A LEAST-/ MEAN- SQUARES APPROACH TO CHANNEL IDENTIFICATION AND EQUALIZATION IN OFDM

Tareq Y. Al-Naffouri[1], Ghazi Al-Rawi[1], Ahmad Bahai[2], and Arogyaswami Paulraj[1]

[1]Electrical Engineering Department, Stanford University, CA 94305

[2]National Semiconductor, Santa Clara, CA 95052

ABSTRACT

This work proposes an iterative least-/mean- squares approach to channel identification and equalization in OFDM. This is achieved by exploiting the natural constraints imposed by the channel (sparsity and maximum delay spread) and those imposed by the transmitter (pilots, cyclic prefix, and the finite alphabet constraint). These constraints are used to reduce the number of pilots needed for channel and data recovery and also to perform this task within one packet. The diagonal nature of the OFDM channel makes it possible to perform optimal (nonlinear) mean-square detection of the data.

1. INTRODUCTION

There has been increasing interest in OFDM as it combines the advantages of high achievable rates and easy implementation. This is reflected by the many standards that considered and adopted OFDM including those for digital audio and video broadcasting, high speed modems over digital subscriber lines, and local area wireless broadband systems. [1]

OFDM divides the communication channel into independent subchannels by appending a cyclic prefix (CP) to the data block transmitted through the channel. This turns out to be a very convenient structure that lends itself to exploiting the various constraints imposed by the transmitter and the channel. As such, many techniques have been proposed in literature to estimate and equalize channels for OFDM transmission (see, e.g., [1], [2], [3], and the references therein). In this paper, we propose a method for semi-blind channel and data recovery. Specifically, we use the natural (sparsity) channel constraints and those imposed by the transmitter to perform channel and data recovery within the same packet and to reduce the number of packets that are eventually needed. These constraints include

- Maximum delay spread and sparsity constraints on the channel which mean that there are only a few active taps.
- Redundancy of the input in the form of cyclic prefix which splits the OFDM channel into two subchannels (redundancy due to the presence of a real code is considered in [4]).
- Finite alphabet constraints on the data.
- The artificial constraint of pilots whose number we are able to reduce by building upon the aforementioned less superficial constraints.

1.1. Notation

We denote scalars with small-case letters, vectors with small-case boldface letters, and matrices with uppercase boldface letters. Caligraphic notation (e.g. $\boldsymbol{\mathcal{X}}$) is reserved for vectors in the frequency domain. When these variables become a function of time, the time index i appears between parentheses for scalars (e.g. $x(i)$) and as a subscript for vectors (e.g. $\boldsymbol{h}_i$).

Now consider a length-N vector $\boldsymbol{x}_i$. We deal with three derivatives associated with this vector. The first two are obtained by partitioning $\boldsymbol{x}_i$ into an upper (prefix) vector $\underline{\boldsymbol{x}}_i$ and a lower (usually longer) vector $\tilde{\boldsymbol{x}}_i$. The third derivative, $\overline{\boldsymbol{x}}_i$, is created by concatenating $\boldsymbol{x}_i$ with a copy of its prefix $\underline{\boldsymbol{x}}_i$. The relation between $\boldsymbol{x}_i$ and its derivatives is summarized by the following relation

$$\overline{\boldsymbol{x}}_i = \begin{bmatrix} \underline{\boldsymbol{x}}_i \\ \boldsymbol{x}_i \end{bmatrix} = \begin{bmatrix} \underline{\boldsymbol{x}}_i \\ \tilde{\boldsymbol{x}}_i \\ \underline{\boldsymbol{x}}_i \end{bmatrix} \tag{1}$$

This notational convention will be extended to matrices as well. Thus, a matrix $\boldsymbol{Q}$ with N rows can be partitioned as

$$\boldsymbol{Q} = \begin{bmatrix} \underline{\boldsymbol{Q}} \\ \tilde{\boldsymbol{Q}} \end{bmatrix} = \begin{bmatrix} \underline{\boldsymbol{Q}}_L \\ \tilde{\boldsymbol{Q}}_{N-L} \end{bmatrix} = \begin{bmatrix} \underline{\boldsymbol{Q}}_{I_1} \\ \tilde{\boldsymbol{Q}}_{I_2} \end{bmatrix} \tag{2}$$

The subscripts stand for the indicator set of the rows in the partitioned matrices or for their number. They are omitted whenever they are understood.

