EIGENFILTER DESIGN OF MIMO EQUALIZERS FOR CHANNEL SHORTENING

Andre Tkacenko and P. P. Vaidyanathan

Dept. of Electrical Engineering
California Institute of Technology, Pasadena, CA 91125, USA
E-mail: andre@systems.caltech.edu, ppvnath@systems.caltech.edu

ABSTRACT

The advent of discrete multitone modulation (DMT) systems in recent years has brought to light the importance of channel shortening equalizers. In this paper, we present a method for the design of one such equalizer for multiple-input multiple-output (MIMO) linear dispersive channels. This method is a generalization of one that was used for shortening of single-input single-output (SISO) channels. Experimental results presented show that our design method performs better than the minimum mean-squared error (MMSE) technique in terms of effective channel energy compaction. [1]

1. INTRODUCTION

The design of time-domain equalizers or TEQs for discrete multitone modulation or DMT systems has received much attention of late. As a result of the long impulse response of typical channels encountered in DMT systems such as twisted pair telephone lines [5], TEQs are necessary to *shorten* the overall channel response to one sample more than the length of the cyclic prefix used. Since the channel encountered in a traditional DMT system is a single-input single-output (SISO) channel, most if not all design methods for TEQs have been only for SISO channels.

More recently, a generalization of the DMT structure to accomodate vector signals was introduced [3] called the discrete matrix multitone (DMMT) system. Unlike the traditional DMT structure, with the DMMT system, a multiple-input multiple-output (MIMO) channel is encountered. If the impulse response of this channel is too long, then as before, an equalizer is required to shorten the overall channel. A method for the design of such an equalizer was very recently proposed in [1] based on a minimum mean-squared error (MMSE) criterion.

In this paper, we consider a different method for the design of MIMO channel shortening equalizers. Our proposed method is a generalization of one originally considered in [4] for SISO channels and later considered by the authors in [7] for single-input multiple-output (SIMO) channels. The equalizer coefficients will be found to be optimal with respect to an objective function which takes into account both the desire to shorten the channel and also to minimize the output noise power with respect to the signal power. It will be seen that the optimal filter coefficients will be related to the components of an eigenvector of a particular matrix.

[1]Work supported in part by the ONR grant N00014-99-1-1002, USA.

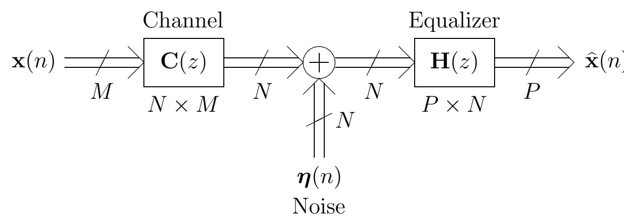

Fig. 1. MIMO channel-equalizer model.

2. EQUALIZER DESIGN PROBLEM

Suppose that we have the MIMO channel-equalizer model shown in Figure 1. The details in the development of the channel shortening problem do not require the restriction $P = M$, even though this is the case in practice where $\widehat{\mathbf{x}}(n)$ represents the received version of $\mathbf{x}(n)$. We make the following assumptions.

- The channel $\mathbf{C}(z)$ is FIR of length L_c.
- The equalizer $\mathbf{H}(z)$ is FIR of length L_e.
- The input signal vector $\mathbf{x}(n)$ is zero-mean and white with autocorrelation sequence $\mathbf{R}_{\mathbf{xx}}(k) = \sigma_x^2 \mathbf{I}\delta(k)$.
- The noise vector sequence $\boldsymbol{\eta}(n)$ is a zero-mean WSS random process uncorrelated with $\mathbf{x}(n)$ and with autocorrelation sequence $\mathbf{R}_{\boldsymbol{\eta\eta}}(k)$.

