LEAST-SQUARES CHANNEL EQUALIZATION PERFORMANCE VERSUS EQUALIZATION DELAY

A.P. LIAVAS

6-99

Preprint no. 6-99/1999

Department of Computer Science
University of Ioannina
451 10 Ioannina, Greece

Least-squares channel equalization performance versus equalization delay

Athanasios P. Liavas

4 March 1999

Abstract

Linear channel equalization has been a successful way for combating intersymbol interference, introduced by physical communication channels, at high enough symbol rates. In this paper, we consider the performance of least-squares equalizers in the single-input/multi-output channel context, when the true channel is composed of an m–th order significant part and long tails of "small" leading and/or trailing terms. Using a perturbation analysis approach, we show that if the diversity of the significant part is sufficiently large, with respect to the size of the tails, then the l–th order least-squares equalizers, with $l \geq m-1$, perform well, for all the delays corresponding to the significant part. On the other hand, we do not have any *a priori* knowledge for the performance of the equalizers, for the delays corresponding to the tails. They may, and usually do, perform poorly. Simulations agree with our theoretical results.

The author is with the Department of Computer Science, University of Ioannina, 45110 Ioannina, Greece.

E-mail: `liavas@cs.uoi.gr`.

1 Introduction

Signals transmitted through physical communication channels are usually distorted by intersymbol interference (ISI) and additive noise [1]. One classical way for combating channel distortions, like ISI, is linear channel equalization. Its target is the computation of an optimum linear filter, called *equalizer*, whose output approximates a (possibly) delayed version of the input. If the quality of the approximation is sufficiently good, then we can recover the input sequence.

It is well known, that in the single-input/single-output (SISO) channel setting, in order to equalize perfectly a mixed-phase finite-order noiseless channel, we need a doubly infinite linear equalizer [2].

On the other hand, in the single-input/multi-output (SIMO) channel setting, derived either by oversampling the channel or by using an array of sensors at the receiver, if the subchannels do not share common zeros, then an L–th order multichannel equalizer can equalize perfectly an M–th order noiseless SIMO channel, with $L \geq M - 1$ [3].

A case commonly encountered in practice is when the M–th order true subchannels possess a significant part of order m, with $m \ll M$, and long tails of "small" leading and/or trailing impulse response terms [4]. Implementation cost considerations force us to investigate which is the smallest possible order that an equalizer should have, in this case, in order to offer acceptable performance. To our knowledge, there does *not* exist a theoretical answer to this question. Furthermore, especially in the SIMO channel context, *no* theoretical explanation has been given to the fact that equalization performance for some delays appears inherently poor, while for some others it is usually satisfactory [5].

In this paper, we consider the least-squares (LS) equalization of SIMO channels, in the cases in which the M–th order true subchannels possess a significant part of order m and long tails of "small" leading and/or trailing terms. Using a perturbation analysis approach, we show that:

- if the diversity of the significant part is sufficiently large, with respect to the size of the tails, then the l–th order LS equalizers, with $l \geq m - 1$, attempting to equalize the M–

th order true channel, offer good performance, for all the delays corresponding to the significant part;

- we do *not* have any *a priori* knowledge for the performance of the LS equalizers, for the delays corresponding to the tails. They may, and usually do, perform poorly.

In Section 2, we present the framework of the LS SIMO channel equalization. In Section 3, we study the performance of LS equalizers for the various delays; simulations are presented in Section 4. Finally, conclusions are drawn in Section 5.

2 LS SIMO channel equalization

In this section, we consider the LS SIMO channel equalization problem. In Fig. 1, we present the single-input/two-output channel setting, resulting either by oversampling the channel, by a factor of 2, or by using 2 sensors at the receiver. Although we present our results in the single-input/two-output channel setting, the extension to the single-input/p-output setting, with $p > 2$, is trivial.

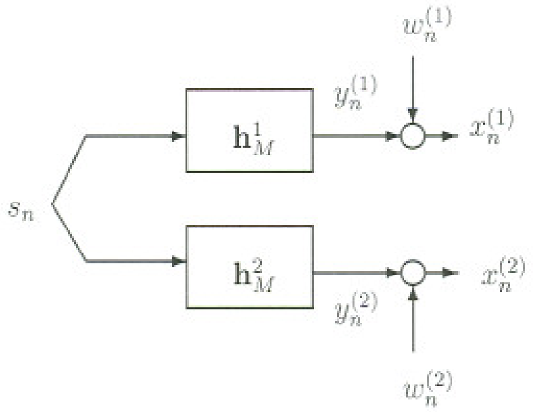

Fig.1 Single-input/two-output channel setting.

If the true channel order is M, then the output of the j–th subchannel, $x_n^{(j)}$, for $j = 1, 2$, is given by

$$x_n^{(j)} = \sum_{k=0}^{M} h_k^{(j)} s_{n-k} + w_n^{(j)},$$

where $\{h_k^{(j)}\}_{k=0}^{M}$ is the impulse response of the j–th subchannel, $\{s_n\}$ is the input sequence and $\{w_n^{(j)}\}$ is the additive subchannel noise. We denote the impulse response vector of the j–th subchannel, for $j = 1, 2$, as $\mathbf{h}_M^j \triangleq [\, h_0^{(j)} \; \cdots \; h_M^{(j)} \,]^T$, where superscript ^{T} denotes transpose, and the entire channel parameter vector as $\mathbf{h}_M \triangleq \begin{bmatrix} \mathbf{h}_M^1 \\ \mathbf{h}_M^2 \end{bmatrix}$. By stacking the $(L+1)$ most recent samples of each subchannel, we construct the data vector

