix} \\
\text{subject to} & \quad ||b||_2 = 1.
\end{align*}
\]  

(25)

We shall now prove that the global solution to the optimization problem expressed by (25) results in a nonzero answer for column vector $c$. The second-order cone $\Omega = \{ c \mid 2b^T A_{12} c + c^T A_{22} c \leq 0 \}$ [14] is a nonempty subspace, therefore $J_1(b, c = 0) = b^T A_{11} b$ is bigger than $J_1(b, c)$ for some nonzero $c$. Therefore, if $b^*$ is the global minimizer of $J_3(b)$, then $J_3(b^*) \leq J_1(b, c = 0)$. This allows concluding that a single unit-norm constraint is sufficient to avoid a nonzero solution for the target echo impulse response.

To solve the problem (25), we incorporate the unit-norm constraint into an auxiliary function

\[
L(b, c, \lambda) = [b^T c^T] \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^T & A_{22} \end{bmatrix} \begin{bmatrix} b \\ c \end{bmatrix} + \lambda (1 - b^T b),
\]

(26)

Applying the optimality conditions yields

\[
\begin{align*}
\frac{\partial L}{\partial b} &= A_{11} b + A_{12} c - \lambda b = 0 \\
\frac{\partial L}{\partial c} &= A_{12}^T b + A_{22} c = 0.
\end{align*}
\]  

(27)

Given $A_{22}$ is invertible, the second optimality condition gives

\[
c = -A_{22}^{-1} A_{12}^T b.
\]  

(28)

This equation explicitly indicates that, given $b$ is nonzero, the solution for $c$ is nonzero. Putting $c$ back into the first optimality condition results in the following equation:

\[
(A_{11} - A_{12} A_{22}^{-1} A_{12}^T) b = \lambda b.
\]  

(29)

This is a standard eigenvalue-eigenvector problem; however, it needs some extra normalization and manipulations to obtain the minimizer of our quadratic optimization. The details are presented in the following algorithm. We incorporated the MATLAB notations for the sake of implementation ease.
Algorithm 2: Unit-norm joint optimization

\[ M = A_{11} - A_{12} A_{22}^{-1} A_{12}^T \]
\[ [X; D] = \text{eig}(M) \]
for \( k = 1 : \text{size}(X, 2) \)
\[ X(:,k) = X(:,k)/\|X(:,k)\|^2 \]
end
\[ Y = -A_{22}^{-1} A_{12} X \]
\[ Z = [X; Y] \]
\[ Z = [Z - Z] \]
\[ J_{\text{min}} = +\infty \]
for \( k = 1 : \text{size}(Z, 2) \)
\[ J(k) = Z(:,k)^T \ast A \ast Z(:,k) \]
if \( J(k) < J_{\text{min}} \)
\[ J_{\text{min}} = J(k) \]
\[ b_{\text{opt}} = Z(1 : m, k) \]
\[ c_{\text{opt}} = Z(m + 1 : \text{end}, k) \]
\[ k_{\text{min}} = k \]
end
end

B. MSSNR

Here, for the sake of completeness and comparison, we overview the maximum shortening SNR approach for the joint shortening filter design. We adopt our notation to the formulation derived in [3]. The equalized channel \( h_{eq} = Hw \). can be partitioned as

\[ h_{\text{win}} = H(\Delta_b + 1 : v_b + \Delta_b + 1,:)w \triangleq H_{\text{win}}w. \]
(30)

and

\[ h_{\text{wall}} = H([1 : \Delta_b, \Delta_b + v_b + 1,:)w \triangleq H_{\text{wall}}w. \]
(31)

Hence, \( h_{\text{win}} \) denotes the part of the equalized channel inside the shortening window, and \( h_{\text{wall}} \) represents the part of the equalized channel outside the shortening window. \( p_{\text{win}} \) and \( p_{\text{wall}} \) are defined similarly.