If our goal is to simply equalize the channel, then clearly the best thing to do in the mean-squared sense is to choose the coefficients of $\mathbf{H}(z)$ to be the Wiener filter [6] for which the desired signal is $\mathbf{x}(n)$. On the other hand, if instead our goal is to shorten the overall channel to a length L_d, then one possibility is to choose $\mathbf{H}(z)$ to be the Wiener filter for which the desired signal is the filtered signal $\mathbf{b}(n) * \mathbf{x}(n)$, where $\mathbf{b}(n)$ is some $P \times M$ FIR filter of length L_d. The problem then resides in the particular choice of $\mathbf{b}(n)$. This design method and problem were considered in [1].

For our method, we address the channel shortening problem in a more direct manner. Denote the impulse responses of $\mathbf{C}(z)$ and $\mathbf{H}(z)$ by $\mathbf{c}(n)$ and $\mathbf{h}(n)$, respectively. The effective channel is $\mathbf{c}_{\text{eff}}(n) = \mathbf{h}(n) * \mathbf{c}(n)$ and has length $L_c + L_e - 1$. Note that the output $\widehat{\mathbf{x}}(n)$ can be expressed as follows.

$$\widehat{\mathbf{x}}(n) = \mathbf{s}(n) + \mathbf{w}(n)$$

where $\mathbf{s}(n)$ and $\mathbf{w}(n)$ are, respectively, the output signal and output noise sequences given by,

$$\mathbf{s}(n) = \mathbf{c}_{\text{eff}}(n) * \mathbf{x}(n)\,,\ \mathbf{w}(n) = \mathbf{h}(n) * \boldsymbol{\eta}(n)$$

Because of the presence of noise, we wish to choose the coefficients of the equalizer to accomplish the following goals.

- Shorten the effective channel $\mathbf{c}_{\text{eff}}(n)$ to a length L_d.
- Minimize the noise power with respect to the signal power.

Note that since $\mathbf{c}_{\text{eff}}(n) = \mathbf{h}(n)*\mathbf{c}(n)$, any particular component of $\mathbf{c}_{\text{eff}}(n)$ depends only upon one row of $\mathbf{h}(n)$. As such, if $\mathbf{c}_{\text{eff},l}(n)$ and $\mathbf{h}_l(n)$ denote, respectively, the l-th rows of $\mathbf{c}_{\text{eff}}(n)$ and $\mathbf{h}(n)$ where $0 \leq l \leq P-1$, then we have $\mathbf{c}_{\text{eff},l}(n) = \mathbf{h}_l(n) * \mathbf{c}(n)$. Hence, it makes sense to choose the coefficients of $\mathbf{H}(z)$ on a row-by-row basis. With the above goals in mind, we opt to choose the rows of $\mathbf{h}(n)$ to minimize the following objective function J.

$$J \triangleq \sum_{l=0}^{P-1} p_l J_l \tag{1}$$

where J_l for $0 \leq l \leq P-1$ is given by,

$$J_l \triangleq \alpha_l J_{\text{short},l} + (1-\alpha_l) J_{\text{noise},l}\,,\ 0 \leq \alpha_l \leq 1 \tag{2}$$

and $J_{\text{short},l}$ and $J_{\text{noise},l}$ are defined as follows.

$$J_{\text{short},l} \triangleq \frac{\sum_n \mathbf{c}_{\text{eff},l}(n)\mathbf{F}_l(n-\Delta)\mathbf{c}^\dagger_{\text{eff},l}(n)}{\sum_n \mathbf{c}_{\text{eff},l}(n)\mathbf{c}^\dagger_{\text{eff},l}(n)} \tag{3}$$

$$J_{\text{noise},l} \triangleq \frac{\sigma^2_{w_l}}{\sigma^2_{s_l}} = \frac{[\mathbf{R}_{\mathbf{ww}}(0)]_{l,l}}{[\mathbf{R}_{\mathbf{ss}}(0)]_{l,l}} \tag{4}$$

Here, p_l is a weighing parameter for the l-th row used to emphasis the design of one row over another. The sequence $\{p_l\}$ is a probability density function or pdf in the sense that $p_l \geq 0$ for all l and $\sum_l p_l = 1$. In addition, the quantity $J_{\text{short},l}$ represents a channel shortening objective function for the l-th row whereas $J_{\text{noise},l}$ is the noise-to-signal ratio observed in the l-th component of the output signal $\widehat{\mathbf{x}}(n)$. The parameter Δ denotes the desired delay for the shortened channel. We must have $0 \leq \Delta \leq L_c + L_e - L_d - 1$ here. In practice, Δ is varied over all of these values and chosen so as to minimize the objective function J in (1).