$$\mathbf{x}_L(n) \triangleq [\, x_n^{(1)} \cdots x_{n-L}^{(1)} \, x_n^{(2)} \cdots x_{n-L}^{(2)} \,]^T,$$

which can be expressed as

$$\mathbf{x}_L(n) = \mathbf{y}_L(n) + \mathbf{w}_L(n) = \mathcal{H}_L(\mathbf{h}_M)\; \mathbf{s}_{L+M}(n) + \mathbf{w}_L(n),$$

where

$$\begin{aligned} \mathbf{y}_L(n) &\triangleq [\, y_n^{(1)} \cdots y_{n-L}^{(1)} \, y_n^{(2)} \cdots y_{n-L}^{(2)} \,]^T, \\ \mathbf{w}_L(n) &\triangleq [\, w_n^{(1)} \cdots w_{n-L}^{(1)} \, w_n^{(2)} \cdots w_{n-L}^{(2)} \,]^T, \\ \mathbf{s}_{L+M}(n) &\triangleq [\, s_n \; \cdots \; s_{n-L-M} \,]^T. \end{aligned}$$

The $2\,(L+1) \times (L+M+1)$ filtering matrix $\mathcal{H}_L(\mathbf{h}_M)$ is defined as

$$\mathcal{H}_L(\mathbf{h}_M) \triangleq \begin{bmatrix} \mathcal{F}_L(\mathbf{h}_M^1) \\ \mathcal{F}_L(\mathbf{h}_M^2) \end{bmatrix},$$

with the $(L+1) \times (L+M+1)$ matrix $\mathcal{F}_L(\mathbf{h}_M^i)$ given by

$$\mathcal{F}_L(\mathbf{h}_M^i) \triangleq \begin{bmatrix} h_0^{(i)} & \cdots & \cdots & h_M^{(i)} & & \\ & \ddots & & & \ddots & \\ & & h_0^{(i)} & \cdots & \cdots & h_M^{(i)} \end{bmatrix}.$$

It is well established that if $L \geq M-1$ and subchannels $\mathbf{h}_M^1$ and $\mathbf{h}_M^2$ do not share common zeros, then $\mathcal{H}_L(\mathbf{h}_M)$ is of full-column rank, i.e.,

$$\operatorname{rank}\left(\mathcal{H}_L(\mathbf{h}_M)\right) = L + M + 1.$$

This means that $\mathcal{H}_L^T(\mathbf{h}_M)$ is of full-row rank, yielding that the canonical vectors $\mathbf{e}_d$, i.e., the vectors with 1 at the d–th position and zeros elsewhere, for $d = 1, \ldots, L+M+1$, belong to the range space of $\mathcal{H}_L^T(\mathbf{h}_M)$. As a consequence, in the absence of noise, the multichannel equalizer defined by

$$\mathbf{g}_{L,d} \triangleq \begin{bmatrix} \mathbf{g}_{L,d}^1 \\ \mathbf{g}_{L,d}^2 \end{bmatrix} = \left(\mathcal{H}_L^T(\mathbf{h}_M)\right)^\sharp \mathbf{e}_d,$$

where superscript $^\sharp$ denotes the Moore-Penrose generalized inverse, equalizes perfectly channel $\mathbf{h}_M$, for delay $(d-1)$. This happens because filtering $\mathbf{x}_L(n)$ through $\mathbf{g}_{L,d}$ gives

$$\mathbf{x}_L^T(n)\,\mathbf{g}_{L,d} = \mathbf{s}_{L+M}^T(n)\,\mathcal{H}_L^T(\mathbf{h}_M)\left(\mathcal{H}_L^T(\mathbf{h}_M)\right)^\sharp \mathbf{e}_d = s_{n-d+1}.$$

In the sequel, we assume that we know *a priori* that the subchannels of $\mathbf{h}_M$ are composed of a significant part of order m, lying between indices m_1 and m_2, i.e., $m = m_2 - m_1$, and long tails occupying the rest of the indices. We partition $\mathbf{h}_M$ as:

$$\mathbf{h}_M = \mathbf{h}_{m_1,m_2}^z + \mathbf{d}_{m_1,m_2}^z,$$

where superscript ^{z} stands for "appropriately zero-padded" and

$$\mathbf{h}_{m_1,m_2}^z \triangleq \begin{bmatrix} \mathbf{h}_{m_1,m_2}^{z\,1} \\ \mathbf{h}_{m_1,m_2}^{z\,2} \end{bmatrix}, \quad \mathbf{d}_{m_1,m_2}^z \triangleq \begin{bmatrix} \mathbf{d}_{m_1,m_2}^{z\,1} \\ \mathbf{d}_{m_1,m_2}^{z\,2} \end{bmatrix},$$

with

$$\mathbf{h}_{m_1,m_2}^{z\,j} \triangleq [\,\underbrace{0\ \cdots\ 0}_{m_1}\ \underbrace{h_{m_1}^{(j)}\ \cdots\ h_{m_2}^{(j)}}_{m+1}\ \underbrace{0\ \cdots\ 0}_{M-m_2}\,]^T, \quad j=1,2,$$

$$\mathbf{d}_{m_1,m_2}^{z\,j} \triangleq [\,\underbrace{h_0^{(j)}\ \cdots\ h_{m_1-1}^{(j)}}_{m_1}\ \underbrace{0\ \cdots\ 0}_{m+1}\ \underbrace{h_{m_2+1}^{(j)}\ \cdots\ h_M^{(j)}}_{M-m_2}\,]^T, \quad j=1,2,$$

and

$$\|\,\mathbf{h}_M\,\|_2 = 1, \quad \left\|\,\mathbf{d}_{m_1,m_2}^z\,\right\|_2 = \epsilon_m \ll 1. \qquad (1)$$