Here, instead of minimizing the mean square error, the optimization criterion is to maximize the ratio of signal energy inside the shortening window to the signal energy spread outside the shortening window. In fact, the equalizer performance depends on how well it concentrates the signal into a target window. The corresponding energies can be found from

\[ E_{\text{win}} = \beta h_{\text{win}}^T h_{\text{win}} + (1 - \beta)p_{\text{win}}^T p_{\text{win}} = w^T B(\beta)w \]
\[ E_{\text{wall}} = \alpha h_{\text{wall}}^T h_{\text{wall}} + (1 - \alpha)p_{\text{wall}}^T p_{\text{wall}} + \sigma_u^2 \]
\[ = w^T A(\alpha)w \]
(32)

where \( \sigma_u^2 \) is the power of the filtered noise \( u_k = w^T \tilde{\nu}_{k+N-1:k} \) and

\[ B(\beta) = \beta H_{\text{win}}^T H_{\text{win}} + (1 - \beta)P_{\text{win}}^T P_{\text{win}} \]
\[ A(\alpha) = \alpha H_{\text{wall}}^T H_{\text{wall}} + (1 - \alpha)P_{\text{wall}}^T P_{\text{wall}} + R_{\text{uv}}. \]
(33)

The parameters \( \alpha \) and \( \beta \) balance the relationship between the energy in and out of the window for both the channel and echo. The shortening SNR is defined as

\[ \frac{E_{\text{win}}}{E_{\text{wall}} + E_{\text{noise}}} = \frac{w^T B(\beta)w}{w^T A(\alpha)w}. \]
(34)

The SSNR can be maximized either by minimizing \( w^T A(\alpha)w \) subject to unit-energy constraint \( w^T B(\beta)w = 1 \), or by maximizing \( w^T [B(\beta)w] \), subject to \( w^T A(\alpha)w = 1 \). If \( B \) is not singular, the solution for the first optimization problem is

\[ w_{\text{opt}} = \left( \sqrt{B} \right)^{-1} e_{\text{min}} \]
(35)

where \( \sqrt{B} \) is the Cholesky decomposition of \( B \) [12], and \( e_{\text{min}} \) is the eigenvector corresponding to the minimum eigenvalue of matrix \( C = \left( \sqrt{B} \right)^{-1} A \left( \sqrt{B} \right)^{-T} \). The maximum shortening signal-to-noise-ratio (MSSNR) is defined as

\[ \text{MSSNR} = \text{SNR}_{\text{max}} = -10 \log_{10}(\lambda_{\text{min}}(C)). \]
(36)

This analysis is based on nonsingularity assumption of matrix \( B \). Note that, \( B \) is singular if the channel shortening filter length is larger than the length of the shortening window, i.e., \( v_b > v_c \), which is the case, generally. One can follow the maximization problem and obtain similar solution.

IV. PERFORMANCE EVALUATIONS

The channel and echo impulse responses of CAT-7A copper cable are used in our simulations. Although the channel impulse response is directly affected by cable length, its effect on echo impulse response is very minor since echo is a measure of the signal reflections occurring along a transmission line and related to impedance mismatch in a cable channel. However, different cable lengths mean different optimum signal bandwidths, and this changes the time spreading of echo signal of 40GBASE-T application over different lengths. For simulation purposes, we use the characteristics of a 50m cable. The signal bandwidth is about 1400 MHz and it is sampled at Nyquist rate. The noise power spectral density is assumed \(-115 \text{ dBm/Hz}\). Other far-end (FEXT) and near-end (NEXT) crosstalks are not considered here. A 250-taps symbol-spaced linear equalizer is assumed. The length of target channel and echo impulse responses are set to 80 and 650, respectively. The optimum delay parameters are \( (\Delta_b, \Delta_c) = (25, 2) \). The minimum MSE after this joint shortening by a linear equalizer is \( 1.33 \times 10^{-6} \). Fig. 5 depicts the original and shortened channel and echo impulse responses. Obviously, we can achieve further reductions in MSE by increasing \( N, v_b \), or \( v_c \).