The $M \times M$ matrix sequence $\mathbf{F}_l(n)$ is a "penalty" matrix function that should satisfy the following conditions.

- $\mathbf{F}_l(n)$ should be positive semidefinite for all l, n.
- $\mathbf{F}_l(n)$ should "penalize" values of $\mathbf{c}_{\text{eff},l}(n)$ that occur outside of the interval $\mathcal{W} \triangleq [\Delta, \Delta + L_d - 1]$.

Since, in addition to shortening the channel, we would also like the effective channel to be of full rank, we need to choose $\mathbf{F}_l(n)$ such that the rows of $\mathbf{h}(n)$ are linearly independent. To help ensure this, we have decided to choose $\mathbf{F}_l(n)$ to penalize the off-diagonal elements of $\mathbf{c}_{\text{eff},l}(n)$ for $n \in \mathcal{W}$. For the purpose of channel shortening, we have chosen to use the following penalty function.

$$\mathbf{F}_l(n) = \begin{cases} \beta\left(\mathbf{I}_M - \mathbf{E}_{l,l}\right), & 0 \leq n \leq L_d - 1 \\ \gamma \mathbf{I}_M, & \text{otherwise} \end{cases}, \ \forall\ l, n$$

where $0 < \beta \leq \gamma$ and $\mathbf{E}_{k,l}$ is the (k,l)-th elementary matrix which is one in the (k,l)-th component and zero elsewhere. Then,

$$J_{\text{short},l} = \frac{\beta \sum_{n\in\mathcal{W}} \left(||\mathbf{c}_{\text{eff},l}(n)||^2 - \left|[\mathbf{c}_{\text{eff},l}(n)]_{0,l}\right|^2 \right) + \gamma \sum_{n \notin \mathcal{W}} ||\mathbf{c}_{\text{eff},l}(n)||^2}{\sum_n ||\mathbf{c}_{\text{eff},l}(n)||^2}$$

and thus minimizing $J_{\text{short},l}$ is tantamount to shortening $\mathbf{c}_{\text{eff},l}(n)$.

Note that J_l is a *convex combination* of the objective functions $J_{\text{short},l}$ and $J_{\text{noise},l}$ and that $J_l \geq 0$ by the assumed positive semidefiniteness of $\mathbf{F}_l(n)$. Here, α_l is a tradeoff parameter between channel shortening and minimizing the output noise-to-signal ratio. We now proceed to analyze the objective function J_l.

3. ANALYSIS OF THE OBJECTIVE FUNCTION J_L

First we will analyze $J_{\text{short},l}$. In order to do so, let us define the following vectors and matrices.

$$\begin{aligned} \mathbf{c}_{\text{eff},l} &\triangleq \begin{bmatrix} \mathbf{c}_{\text{eff},l}(0) & \mathbf{c}_{\text{eff},l}(1) & \cdots & \mathbf{c}_{\text{eff},l}(L_c+L_e-2) \end{bmatrix} \\ \mathbf{h}_l &\triangleq \begin{bmatrix} \mathbf{h}_l(0) & \mathbf{h}_l(1) & \cdots & \mathbf{h}_l(L_e-1) \end{bmatrix} \end{aligned}$$

$$\mathbf{C} \triangleq \begin{bmatrix} \mathbf{c}(0) & \mathbf{c}(1) & \cdots & \mathbf{c}(L_c-1) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{c}(0) & \mathbf{c}(1) & \cdots & \mathbf{c}(L_c-1) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{c}(0) & \mathbf{c}(1) & \cdots & \mathbf{c}(L_c-1) \end{bmatrix}$$

$$\boldsymbol{\Lambda}_l \triangleq \begin{bmatrix} \mathbf{F}_l(0-\Delta) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_l(1-\Delta) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{F}_l((L_c+L_e-2)-\Delta) \end{bmatrix}$$

These quantities have the following sizes.