With $\mathbf{h}_{m_1,m_2}$ we denote the corresponding truncated vectors, i.e.,

$$\mathbf{h}_{m_1,m_2} \triangleq \begin{bmatrix} \mathbf{h}_{m_1,m_2}^1 \\ \mathbf{h}_{m_1,m_2}^2 \end{bmatrix}, \quad \mathbf{h}_{m_1,m_2}^j \triangleq [\,h_{m_1}^{(j)}\ \cdots\ h_{m_2}^{(j)}\,]^T, \quad j=1,2.$$

We are interested in the performance levels offered by l–th order equalizers, with $m-1 \leq l < M$. In general, we can *not* equalize perfectly $\mathbf{h}_M$, by using l–th order equalizers. The best we can do, in the LS sense, is to compute the l–th order LS equalizers, for $d = 1, \ldots, l+M+1$, given by

$$\mathbf{g}_{l,d} = \left(\mathbf{H}_l^T(\mathbf{h}_M)\right)^{\sharp} \mathbf{e}_d,$$

leading to

$$\mathbf{x}_l^T(n)\,\mathbf{g}_{l,d} = \mathbf{s}_{l+M}^T(n)\,\mathcal{H}_l^T(\mathbf{h}_M)\left(\mathcal{H}_l^T(\mathbf{h}_M)\right)^{\sharp} \mathbf{e}_d \approx s_{n-d+1}.$$

In Fig. 2, we plot a portion of the real part of the two subchannels, constructed by the oversampled, by a factor of two, complex-valued microwave radio impulse response *chan2.mat*, found at the address *http://spib.rice.edu/spib/microwave.html*. The partitioning into the significant part and the tails is clear.

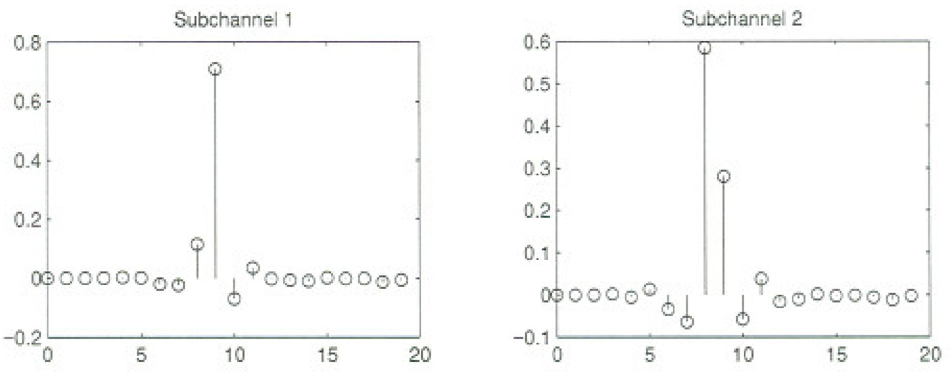

Fig. 2. Portion of the real part of the subchannels of *chan2.mat*.

In Fig. 3, we plot the vector 2-norm of the residuals of the l-th order LS equalizers, for delays $\{0, \ldots, 19\}$, i.e.,

$$\|\mathbf{r}_{l,d}\|_2 = \left\| \mathbf{e}_d - \mathcal{H}_l^T(\mathbf{h}_M)\,\mathbf{g}_{l,d} \right\|_2, \quad \text{for } d = 1, \ldots, 20,$$

and $l = 3, 6$. We observe that for certain delays the LS performance is satisfactory, while for delays outside a specific range it is *not*. Also, we observe that the performance of the 6-th order

LS equalizer is satisfactory for more delays than that of the 3-rd order LS equalizer. In the sequel, we provide an explanation for this phenomenon.

Fig. 3. 2-norm of residuals of l-th order LS equalizers, for delays $\{0, \ldots, 19\}$, for *chan2.mat* ($l = 3, 6$).

3 LS equalization performance vs. delay

Our "real-world" problem is the assessment of the performance of the LS solution of the equation

$$\mathcal{H}_l^T(\mathbf{h}_M)\ \mathbf{g}_{l,d} = \mathbf{e}_d,$$

for $d = 1, \ldots, l + M + 1$. From the dimensions of the $(l + M + 1) \times 2\,(l + 1)$ matrix $\mathcal{H}_l^T(\mathbf{h}_M)$, we obtain that

$$\operatorname{rank}(\mathcal{H}_l^T(\mathbf{h}_M)) \leq \min\{l + M + 1, 2\,(l + 1)\},$$

which gives that, for $l \leq M - 1$,

$$\operatorname{rank}(\mathcal{H}_l^T(\mathbf{h}_M)) \leq 2\,(l + 1).$$

This gives that out of the set of the $(l + M + 1)$ different canonical vectors corresponding to the $(l + M + 1)$ different possible delays, at most $2\,(l + 1)$ may lie into or close the range space of $\mathcal{H}_l^T(\mathbf{h}_M)$. Thus, the biggest number of delays for which we may expect sufficiently good LS equalization, with an equalizer of order l, is $2\,(l + 1)$.

Towards developing a study of the performance of the l–th order LS equalizers, attempting to equalize $\mathbf{h}_M$, it proves convenient to decompose our "real-world" problem into an "ideal" part and a perturbation. In order to develop a successfull analysis, the "ideal" problem and the perturbation should fulfill the following conditions:

1. The "ideal" problem should have a well-defined and informative solution.

2. The perturbation should be "small" with respect to the "ideal" quantities.

For the $(l+m+1)$ delays corresponding to the significant part, i.e., $d = m_1+1, \ldots, m_2+l+1$, we consider as "ideal" problem the equalization of the significant part of the channel, $\mathbf{h}^z_{m_1,m_2}$, by the l–th order LS equalizers, with $l \geq m-1$. We show that if the diversity of the significant part is sufficiently large with respect to the size of the tails, then the l–th order LS equalizers, for these delays, equalize sufficiently well our "real-world" channel, $\mathbf{h}_M$.