1.2. A Note On (Regularized) Least-Squares (LS)

Since LS will be heavily used in this paper, it would be conducive to briefly introduce it here. We solve the system of equations $\boldsymbol{A}\underline{\boldsymbol{h}} = \boldsymbol{\mathcal{H}}$ in the LS sense by performing the minimization

$$\min_{\underline{\boldsymbol{h}}} \|\underline{\boldsymbol{h}} - \underline{\boldsymbol{h}}_{-1}\|^2_{\Pi_{-1}} + \|\boldsymbol{A}\underline{\boldsymbol{h}} - \boldsymbol{\mathcal{H}}\|^2_W \tag{3}$$

where $\underline{\boldsymbol{h}}_{-1}$ is the a priori information about $\underline{\boldsymbol{h}}$, and $\boldsymbol{\Pi}_{-1}$ ($\boldsymbol{W}$) is the confidence we have about $\underline{\boldsymbol{h}}_{-1}$ ($\boldsymbol{\mathcal{H}}$). The solution of (3) and the corresponding estimation error are given, respectively, by [5]

$$\underline{\boldsymbol{h}}_0 = \underline{\boldsymbol{h}}_{-1} + \left(\boldsymbol{\Pi}_{-1} + \boldsymbol{A}^*\boldsymbol{W}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^*\boldsymbol{W}\left(\boldsymbol{\mathcal{H}} - \boldsymbol{A}\underline{\boldsymbol{h}}_{-1}\right) \tag{4}$$

$$\boldsymbol{\Pi}_0^{-1} = \left(\boldsymbol{\Pi}_{-1} + \boldsymbol{A}^*\boldsymbol{W}\boldsymbol{A}\right)^{-1} \tag{5}$$

2. ESSENTIAL ELEMENTS OF OFDM

In OFDM, a data sequence $\{\mathcal{X}(i)\}$ is transmitted in packets $\boldsymbol{\mathcal{X}}_i$ of length N. Each packet undergoes an IDFT operation to produce the transform vector $\boldsymbol{x}_i$:

$$\boldsymbol{x}_i = \boldsymbol{Q}\boldsymbol{\mathcal{X}}_i \quad \text{where} \quad \boldsymbol{Q} = \left[e^{j\frac{2\pi}{N}lm}\right] \tag{6}$$

If the underlying sequence $\{x(i)\}$ is transmitted through a (length $L{+}1$ FIR) channel $\underline{\boldsymbol{h}}$, it will be subject to intersymbol interference ISI. To go around this, a guard band in the form of a CP is inserted between any consecutive packets, $\boldsymbol{x}_{i-1}$ and $\boldsymbol{x}_i$. Thus, instead of transmitting $\boldsymbol{x}_i$, we transmit the length $N+L$ packet $\overline{\boldsymbol{x}}_i$ as defined in (1). This induces the sequence $\{\overline{x}(i)\}$ which in turn produces the sequence $\{\overline{y}(i)\}$ at the channel output. Motivated by the packet structure of the input, it is also convenient to deal with the output in the form of packets of length $M = N+L$, and further split each packet into a length-N packet $\boldsymbol{y}_i$ and a prefix associated with it $\underline{\boldsymbol{y}}_i$, i.e.

$$\overline{\boldsymbol{y}}_i = \begin{bmatrix} \underline{\boldsymbol{y}}_i \\ \boldsymbol{y}_i \end{bmatrix} \tag{7}$$