- $\mathbf{c}_{\text{eff},l} : 1 \times (L_c+L_e-1)M$
- $\mathbf{h}_l : 1 \times L_e N$
- $\mathbf{C} : L_e N \times (L_c+L_e-1)M$
- $\boldsymbol{\Lambda}_l : (L_c+L_e-1)M \times (L_c+L_e-1)M$

Note that by the convolution $\mathbf{c}_{\text{eff},l}(n) = \mathbf{h}_l(n) * \mathbf{c}(n)$, we have,

$$\mathbf{c}_{\text{eff},l} = \mathbf{h}_l \mathbf{C} \tag{5}$$

Using (5) and the definition of $J_{\text{short},l}$ from (3), we get,

$$J_{\text{short},l} = \frac{\mathbf{c}_{\text{eff},l}\boldsymbol{\Lambda}_l\mathbf{c}^\dagger_{\text{eff},l}}{\mathbf{c}_{\text{eff},l}\mathbf{c}^\dagger_{\text{eff},l}} = \frac{\mathbf{h}_l\mathbf{C}\boldsymbol{\Lambda}_l\mathbf{C}^\dagger\mathbf{h}_l^\dagger}{\mathbf{h}_l\mathbf{C}\mathbf{C}^\dagger\mathbf{h}_l^\dagger} \tag{6}$$

Assuming that $L_e N \leq (L_c+L_e-1)M$ and that $\mathbf{C}$ has a full rank of $L_e N$, then the matrix $\mathbf{A} \triangleq \mathbf{C}\mathbf{C}^\dagger$ is positive definite. As such, there exists a nonsingular $L_e N \times L_e N$ matrix $\mathbf{G}$ such that [2],

$$\mathbf{A} = \mathbf{G}^\dagger \mathbf{G}$$

The choice of $\mathbf{G}$ is not unique here. One such $\mathbf{G}$ is upper triangular and is called the *Cholesky decomposition* [2] of $\mathbf{A}$. Using the above factorization, we can express $J_{\text{short},l}$ as a Rayleigh quotient [2]. Defining the $L_e N \times 1$ column vector $\mathbf{v}_l$ as follows,

$$\mathbf{v}_l \triangleq \mathbf{G}\mathbf{h}_l^\dagger \Longleftrightarrow \mathbf{h}_l = \mathbf{v}_l^\dagger\left(\mathbf{G}^{-1}\right)^\dagger\,,\ 0 \leq l \leq P-1 \tag{7}$$

we have from (6),

$$J_{\text{short},l} = \frac{\mathbf{v}_l^\dagger \left(\mathbf{G}^{-1}\right)^\dagger \mathbf{C}\mathbf{\Lambda}_l\mathbf{C}^\dagger \left(\mathbf{G}^{-1}\right)\mathbf{v}_l}{\mathbf{v}_l^\dagger \mathbf{v}_l} = \frac{\mathbf{v}_l^\dagger \mathbf{P}_l \mathbf{v}_l}{\mathbf{v}_l^\dagger \mathbf{v}_l} \quad (8)$$

where $\mathbf{P}_l \triangleq \left(\mathbf{G}^{-1}\right)^\dagger \mathbf{C}\mathbf{\Lambda}_l\mathbf{C}^\dagger \left(\mathbf{G}^{-1}\right)$. Evidently $\mathbf{P}_l$ is Hermitian and so we have expressed $J_{\text{short},l}$ as a Rayleigh quotient.