For the delays corresponding to the tails, we must consider in the "ideal" problem not only the significant part but also certain "small" terms; otherwise, the perturbation can *not* be "small" with respect to the "ideal" quantities. We show that, for these delays, we do *not* have any *a priori* knowledge about the performance of the corresponding l–th order LS equalizers, attempting to equalize $\mathbf{h}_M$. In practice, they perform poorly in the majority of the cases.

3.1 Delays corresponding to the significant part

In this subsection, we consider the performance of the l–th order LS equalizers, for the delays corresponding to the significant part of the channel. Our analysis is performed in three steps. The first two steps are hypothetical but they lead to the resolution of our problem in the third step.

In the first step, we assume that our channel is $\mathbf{h}_{m_1,m_2}$, i.e., the truncated significant part of the true channel; we recall that its order is m. If $\mathbf{h}^1_{m_1,m_2}$ and $\mathbf{h}^2_{m_1,m_2}$ do not share common zeros, then $\mathcal{H}_l(\mathbf{h}_{m_1,m_2})$ is of full-column rank, i.e.,

$$\operatorname{rank}\left(\mathcal{H}_l(\mathbf{h}_{m_1,m_2})\right) = l + m + 1.$$

Thus, $\mathcal{H}_l^T(\mathbf{h}_{m_1,m_2})$ is of full-row rank, giving that

$$\mathcal{R}\left(\mathcal{H}_l^T(\mathbf{h}_{m_1,m_2})\right) = \mathcal{R}\left(\mathbf{I}_{l+m+1}\right),$$

where $\mathcal{R}(\mathcal{A})$ denotes the range space of matrix $\mathcal{A}$ and $\mathbf{I}_n$ denotes the n-dimensional identity matrix. This means that the canonical vectors $\mathbf{e}_d$, for $d = 1, \ldots, l+m+1$, belong to the range space of $\mathcal{H}_l^T(\mathbf{h}_{m_1,m_2})$, i.e.,

$$\mathbf{e}_d \in \mathcal{R}\left(\mathcal{H}_l^T(\mathbf{h}_{m_1,m_2})\right).$$

Consequently, equation

$$\mathcal{H}_l^T(\mathbf{h}_{m_1,m_2})\ \mathbf{g}_{l,d} = \mathbf{e}_d$$

has always a solution, yielding that, in the absence of noise, channel $\mathbf{h}_{m_1,m_2}$ can be equalized perfectly by an l–th order equalizer, with $l \geq m-1$. The minimum norm solution is given by

$$\mathbf{g}_{l,d} = \left(\mathcal{H}_l^T(\mathbf{h}_{m_1,m_2})\right)^{\sharp}\ \mathbf{e}_d.$$

In the second step, we assume that our channel is $\mathbf{h}^z_{m_1,m_2}$, i.e., (the appropriately zero-padded version of) the significant part of the true channel. It is easy to see that

$$\text{rank}\left(\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})\right) = l+m+1 \quad \text{and} \quad \mathcal{R}\left(\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})\right) = \mathcal{R}\left(\begin{bmatrix} \mathbf{O}_{m_1,l+m+1} \\ \mathbf{I}_{l+m+1} \\ \mathbf{O}_{M-m_2,l+m+1} \end{bmatrix}\right),$$

where $\mathbf{O}_{n,m}$ denotes the $(n \times m)$ zero matrix. This means that $\mathcal{R}\left(\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})\right)$ contains the canonical vectors $\mathbf{e}_d$, for $d = m_1+1, \ldots, m_2+l+1$. For the corresponding delays, $\mathbf{h}^z_{m_1,m_2}$ can be equalized perfectly, in the noiseless case, by the minimum norm equalizers

$$\mathbf{g}_{l,d} = \left(\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})\right)^{\sharp}\ \mathbf{e}_d.$$

In the third step, we consider $\mathcal{H}_l^T(\mathbf{h}_M)$ as the result of the perturbation $\mathcal{H}_l^T(\mathbf{d}^z_{m_1,m_2})$ acting on $\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})$, and we study the performance of the l–th order LS equalizers, attempting to equalize $\mathbf{h}_M$. We denote the matrix 2-norm of the perturbation as

$$\mathcal{E}_l^{m_1,m_2} \triangleq \left\|\mathcal{H}_l^T(\mathbf{d}^z_{m_1,m_2})\right\|_2.$$

In order to relate $\mathcal{E}_l^{m_1,m_2}$ to the size of the tails, we use the structure of $\mathcal{H}_l^T(\mathbf{d}_{m_1,m_2}^z)$ and (1), to obtain

$$\left\|\mathcal{H}_l^T(\mathbf{d}_{m_1,m_2}^z)\right\|_F = \sqrt{l+1}\,\epsilon_m,$$

where $\|\cdot\|_F$ denotes the matrix Frobenious norm. Then, using the matrix 2-norm/Frobenious-norm inequalities [6, p. 57, 72], we obtain

$$\frac{1}{\sqrt{2}}\,\epsilon_m = \frac{1}{\sqrt{2\,(l+1)}}\left\|\mathcal{H}_l^T(\mathbf{d}_{m_1,m_2}^z)\right\|_F \leq \mathcal{E}_l^{m_1,m_2} \leq \left\|\mathcal{H}_l^T(\mathbf{d}_{m_1,m_2}^z)\right\|_F = \sqrt{l+1}\,\epsilon_m. \tag{2}$$