This partition is natural because the prefix $\underline{\boldsymbol{y}}_i$ takes the burden of interference between $\overline{\boldsymbol{x}}_{i-1}$ and $\overline{\boldsymbol{x}}_i$, while the remaining part, $\boldsymbol{y}_i$, depends on the ith input packet $\boldsymbol{x}_i$ only. These facts and more can be seen from the relationship

$$\begin{bmatrix} \boldsymbol{y}_{i-1} \\ \underline{\boldsymbol{y}}_i \\ \boldsymbol{y}_i \end{bmatrix} = \begin{bmatrix} \boldsymbol{n}_{i-1} \\ \underline{\boldsymbol{n}}_i \\ \boldsymbol{n}_i \end{bmatrix} + \tag{8}$$

$$\left[\begin{array}{cc|cc} \multicolumn{2}{c}{\overline{\boldsymbol{H}}} & \boldsymbol{O}_{N\times L} & \boldsymbol{O}_{N\times N} \\ \multicolumn{2}{c}{\cdots} & \multicolumn{2}{c}{\cdots} \\ \boldsymbol{O}_{L\times N} & \underline{\boldsymbol{H}}^B & \underline{\boldsymbol{H}}^F & \boldsymbol{O}_{L\times N} \\ \multicolumn{2}{c}{\cdots} & \multicolumn{2}{c}{\cdots} \\ \boldsymbol{O}_{N\times N} & \boldsymbol{O}_{N\times L} & \multicolumn{2}{c}{\overline{\boldsymbol{H}}} \end{array}\right] \begin{bmatrix} \underline{\boldsymbol{x}}_{i-1} \\ \tilde{\boldsymbol{x}}_{i-1} \\ \underline{\boldsymbol{x}}_{i-1} \\ \underline{\boldsymbol{x}}_i \\ \tilde{\boldsymbol{x}}_i \\ \underline{\boldsymbol{x}}_i \end{bmatrix}$$

where $\boldsymbol{n}$ is the output noise which we take to be white Gaussian with variance σ_n^2. The matrices $\overline{\boldsymbol{H}}$, $\underline{\boldsymbol{H}}^B$, and $\underline{\boldsymbol{H}}^F$ are convolution (Toeplitz) matrices of proper sizes created from the vector $\underline{\boldsymbol{h}}$. Because of the redundancy in the input, the convolution in (8) can be decomposed into two distinct constituent convolution operations or subchannels.

2.1. Circular Convolution (Subchannel)

From (8), we can write

$$\boldsymbol{y}_i = \overline{\boldsymbol{H}} \begin{bmatrix} \underline{\boldsymbol{x}}_i \\ \tilde{\boldsymbol{x}}_i \\ \underline{\boldsymbol{x}}_i \end{bmatrix} = \overline{\boldsymbol{H}}\,\overline{\boldsymbol{x}}_i + \boldsymbol{n}_i \tag{9}$$

Thus, $\boldsymbol{y}_i$ is created solely from $\overline{\boldsymbol{x}}_i$ through convolution. Moreover, the existence of a cyclic prefix in $\overline{\boldsymbol{x}}_i$ renders this convolution cyclic:

$$\boxed{\boldsymbol{y}_i = \boldsymbol{h} \circledast \boldsymbol{x}_i + \boldsymbol{n}_i} \tag{10}$$

where $\boldsymbol{h}$ is a length-N zero-padded version of $\underline{\boldsymbol{h}}$

$$\boldsymbol{h} \triangleq \begin{bmatrix} \underline{\boldsymbol{h}} \\ \boldsymbol{O}_{(N-L-1)\times 1} \end{bmatrix} \tag{11}$$

In the frequency domain, the cyclic convolution (10) reduces to the element-by-element operation

$$\boxed{\boldsymbol{\mathcal{Y}}_i = \boldsymbol{\mathcal{H}} \odot \boldsymbol{\mathcal{X}}_i + \boldsymbol{\mathcal{N}}_i} \tag{12}$$

where $\boldsymbol{\mathcal{H}}$, $\boldsymbol{\mathcal{X}}_i$, and $\boldsymbol{\mathcal{Y}}_i$ are the DFT's of $\boldsymbol{h}$, $\boldsymbol{x}_i$, and $\boldsymbol{y}_i$, respectively. In other words, we have

$$\boldsymbol{h} = \boldsymbol{Q}\boldsymbol{\mathcal{H}}, \quad \boldsymbol{x}_i = \boldsymbol{Q}\boldsymbol{\mathcal{X}}_i, \quad \text{and} \quad \boldsymbol{y}_i = \boldsymbol{Q}\boldsymbol{\mathcal{Y}}_i \tag{13}$$