Now we proceed to analyze the noise objective function $J_{\text{noise},l}$. As $\mathbf{s}(n) = \mathbf{c}_{\text{eff}}(n) * \mathbf{x}(n)$ and $\mathbf{R}_{\mathbf{xx}}(k) = \sigma_x^2 \mathbf{I}\delta(k)$, we have,

$$\mathbf{R}_{\mathbf{ss}}(0) = \sigma_x^2 \sum_n \mathbf{c}_{\text{eff}}(n)\mathbf{c}_{\text{eff}}^\dagger(n)$$

Similarly, as $\mathbf{w}(n) = \mathbf{h}(n) * \boldsymbol{\eta}(n)$, we have,

$$\mathbf{R}_{\mathbf{ww}}(0) = \sum_{m,n} \mathbf{h}(m)\mathbf{R}_{\boldsymbol{\eta\eta}}(n-m)\mathbf{h}^\dagger(n)$$

From this, a straightforward calculation shows that we have,

$$[\mathbf{R}_{\mathbf{ss}}(0)]_{k,l} = \sigma_x^2 \sum_n \mathbf{c}_{\text{eff},k}(n)\mathbf{c}_{\text{eff},l}(n) \quad (9)$$

$$[\mathbf{R}_{\mathbf{ww}}(0)]_{k,l} = \sum_{m,n} \mathbf{h}_k(m)\mathbf{R}_{\boldsymbol{\eta\eta}}(n-m)\mathbf{h}_l^\dagger(n) \quad (10)$$

for $0 \leq k, l \leq P-1$. Using (9) and (7) yields the following.

$$\begin{aligned} \sigma_{s_l}^2 &= [\mathbf{R}_{\mathbf{ss}}(0)]_{l,l} = \sigma_x^2 \sum_n \mathbf{c}_{\text{eff},l}(n)\mathbf{c}_{\text{eff},l}(n) \\ &= \sigma_x^2 \mathbf{c}_{\text{eff},l}\mathbf{c}_{\text{eff},l}^\dagger = \sigma_x^2 \mathbf{h}_l \mathbf{C}\mathbf{C}^\dagger \mathbf{h}_l^\dagger = \sigma_x^2 \mathbf{v}_l^\dagger \mathbf{v}_l \quad (11) \end{aligned}$$

Defining the $L_e N \times L_e N$ matrix $\mathbf{R}_{\boldsymbol{\eta}}$ as follows,

$$\mathbf{R}_{\boldsymbol{\eta}} \triangleq \begin{bmatrix} \mathbf{R}_{\boldsymbol{\eta\eta}}(0) & \mathbf{R}_{\boldsymbol{\eta\eta}}(1) & \cdots & \mathbf{R}_{\boldsymbol{\eta\eta}}(L_e-1) \\ \mathbf{R}_{\boldsymbol{\eta\eta}}(-1) & \mathbf{R}_{\boldsymbol{\eta\eta}}(0) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{R}_{\boldsymbol{\eta\eta}}(1) \\ \mathbf{R}_{\boldsymbol{\eta\eta}}(-(L_e-1)) & \cdots & \mathbf{R}_{\boldsymbol{\eta\eta}}(-1) & \mathbf{R}_{\boldsymbol{\eta\eta}}(0) \end{bmatrix}$$

we can express $\sigma_{w_l}^2$ as follows using (10) and (7).

$$\begin{aligned} \sigma_{w_l}^2 &= [\mathbf{R}_{\mathbf{ww}}(0)]_{l,l} = \sum_{m,n} \mathbf{h}_l(m)\mathbf{R}_{\boldsymbol{\eta\eta}}(n-m)\mathbf{h}_l^\dagger(n) \\ &= \mathbf{h}_l \mathbf{R}_{\boldsymbol{\eta}} \mathbf{h}_l^\dagger = \mathbf{v}_l^\dagger \left(\mathbf{G}^{-1}\right)^\dagger \mathbf{R}_{\boldsymbol{\eta}} \left(\mathbf{G}^{-1}\right)\mathbf{v}_l \quad (12) \end{aligned}$$

Combining (11) and (12) with the definition of $J_{\text{noise},l}$ given in (4), we have the following.