If $\mathcal{E}_l^{m_1,m_2} < \sigma_{l+m+1}\left(\mathcal{H}_l^T(\mathbf{h}_{m_1,m_2}^z)\right)$, where $\sigma_i(\mathcal{A})$ denotes the i–th singular value of matrix $\mathcal{A}$, then $\operatorname{rank}\left(\mathcal{H}_l^T(\mathbf{h}_M)\right) \geq l+m+1$. That is, perturbation $\mathcal{H}_l^T(\mathbf{d}_{m_1,m_2}^z)$ can increase the rank of our filtering matrix, but it can *not* decrease it. In this case, we denote by $\mathcal{R}\left(\mathcal{H}_l^T(\mathbf{h}_M)\right)_{l+m+1}$ the $(l+m+1)$–st dimensional subspace of $\mathcal{R}\left(\mathcal{H}_l^T(\mathbf{h}_M)\right)$, spanned by the left singular vectors associated with the $(l+m+1)$ largest singular values of $\mathcal{H}_l^T(\mathbf{h}_M)$. Thus, $\mathcal{R}\left(\mathcal{H}_l^T(\mathbf{h}_M)\right)_{l+m+1}$ may be considered as the perturbed subspace corresponding to $\mathcal{R}\left(\mathcal{H}_l^T(\mathbf{h}_{m_1,m_2}^z)\right)$.

The l–th order LS equalizer provides a combined channel–equalizer impulse response, $\widehat{\mathbf{e}}_{l,d}$, with

$$\widehat{\mathbf{e}}_{l,d} \in \mathcal{R}\left(\mathcal{H}_l^T(\mathbf{h}_M)\right),$$

that is closest, with respect to the vector 2-norm, to $\mathbf{e}_d$. In the sequel, we give the conditions under which even if we constrain our search to $\mathcal{R}\left(\mathcal{H}_l^T(\mathbf{h}_M)\right)_{l+m+1}$, that is, a subspace of $\mathcal{R}\left(\mathcal{H}_l^T(\mathbf{h}_M)\right)$, we can find a vector $\widetilde{\mathbf{e}}_{l,d}$ that is "close" to $\mathbf{e}_d$. Since the LS solution, $\widehat{\mathbf{e}}_{l,d}$, can only do better than $\widetilde{\mathbf{e}}_{l,d}$, the fact that $\widetilde{\mathbf{e}}_{l,d}$ is "close" to $\mathbf{e}_d$ means that the l–th order LS equalizers, attempting to equalize $\mathbf{h}_M$, for the delays corresponding to the significant part, perform "well".

In order to proceed, we need a measure of the distance between two linear subspaces $\mathcal{X}$ and $\mathcal{Y}$. Such a measure, commonly encountered in numerical analysis, is the sine of their *canonical angles*, denoted as $\|\sin\angle(\mathcal{X},\mathcal{Y})\|_2$. It is well known that [7, p. 92]

$$\rho_{g,2}(\mathcal{X},\mathcal{Y}) = \|\sin\angle(\mathcal{X},\mathcal{Y})\|_2\,, \tag{3}$$

where $\rho_{g,2}$ is the 2-*gap* between $\mathcal{X}$ and $\mathcal{Y}$, defined as [7, p. 91]

$$\rho_{g,2}(\mathcal{X},\mathcal{Y}) \triangleq \max\left\{ \max_{\substack{x\in\mathcal{X}\\ \|x\|_2=1}} \delta_2(x,\mathcal{Y}),\ \max_{\substack{y\in\mathcal{Y}\\ \|y\|_2=1}} \delta_2(y,\mathcal{X}) \right\}, \tag{4}$$

with

$$\delta_2(x,\mathcal{Y}) \triangleq \min_{y\in\mathcal{Y}} \| x - y \|_2. \tag{5}$$

The theorem that follows provides an upper bound for the distance between $\mathcal{R}\left(\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})\right)$ and $\mathcal{R}\left(\mathcal{H}_l^T(\mathbf{h}_M)\right)_{l+m+1}$.

Theorem 1: *Let $\mathcal{R}\left(\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})\right)$ denote the $(l+m+1)$–st dimensional range space of $\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})$, $\sigma_{l+m+1}\left(\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})\right)$ denote the smallest nonzero singular value of $\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})$ and $\mathcal{R}\left(\mathcal{H}_l^T(\mathbf{h}_M)\right)_{l+m+1}$ denote the $(l+m+1)$–st dimensional subspace spanned by the left singular vectors of $\mathcal{H}_l^T(\mathbf{h}_M)$ associated with its $(l+m+1)$ largest singular values. Let $\mathcal{E}_l^{m_1,m_2}$ be the matrix 2-norm of the perturbation $\mathcal{H}_l^T(\mathbf{d}^z_{m_1,m_2})$. If $\mathcal{E}_l^{m_1,m_2} \le \frac{\sigma_{l+m+1}\left(\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})\right)}{2}$, then*

$$\left\| \sin\angle\left(\mathcal{R}\left(\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})\right), \mathcal{R}\left(\mathcal{H}_l^T(\mathbf{h}_M)\right)_{l+m+1}\right)\right\|_2 \le \frac{\mathcal{E}_l^{m_1,m_2}}{\sigma_{l+m+1}\left(\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})\right) - \mathcal{E}_l^{m_1,m_2}} \tag{6}$$

Otherwise, the upper bound is equal to 1.