From (11) and (13), we can show, using the unitary nature of $\boldsymbol{Q}$, that

$$\underline{\boldsymbol{Q}}^*_{L+1}\underline{\boldsymbol{h}} = \boldsymbol{\mathcal{H}} \tag{14}$$

This in turn enables us to write (12) in the *time-frequency form*

$$\boxed{\boldsymbol{\mathcal{Y}}_i = \operatorname{diag}(\boldsymbol{\mathcal{X}}_i)\,\underline{\boldsymbol{Q}}^*_{L+1}\underline{\boldsymbol{h}} + \boldsymbol{\mathcal{N}}_i} \tag{15}$$

2.2. Mean-Square Estimation of Data

The diagonal nature of the cyclic subchannel makes it possible to perform MMSE equalization with low complexity. This has the added advantage of enabling one to perform truly optimal (as opposed to linear) MMSE. Thus, assume that the sample $\mathcal{X}_i(l)$, in the lth bin, takes on its value from the alphabet $A = \{A_1, A_2, \ldots, A_{|A|}\}$, with equal probability. We can show that the MMSE estimate of $\mathcal{X}_i(l)$, given $\mathcal{Y}_i(l)$, is

$$\hat{\mathcal{X}}_i^{\mathsf{MMSE}}(l) = E\left[\mathcal{X}_i(l)|\mathcal{Y}_i(l)\right] = \frac{\sum_{j=1}^{j=|A|} A_j e^{-\frac{|\mathcal{Y}_i(l)-\mathcal{H}(l)A_j|^2}{\sigma_n^2}}}{\sum_{j=1}^{|A|} e^{-\frac{|\mathcal{Y}_i(l)-\mathcal{H}(l)A_j|^2}{\sigma_n{}^2}}} \tag{16}$$

The estimate is subsequently thresholded to the nearest alphabet

$$\hat{\boldsymbol{\mathcal{X}}}_i = \lfloor \hat{\boldsymbol{\mathcal{X}}}^{\mathsf{MMSE}} \rceil \tag{17}$$

Since joint channel and data recovery is being sought, the channel estimate is usually uncertain, i.e. instead of (12), we should actually have

$$\mathcal{Y}_i(l) = (\mathcal{H}(l) + \delta\mathcal{H}(l))\,\mathcal{X}_i(l) + \mathcal{N}_i(l)$$

If we assume that the channel uncertainty is zero-mean Gaussian and independent of the data and noise, we can show that the optimal (nonlinear) MMSE estimate is

$$\hat{\mathcal{X}}_i^{\mathsf{MMSE}}(l) = \frac{\sum_{j=1}^{j=|A|} A_j e^{-\frac{|\mathcal{Y}_i(l)-\mathcal{H}(l)A_j|^2}{\sigma_n^2+|A_j|^2\sigma^2_{\mathcal{H}(l)}}}}{\sum_{j=1}^{j=|A|} e^{-\frac{|\mathcal{Y}_i(l)-\mathcal{H}(l)A_j|^2}{\sigma_n^2+|A_j|^2\sigma^2_{\mathcal{H}(l)}}}} \tag{18}$$

Fortunately, the variance $\sigma^2_{\mathcal{H}(l)}$ comes as a byproduct of the channel LS estimate. More specifically, given the error covariance matrix (5) of the estimate $\hat{\underline{\boldsymbol{h}}}$, we can use (14) to calculate the covariance matrix of $\hat{\boldsymbol{\mathcal{H}}}$

$$\mathrm{Cov}_{\mathcal{H}} \triangleq E\left[\boldsymbol{\delta\mathcal{H}\delta\mathcal{H}}^*\right] = \underline{\boldsymbol{Q}}^*_{L+1}\boldsymbol{\Pi}_0^{-1}\underline{\boldsymbol{Q}}_{L+1} \tag{19}$$