$$J_{\text{noise},l} = \frac{\mathbf{v}_l^\dagger \left[\frac{1}{\sigma_x^2}\left(\mathbf{G}^{-1}\right)^\dagger \mathbf{R}_{\boldsymbol{\eta}} \left(\mathbf{G}^{-1}\right)\right]\mathbf{v}_l}{\mathbf{v}_l^\dagger \mathbf{v}_l} = \frac{\mathbf{v}_l^\dagger \mathbf{Q}_l \mathbf{v}_l}{\mathbf{v}_l^\dagger \mathbf{v}_l} \quad (13)$$

where $\mathbf{Q}_l \triangleq \frac{1}{\sigma_x^2}\left(\mathbf{G}^{-1}\right)^\dagger \mathbf{R}_{\boldsymbol{\eta}} \left(\mathbf{G}^{-1}\right)$. Since $\mathbf{R}_{\boldsymbol{\eta}}$ is Hermitian (as $\mathbf{R}_{\boldsymbol{\eta\eta}}^\dagger(-k) = \mathbf{R}_{\boldsymbol{\eta\eta}}(k)$), it follows that the matrix $\mathbf{Q}_l$ is as well. Thus, we have expressed $J_{\text{noise},l}$ as a Rayleigh quotient. Combining (8) and (13), we obtain from the definition of J_l in (2),

$$J_l = \frac{\mathbf{v}_l^\dagger \left[\alpha_l \mathbf{P}_l + (1-\alpha_l)\mathbf{Q}_l\right]\mathbf{v}_l}{\mathbf{v}_l^\dagger \mathbf{v}_l} = \frac{\mathbf{v}_l^\dagger \mathbf{T}_l \mathbf{v}_l}{\mathbf{v}_l^\dagger \mathbf{v}_l}$$

where $\mathbf{T}_l \triangleq \alpha_l \mathbf{P}_l + (1-\alpha_l)\mathbf{Q}_l$. Clearly $\mathbf{T}_l$ is Hermitian since $\mathbf{P}_l$ and $\mathbf{Q}_l$ are Hermitian and α_l is real. As $\mathbf{T}_l$ is Hermitian, we have thus expressed J_l as a Rayleigh quotient. By *Rayleigh's principle* [2], it follows that as $\mathbf{v}_l$ varies over all nonzero vectors, the minimum value of J_l is $\lambda_{l,\text{min}}$ where $\lambda_{l,\text{min}}$ denotes the smallest eigenvalue of $\mathbf{T}_l$. Furthermore, this minimum value is achieved if $\mathbf{v}_l = \mathbf{v}_{l,\text{min}}$, where $\mathbf{v}_{l,\text{min}}$ denotes an eigenvector of $\mathbf{T}_l$ corresponding to $\lambda_{l,\text{min}}$. (More generally, the minimum value of J_l is achieved iff $\mathbf{v}_l$ is in the eigenspace corresponding to $\lambda_{l,\text{min}}$. However, for sake of clarity, we will ignore this scenario.) Thus, if $J_{l,\text{opt}}$ and $\mathbf{h}_{l,\text{opt}}$ denote the optimum value of J_l and optimizing equalizer coefficients, respectively, then we have the following.

$$\boxed{\begin{aligned} J_{l,\text{opt}} &= \lambda_{l,\text{min}} \\ \mathbf{h}_{l,\text{opt}} &= \mathbf{v}_{l,\text{min}}^\dagger \left(\mathbf{G}^{-1}\right)^\dagger \end{aligned}}$$

The vector $\mathbf{h}_{l,\text{opt}}$ is referred to as an *eigenfilter* [8, 7] as its elements are filter coefficients derived from an eigenvector of a matrix. By computing $\mathbf{h}_{l,\text{opt}}$ for all l, we will have obtained the optimal filter coefficients of $\mathbf{H}(z)$ which minimize the objective function J in (1) for a fixed weighing sequence $\{p_l\}$.

4. EXPERIMENTAL RESULTS

To test our proposed channel shortening method, we have decided to compare it with the MMSE method of [1]. Here we used the following data regarding the test channel and input statistics.