Proof: The theorem can be proved easily by using the "*generalized* $\sin\theta$ *theorem*" of [8]. ■

From (6), (3), (4), (5) and the fact that $\mathbf{e}_d \in \mathcal{R}\left(\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})\right)$, for $d = m_1+1,\ldots,m_2+l+1$, we deduce that there is an $\widetilde{\mathbf{e}}_{l,d} \in \mathcal{R}\left(\mathcal{H}_l^T(\mathbf{h}_M)\right)_{l+m+1}$ such that

$$\|\mathbf{e}_d - \widetilde{\mathbf{e}}_{l,d}\|_2 \le \frac{\mathcal{E}_l^{m_1,m_2}}{\sigma_{l+m+1}\left(\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})\right) - \mathcal{E}_l^{m_1,m_2}}.$$

Using the fact that the LS solution, $\widehat{\mathbf{e}}_{l,d}$, can only do better than $\widetilde{\mathbf{e}}_{l,d}$, we obtain

$$\|\mathbf{e}_d - \widehat{\mathbf{e}}_{l,d}\|_2 \le \|\mathbf{e}_d - \widetilde{\mathbf{e}}_{l,d}\|_2 \le \frac{\mathcal{E}_l^{m_1,m_2}}{\sigma_{l+m+1}\left(\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})\right) - \mathcal{E}_l^{m_1,m_2}}. \tag{7}$$

Bound (7) is a worst-case quantity. It means that if $\sigma_{l+m+1}\left(\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})\right)$ is sufficiently large with respect to $\mathcal{E}_l^{m_1,m_2}$, then the l–th order LS equalizers, attempting to equalize $\mathbf{h}_M$, perform well, for *all* the delays corresponding to the significant part. Of course, assessment of the best-case performance remains a very interesting problem, especially in the cases in which $\sigma_{l+m+1}\left(\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})\right)$ is "small".

Term $\sigma_{l+m+1}\left(\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})\right)$, being the distance, in the matrix 2-norm, of $\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})$ from the matrices with rank $(l+m)$, measures "how well" is fulfilled our assumption about $\text{rank}\left(\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})\right)$ or, equivalently, $\text{rank}\left(\mathcal{H}_l^T(\mathbf{h}_{m_1,m_2})\right)$. Analogous quantities have appeared in studies of the robustness of blind channel approximation methods, with respect to effective channel undermodeling/overmodeling [9], [10]. These quantities measure the distance of certain filtering matrices from the matrices with rank one less than the assumed rank. Thus, they may be interpreted as measures of *diversity* of the significant part of the channel. For varying l, $\sigma_{l+m+1}\left(\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})\right)$ are not orderable, in general; extensive simulations have shown that they are reasonably close each other.

Using (2), if $\epsilon_m \leq \dfrac{\sigma_{l+m+1}\left(\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})\right)}{2\sqrt{l+1}}$, we can derive a simpler but looser bound as

$$\|\, \mathbf{e}_d - \widehat{\mathbf{e}}_{l,d} \,\|_2 \leq \frac{2\,\mathcal{E}_l^{m_1,m_2}}{\sigma_{l+m+1}\left(\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})\right)} \leq \frac{2\sqrt{l+1}\,\epsilon_m}{\sigma_{l+m+1}\left(\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})\right)}. \tag{8}$$

Taking into account that, for varying l, quantities $\sigma_{l+m+1}\left(\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})\right)$ are *not* orderable, while the numerator at the right-hand side of inequality (8) increases with increasing l, our analysis favors the use of equalizers with the smallest possible order, i.e., $l = m-1$. This is to be expected, due to the "worst-case" character of our approach.

Of course, one may consider equalizers with order smaller than $(m-1)$. However, in this case, our "ideal" and "real-world" problems can *not* be related through a "small," i.e., $O(\epsilon_m)$, perturbation. Our analysis shows that, in this case, we should *not* expect sufficiently good performance, in general.

A seemingly annoying aspect of our results is that they have been derived by assuming the knowledge of the true impulse response, $\mathbf{h}_M$. However, since during our analysis we used only the size and not the structure of the perturbation, $\mathcal{H}_l^T(\mathbf{d}^z_{m_1,m_2})$, our results hold also for the cases in which the impulse response is known to within an $O(\epsilon_m)$ estimation error. Derivation of results by exploiting the structure of the perturbation remains a very interesting problem.

Recapitulating, we may say that we obtained, in a indirect way, an explanation for the observed satisfactory performance of LS equalizers, for the delays corresponding to the significant part of the channel. A direct approach, using pseudoinverse perturbation results, might appear complicated, due to the fact that perturbations on $\mathcal{H}_l^T(\mathbf{h}^z_{m_1,m_2})$ are *not* necessarily *acute* [7, pp. 139, 140].

3.2 Delays corresponding to the tails

In the previous subsection, we saw that, under certain conditions, the $(l+m+1)$ delays corresponding to the significant part of the channel lead to sufficiently good equalization of $\mathbf{h}_M$, by the l–th order LS equalizers, for $l \geq m-1$. Also, we saw that the dimensions of $\mathcal{H}_l^T(\mathbf{h}_M)$ imply that at most $2\,(l+1)$ delays may lead to sufficiently good equalization performance. This means that for equalizer order $l = m-1$, we should *not* expect any other delays to lead to sufficiently good equalization performance. However, for $l \geq m$, it is not immediately clear, from our analysis until now, whether there exist other delays which may lead to sufficiently good performance, or not.