A comparison to the proposed method is made considering both MMSE unit-tap joint shortening and MSSNR joint shortening. The unit-tap MMSE joint shortening presented in [11] was performed with the same set of parameters. The minimum
achievable MSE was $2.56 \times 10^{-5}$ with optimum delay parameters $(\Delta_t, \Delta_e) = (25, 4)$. As we can see, the proposed method outperforms the unit-norm joint shortening by 13 dB. Both methods have very similar estimates of the optimum delay parameters. We also carried out the MSSNR joint-shortening methods have very similar estimates of the optimum delay parameters. We also carried out the MSSNR joint-shortening [14] using the same $(\Delta_t, \Delta_e)$ obtained from our proposed method. The resulting MMSE was $1.22 \times 10^{-6}$ which is only $.37$ dB better than our proposed joint shortening.

V. CONCLUSIONS

To support a data rate of 40 Gbps over copper wire, for 40GBASE-T application, a very long FIR echo canceller is required to achieve a proper cancellation level. We propose the MMSE joint channel shortening technique to preserve the implementation costs, power consumption and convergence of (possibly adaptive) FIR echo cancellers. The previously proposed unit-tap MMSE joint channel and echo impulse response shortening is revisited and the computational complexity of search algorithm is reduced by carefully examining the structure of underlying matrices. The unit-norm constraint is proposed for target channel impulse response in the MMSE joint shortening framework, and the corresponding algorithm is offered. The unit-norm constraint is already proved to outperform the unit-tap constraint, which was the main motivation of our work. We also present a simple proof that the unit-norm constraint on target channel impulse response is sufficient to avoid nontrivial answer for target echo impulse response. The analytical solution for MMSE joint shortening problem subjected to a single norm-constraint is formulated and the MATLAB algorithm is presented.

APPENDIX A

This Appendix presents the unit-tap constraint quadratic optimization. We are interested in the following optimization problem:

$$\begin{align*}
\text{minimize} \quad & f_0(x) = x^t A x + 2b^t x \\
\text{subject to} \quad & x_j = 1, \ \exists j \in \mathcal{J}
\end{align*}$$

(A.1)

where $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{n \times n}$ is symmetric positive definite, $b \in \mathbb{R}^n$, and finally $\mathcal{J}$ is any subset of indices $\{1, 2, \ldots, n\}$. The constraint $x_j = 1$ can be replaced by $x^t \epsilon_j = 1$ where $\epsilon_j$ is the $j$th unit vector. The standard approach to solve this problem is Lagrange Multiplier method as follows:

$$L(x, \lambda) = x^t A x + 2b^t x + \lambda (1 - x^t \epsilon_j).$$

(A.2)

The Euler optimality condition, i.e., $\partial L/\partial x = 0$ gives

$$x = A^{-1} (\epsilon_j - b).$$

(A.3)

Applying the given constraint, the optimum values for $\lambda$ and $x$ are

$$\lambda^* = \frac{1 + b^t A^{-1} (\epsilon_j)}{A^{-1}(j, j)},$$

$$x^* = A^{-1} \left( \frac{1 + b^t A^{-1} (\epsilon_j)}{A^{-1}(j, j)} \right) \epsilon_j - b.$$  

(A.4)

After simple manipulation, the minimum value of quadratic objective function $f_0(x^*)$ can be calculated as

$$f_0(x^*) = \frac{(1 + b^t A^{-1} (\epsilon_j))^2}{A^{-1}(j, j)} - b^t A^{-1} b.$$  

(A.5)

The second term, $b^t A^{-1} b$, is independent of $j$, therefore the optimum $j$ that results in a global minimum is

$$j^* = \arg \min_{j \in \mathcal{J}} \frac{(1 + b^t A^{-1} (\epsilon_j))^2}{A^{-1}(j, j)}.$$  

(A.6)

APPENDIX B

In this Appendix, we review the following norm-constraint optimization problem:

$$\begin{align*}
\text{minimize} \quad & f_0(x) = x^t A x + 2b^t x \\
\text{subject to} \quad & \|x\|_2 = 1.
\end{align*}$$

(B.1)

The Lagrangian of (B.1) is defined by

$$L(x, \lambda) = x^t (A - \lambda I) x + 2b^t x + \lambda.$$  

(B.2)

Setting the gradient of the Lagrange function to zero results in the normal equation

$$(A - \lambda I) x = -b.$$  

(B.3)