- The channel $\mathbf{C}(z)$ is a 2×2 non-minimum phase system of length $L_c = 10$ (See [9] for the channel coefficients).
- The desired length of the effective channel is $L_d = 3$.
- The input signal $\mathbf{x}(n)$ is white with $\mathbf{R}_{\mathbf{xx}}(k) = \sigma_x^2 \mathbf{I}\delta(k)$.
- The input noise $\boldsymbol{\eta}(n)$ is also white with $\mathbf{R}_{\boldsymbol{\eta\eta}}(k) = \sigma_{\boldsymbol{\eta}}^2 \mathbf{I}\delta(k)$.
- The input signal-to-noise ratio $\frac{\sigma_x^2}{\sigma_{\boldsymbol{\eta}}^2}$ was chosen to be 10 dB.

Using our design method, we chose the following parameters.

- $P = M = 2$ (i.e. the number of equalizer outputs equals the number of inputs).
- $p_l = \frac{1}{P}$ for all l (i.e. all rows were weighed equally).
- $\alpha_l = \alpha$ for all l (i.e. the same tradeoff parameter was used for each row).
- For the penalty function, we chose $\beta = 0.5$ and $\gamma = 1$.

For the MMSE equalizer, the FIR filter $\mathbf{b}(n)$ was chosen using the orthonormality constraint (ONC) described in [1].

To compare these methods, we considered two figures of merit. The first was the energy compaction ratio, defined as follows.

$$\rho \triangleq \frac{\sum_{n \in \mathcal{W}} ||\mathbf{c}_{\text{eff}}(n)||_F^2}{\sum_n ||\mathbf{c}_{\text{eff}}(n)||_F^2}$$

Here, $||\mathbf{A}||_F$ denotes the *Frobenius norm* [2] of a matrix $\mathbf{A}$ given by $||\mathbf{A}||_F = \sqrt{\text{Tr}\left[\mathbf{A}^\dagger \mathbf{A}\right]}$. The energy compaction ratio is a measure of how much of the effective channel energy is contained within the desired window of interest. Clearly we have $0 \leq \rho \leq 1$ and for the purposes of channel shortening, a larger ρ is desired.

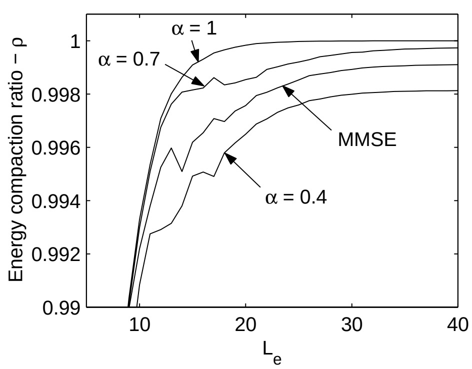

Fig. 2. Energy compaction as a function of equalizer length.

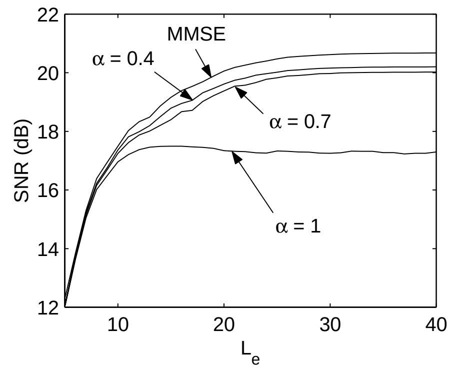

Fig. 3. Overall SNR as a function of equalizer length.

In addition to this figure of merit, we also considered the overall signal-to-noise ratio (SNR) defined as follows.

$$\text{SNR} \triangleq \frac{\text{Tr}\left[\mathbf{R}_{\mathbf{s}_{\text{des}}\mathbf{s}_{\text{des}}}(0)\right]}{\text{Tr}\left[\mathbf{R}_{\mathbf{s}_{\text{res}}\mathbf{s}_{\text{res}}}(0)\right] + \text{Tr}\left[\mathbf{R}_{\mathbf{ww}}(0)\right]} \qquad (14)$$