Thus, a natural question arises: *"Are there any other delays, which may lead generically to sufficiently good LS equalization?"*

In order to answer this question, we perform a perturbation analysis similar to that in the previous subsection. However, now, in our "ideal" problem we must consider not only the significant part of the channel but also certain "small" terms. Otherwise, the perturbation can *not* be small with respect to the "ideal" quantities. For example, in order to study the performance of the l–th order LS solution to

$$\mathcal{H}_l^T(\mathbf{h}_M)\,\mathbf{g}_{l,m_1^*} = \mathbf{e}_{m_1^*},$$

with $m_1^* < m_1$, we must include in our "ideal" problem terms $h^{(1)}_{m_1^*}$ and $h^{(2)}_{m_1^*}$. This happens because if we do *not* include these terms, it is *not* possible to have a nonzero term at the (m_1^*+1)–st position of the right-hand side of our "ideal" problem. As a result, our "real-world" problem, which has a 1 at the (m_1^*+1)–st position of the right-hand side, can *not* be a "small" perturbation of our "ideal" problem.

A possible "ideal" problem is

$$\mathcal{H}_l^T(\mathbf{h}_{m_1^*,m_2}^z)\,\mathbf{g}_{l,m_1^*} = \mathbf{e}_{m_1^*},$$

with $\mathbf{h}_{m_1^*,m_2}^z$ denoting the part of the true channel lying between indices m_1^* and m_2, appropriately zero-padded.

The first implication of this fact is that we must consider equalizers of order $l^* \geq m^* - 1$, with $m^* = m_2 - m_1^*$. That is, in this case, we need equalizers longer than those used in the "ideal" problems of the previous subsection; otherwise, we can *not* equalize perfectly, in general, $\mathbf{h}_{m_1^*,m_2}$ or, equivalently, $\mathbf{h}_{m_1^*,m_2}^z$. The corresponding perturbation is $\mathcal{H}_{l^*}^T(\mathbf{d}_{m_1^*,m_2}^z)$. Continuing similarly to the analysis of the previous subsection, we obtain that the key terms are $\mathcal{E}_{l^*}^{m_1^*,m_2}$, that is, the matrix 2-norm of $\mathcal{H}_{l^*}^T(\mathbf{d}_{m_1^*,m_2}^z)$, and $\sigma_{l^*+m^*+1}\left(\mathcal{H}_{l^*}^T(\mathbf{h}_{m_1^*,m_2}^z)\right)$, that is, the smallest nonzero singular value of $\mathcal{H}_{l^*}^T(\mathbf{h}_{m_1^*,m_2}^z)$. Defining $\epsilon_{m^*} \triangleq \left\| \mathbf{d}_{m_1^*,m_2}^z \right\|_2$, we obtain similarly to (2)

$$\frac{1}{\sqrt{2}}\,\epsilon_{m^*} \leq \mathcal{E}_{l^*}^{m_1^*,m_2} \leq \sqrt{l^*+1}\,\epsilon_{m^*}. \tag{9}$$

Furthermore, since the small leading and/or trailing terms are usually of the same order of magnitude [4], we usually have that $\epsilon_m \approx \epsilon_{m^*}$.

As we saw in the previous subsection, if term $\mathcal{E}_{l^*}^{m_1^*,m_2}$ is sufficiently small with respect to $\sigma_{l^*+m^*+1}\left(\mathcal{H}_{l^*}^T(\mathbf{h}_{m_1^*,m_2}^z)\right)$, then we expect $\mathbf{e}_{m_1^*}$ to be close to $\mathcal{R}\left(\mathcal{H}_{l^*}^T(\mathbf{h}_M)\right)_{l^*+m^*+1}$, implying that the l^*–th order LS equalizer, attempting to equalize $\mathbf{h}_M$, for delay m_1^*, performs well. In the sequel, we use a result of [9], showing that this does *not* happen for m_1^* corresponding to the tails.

Theorem 2: *If $\sigma_{l+m+1}\left(\mathcal{H}_l^T(\mathbf{h}_{m_1,m_2}^z)\right)$ denotes the smallest nonzero singular value of the rank-$(l+m+1)$ matrix $\mathcal{H}_l^T(\mathbf{h}_{m_1,m_2}^z)$, then*

$$\sigma_{l+m+1}\left(\mathcal{H}_l^T(\mathbf{h}_{m_1,m_2}^z)\right) \leq \min\left\{ \sqrt{\left|h_{m_1}^{(1)}\right|^2 + \left|h_{m_1}^{(2)}\right|^2},\ \sqrt{\left|h_{m_2}^{(1)}\right|^2 + \left|h_{m_2}^{(2)}\right|^2} \right\}. \tag{10}$$

In this case, (1) and the fact that $h^{(1)}_{m_1^*}$ and $h^{(2)}_{m_1^*}$ belong to the *true* channel tails, i.e., $\mathbf{d}^z_{m_1,m_2}$, result in

$$\sigma_{l^*+m^*+1}\left(\mathcal{H}^T_{l^*}(\mathbf{h}^z_{m_1^*,m_2})\right) \leq \epsilon_m. \tag{11}$$

Theorem 1, relations (9), (11) and the fact that $\epsilon_{m^*} \approx \epsilon_m$, yield that, usually, we do *not* have any *a priori* knowledge about the distance between $\mathcal{R}\left(\mathcal{H}^T_{l^*}(\mathbf{h}^z_{m_1^*,m_2})\right)$ and $\mathcal{R}\left(\mathcal{H}^T_{l^*}(\mathbf{h}_M)\right)_{l^*+m^*+1}$. In fact, in the majority of the cases, the only upper bound we can give for the sine of their canonical angles is 1, meaning that it is *not* guaranteed that there is a vector in $\mathcal{R}\left(\mathcal{H}^T_{l^*}(\mathbf{h}_M)\right)_{l^*+m^*+1}$ that is close to $\mathbf{e}_{m_1^*}$.