In general, there are many pairs $(\lambda, x)$ satisfying the normal equation. Moreover, for fixed multiplier $\lambda = \mu_j$ with the $j$th eigenvalue of $A$, the unconstrained normal equation may have multiple solutions. If the multiplier $\lambda = \mu_j$ for all $j = 1, \ldots, n$, the corresponding vector $x$ is uniquely determined by $\lambda$ and $x = (A - \lambda I)^{-1} b.$
In this case, the multipliers required are the roots of the secular equation
\[ f(\lambda) = \|(A - \lambda I)^{-1}b\|_2 = 1. \] (B.5)

We are interested in the solutions \((\lambda, x)\) with respect to the optimal multiplier \(\lambda^*\). It is shown in [15] that the minimizer of problem (B.1) is given as the solution of the normal equation with the largest. We would not get into many other details on the feasibility conditions and solving the normal equation as they are discussed comprehensively in [15].

Another approach to solve the problem (B.1) is as follows. The dual function of the primal problem is
\[ g(\lambda) = \inf_{x} L(x, \lambda) = \begin{cases} \lambda - b'(A - \lambda I)^{-1}b, & (A - \lambda I) \geq 0 \\ -\infty, & \text{otherwise} \end{cases} \] (B.6)

Using a Schur complement [14], [16], we can express the dual problem as
\[
\begin{align*}
\text{maximize} & \quad \gamma \\
\text{subject to} & \quad \lambda \geq \lambda^* \\
& \quad \left[ \begin{array}{cc}
A - \lambda I & b' \\
-\gamma & -\gamma
\end{array} \right] \succeq 0.
\end{align*}
\] (B.7)

This is a semidefinite program for which there are very efficient algorithms already developed to solve them [15]. After obtaining the optimum \(\lambda^*\), one can find the optimum \(x^*\) from (B.2) applying the Euler optimality condition, i.e., \(\partial L/\partial x = 0\).
\[
x^* = -(A - \lambda^* I)^{-1}b. \quad (B.8)
\]

ACKNOWLEDGMENT

The authors gratefully acknowledge the kind support and helpful discussions with Nexans’ engineers and the tools and services of CICTR and high-performance computing services at The Pennsylvania State University.

REFERENCES


Ali Enteshari (S’05) received the B.Sc. and M.Sc. degrees in electrical engineering from Sharif University of Technology, Tehran, Iran, in 1999 and 2001, respectively. He is currently pursuing the Ph.D. degree at The Pennsylvania State University, University Park.

Between September 1999 and September 2001, he was with Signal Company, Tehran, developing signal processing algorithms and implementing high-speed circuits. He was later a Senior Engineer with Micro Modje, Inc., Tehran, involved in microwave radio design and baseband processing. From 2003 to 2006, he worked on mixed-signal chip design, image sensors and linear high-efficiency RF amplifiers (LINC). His research interests include communications theory, equalization and precoding, coding, multi-input multi-output (MIMO) systems, and OFDM techniques.

Jarir M. Fadullah (S’08) received the B.S.E.E. degree in 2004 from Bangladesh University of Engineering and Technology, Dhaka, Bangladesh. He is currently pursuing the Ph.D. degree in electrical engineering at The Pennsylvania State University, University Park.

From 2004 to 2005, he was a Lecturer in the Department of Electrical Engineering, Islamic University of Technology, Gazipur, Bangladesh. Currently, he is with the Center for Information and Communications Technology Research (CICTR), Penn State. His research interests are multiple-input multiple-output communication system analysis, wireless MIMO channel modeling, and MIMO signal processing. He has experience working with WiMAX systems.

Mohsen Kavehrad (F’92) is the Weiss Chair Professor and Director of the Center for Information and Communications Technology Research, Electrical Engineering Department, The Pennsylvania State University, University Park. He has over 350 published papers, book chapters, books, and key patents. His research interests are in the areas of wireless and optical communications networked systems.

Dr. Kavehrad received three Bell Labs awards, a 1990 TRIO Feedback Award for a patent on optical interconnect, a 2001 IEEE VTS Best Paper Award, three IEEE LEOS Best Paper Awards, and a Canada NSERC Ph.D. dissertation award (with his graduate students) for contributions to wireless and optical networks.