Here, $\mathbf{s}_{\text{des}}(n)$ is the *desired* signal component given by $\mathbf{s}_{\text{des}}(n) = \mathbf{c}_{\text{des}}(n) * \mathbf{x}(n)$, where we have,

$$\mathbf{c}_{\text{des}}(n) = \begin{cases} \mathbf{c}_{\text{eff}}(n)\,, & n \in \mathcal{W} \\ \mathbf{0}\,, & \text{otherwise} \end{cases}$$

Similarly, $\mathbf{s}_{\text{res}}(n)$ is the *residual* signal component given by $\mathbf{s}_{\text{res}}(n) = \mathbf{c}_{\text{res}}(n) * \mathbf{x}(n)$, where we have,

$$\mathbf{c}_{\text{res}}(n) = \mathbf{c}_{\text{eff}}(n) - \mathbf{c}_{\text{des}}(n) = \begin{cases} \mathbf{0}\,, & n \in \mathcal{W} \\ \mathbf{c}_{\text{eff}}(n)\,, & \text{otherwise} \end{cases}$$

The SNR of (14) is the ratio of the desired signal power to the sum of the residual intersymbol interference (ISI) and noise power.

Plots of the energy compaction ratio and overall SNR as a function of the length of the equalizer L_e are shown in Figures 2 and 3, respectively. For both our method and the MMSE method, the value of Δ was varied over all admissible values and chosen to optimize the respective objective function for each method.

From Figure 2, we can see that for most of the tradeoff parameters chosen (namely $\alpha = 0.7, 1$), our method performed better than the MMSE method in terms of energy compaction. This improvement however came at the price of a lower overall SNR as can be seen in Figure 3. Depending upon the application, this discrepancy in overall SNR may not be too severe, especially since the discrepancy itself is rather small (less than 1 dB for $\alpha = 0.4, 0.7$).

Typically, TEQs in traditional DMT systems are compared on the basis of achievable bit rate [5]. Though the DMMT has not yet been implemented in a practical setting, as more is known about it, most likely MIMO channel shortening equalizers will likewise be compared on the basis of achievable bit rate. The versatility that our method possesses in contrast to the MMSE method may then prove to be useful for this application. Only when these bit rates are calculated in a practical setting will our comparison of both methods be truly complete.

5. CONCLUDING REMARKS

We have shown that our channel shortening equalizer design method can outperform the MMSE method if our objective is to compact the energy of the overall channel. In addition, our method admits a greater degree of versatility than the MMSE method as a result of the tradeoff and weighing parameters present. These advantages may prove to be useful if, for example, our goal is to maximize achievable bit rate, as is the case for TEQs in DMT systems [5].

6. REFERENCES

[1] N. Al-Dhahir, "FIR channel-shortening equalizers for MIMO ISI channels," *IEEE Trans. on Comm.*, 49(2):213-218, Feb. 2001.

[2] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge Univ. Press, Cambridge, U.K., 1985.

[3] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. on Comm.*, 46(3):357-366, Mar. 1998.

[4] R. Schur and J. Speidel, "An efficient equalization method to minimize delay spread in OFDM/DMT systems," in *Proc. IEEE ICC '01*, Helsinki, Vol. 1, pp. 1-5, Jun. 2001.

[5] T. Starr, J. M. Cioffi, and P. J. Silverman, *Understanding Digital Subscriber Line Technology*, Prentice-Hall, Inc., Upper Saddle River, NJ, 1999.

[6] C. W. Therrien, *Discrete Random Signals and Statistical Signal Processing*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1992.

[7] A. Tkacenko and P. P. Vaidyanathan, "Noise optimized eigenfilter design of time-domain equalizers for DMT systems," submitted to *Proc. IEEE ICC '02*, New York, Apr.-May 2002.

[8] P. P. Vaidyanathan and T. Q. Nguyen, "Eigenfilters: a new approach to least-squares FIR filter design and applications including Nyquist filters," *IEEE Trans. on Circuits and Systems*, 34(1):11-23, Jan. 1987.

[9] `http://www.systems.caltech.edu/dsp/students/andre/papers/icassp2002.html`