One may wander if $\mathbf{e}_{m_1^*}$ may be generically close to the subspace spanned by the left singular vectors of $\mathcal{H}^T_{l^*}(\mathbf{h}_M)$ corresponding to its remaining nonzero singular values. It turns out that this does *not* happen, for a counterexample can be easily constructed. It can be easily seen that we can null terms $h^{(1)}_{m_1^*}$ and $h^{(2)}_{m_1^*}$ of the m_{1^*}–th row of $\mathcal{H}^T_{l^*}(\mathbf{h}^z_{m_1^*,m_2})$, by adding a small, i.e., $O(\epsilon_m)$, perturbation matrix, composed of terms $-h^{(1)}_{m_1^*}$ and $-h^{(2)}_{m_1^*}$ at the appropriate positions of the m_{1^*}–th row, and zeros elsewhere. This small perturbation makes $\mathbf{e}_{m_1^*}$ orthogonal to the range space of the resulting perturbed matrix. Of course, this perturbation does *not* have the structure of $\mathcal{H}^T_{l^*}(\mathbf{d}^z_{m_1^*,m_2})$. However, it is very informative, in our framework, in which we repeat that we use only the size and not the structure of the perturbation, because it implies that for the delays corresponding to the tails, we can *not* derive a worst-case bound smaller than 1.

Analogous arguments hold for the $d > m_2 + l + 1$ case.

4 Simulations

In the previous section, we derived bounds for the performance of LS equalizers, for the various equalization delays. In this section, we perform simulations to check our theoretical results.

In Fig. 4a, we plot a 10-th order 2-channel impulse response, composed of a 2-nd order significant part, lying between positions $m_1 = 3$ and $m_2 = 5$, and tails. In Fig. 4b, we plot the 2-norm of the residuals of the 3-rd order LS equalizers, for the various delays, and bound (7) (thick line). We observe that our bound is able to predict the performance of the LS equalizers for the various delays. Also, we observe that for some specific delays corresponding to the tails, the 2-norm of the residual of the LS equalizers is "close" to 1, supporting the arguments of the previous subsection.

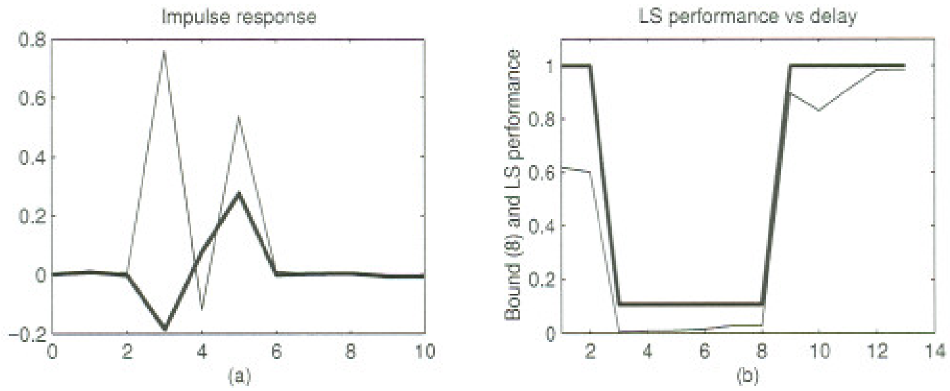

Fig. 4. (a) 10-th order impulse response (b) Bound (7) (thick line) and 2-norm of residuals of the 3-rd order LS equalizers versus delay.

5 Conclusions

We considered the relationship between LS equalization performance and equalization delay, in the cases in which the M–th order true subchannels possess an m–th order significant part, with $m \ll M$, and long tails of leading and/or trailing terms. Using a perturbation analysis approach, we showed that if the diversity of the significant part is sufficiently large with respect to the size of the tails, then the l–th order LS equalizers, with $l \geq m - 1$, perform well, for all the delays corresponding to the significant part. On the other hand, we do *not* have any *a priori* knowledge for the performance of the LS equalizers for the delays corresponding to the tails. They may, and usually do, perform poorly. Our results offer an explanation of the observed behavior of LS equalizers in "realistic" cases [5].

References

[1] J. G. Proakis. *Digital Communications.* 3rd ed. New York: McGraw-Hill, 1995.

[2] A. Benveniste, M. Goursat and G. Ruget, "Robust Identification of a Nonminimum Phase System: Blind Adjustment of a Linear Equalizer in Data Communications," *IEEE Trans. Automatic Control*, vol. AC-25, no. 3, pp. 385–399, June 1980.

[3] E. Moulines, P. Duhamel, J. F. Cardoso and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, pp. 516–525, February 1995.

[4] J. R. Treichler, I. Fijalkow and C. R. Johnson, Jr., "Fractionally spaced equalizers. How long should they really be?," *IEEE Signal Processing Magazine*, pp. 65–81, May 1996.

[5] C. R. Johnson, P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown and R. A. Casas, "Blind Equalization Using the Constant Modulus Criterion: A Review," *Proc. IEEE*, vol. 86, no. 10, pp. 1927–1950, October 1998.

[6] G. Golub and C. Van Loan. *Matrix Computations.* 2nd ed. The Johns Hopkins University Press, 1991.

[7] G. Stewart and J. Sun. *Matrix perturbation theory.* Academic Press, 1990.

[8] P. A. Wedin, "Perturbation bounds in connection with singular value decomposition," *BIT*, vol. 12, pp. 99–111, 1972.

[9] A. P. Liavas, P. A. Regalia and J. P. Delmas, "Robustness of least-squares and subspace methods for blind channel identification/equalization with respect to effective channel undermodeling/overmodeling," *IEEE Trans. Signal Processing*, June 1999 (to appear).

[10] A. P. Liavas, P. A. Regalia and J. P. Delmas, "Robustness of the linear prediction method for blind channel identification/equalization with respect to effective channel undermodeling/overmodeling," submitted to the *IEEE Trans. Signal Processing*, April 1998.