2.3. Linear Convolution (Subchannel)

Starting from (8), we can show that the underlying prefix sequences $\{\underline{x}(i)\}$ and $\{\underline{y}(i)\}$ are related together with the channel through linear convolution [6], i.e.

$$\boxed{\underline{y}(i) = \underline{h}(i) * \underline{x}(i) + \underline{n}(i)} \tag{20}$$

Alternatively, we can write

$$\underline{\boldsymbol{y}}_i = \underline{\boldsymbol{X}}_i \underline{\boldsymbol{h}} + \underline{\boldsymbol{n}}_i \tag{21}$$

where

$$\underline{\boldsymbol{X}}_i = \begin{bmatrix} \underline{x}_i(0) & \underline{x}_{i-1}(L-1) & \cdots & \underline{x}_{i-1}(0) \\ \underline{x}_i(1) & \underline{x}_i(0) & \cdots & \underline{x}_{i-1}(1) \\ \vdots & \vdots & \cdots & \underline{x}_{i-1}(L-2) \\ \underline{x}_i(L-1) & \underline{x}_i(L-2) & \cdots & \underline{x}_{i-1}(L-1) \end{bmatrix}$$

3. CHANNEL ESTIMATION

To perform channel estimation, we utilize the constraints imposed by natural (sparse) structure of the channel and those induced by the transmitter. In what follows, we introduce both of these classes of constraints and explain how they can be used for channel estimation.

3.1. Channel Induced Constraints

Maximum delay spread: This has already been utilized to transform the system of equations (12) in $\boldsymbol{\mathcal{H}}$ into the overdetermined system (15) in $\underline{\boldsymbol{h}}$. Although the channel frequency response $\boldsymbol{\mathcal{H}}$ has N elements (or frequency bins), these elements have at most $L+1$ degrees of freedom as reflected by the DFT relationship (14). Thus, estimating $\underline{\boldsymbol{h}}$ from (15) instead of estimating $\boldsymbol{\mathcal{H}}$ results in an improved channel estimate.

Sparsity: In addition to a maximum delay spread constraint, the channel is usually sparse consisting of a few active taps (indexed by a set I_s). Even if the channel is unknown, the location of these taps can usually be determined a priori. [1] We can thus simply identify those and set the rest to zero. In this case, (14) takes the alternative form

$$\underline{\boldsymbol{Q}}^*_{I_s} \underline{\boldsymbol{h}}_{I_s} = \boldsymbol{\mathcal{H}} \tag{22}$$

With this in mind, relations (15) and (21) can be rewritten and augmented as

$$\begin{bmatrix} \boldsymbol{\mathcal{Y}}_i \\ \underline{\boldsymbol{y}}_i \end{bmatrix} = \begin{bmatrix} \operatorname{diag}(\boldsymbol{\mathcal{X}}_i)\underline{\boldsymbol{Q}}^*_{I_s} \\ \underline{\boldsymbol{X}}^s_i \end{bmatrix} \underline{\boldsymbol{h}}_{I_s} + \begin{bmatrix} \boldsymbol{\mathcal{N}}_i \\ \underline{\boldsymbol{n}}_i, \end{bmatrix} \tag{23}$$

where $\underline{\boldsymbol{X}}^s_i$ is nothing but $\underline{\boldsymbol{X}}_i$ stripped of the columns corresponding to the zero elements of $\underline{\boldsymbol{h}}$. This system can be solved in the LS sense to obtain an estimate of $\underline{\boldsymbol{h}}_{I_s}$ (or $\underline{\boldsymbol{h}}$) (4) and the corresponding error covariance (5). In solving (3), we set $\boldsymbol{W} = \boldsymbol{R}_n^{-1}$, the inverse noise covariance matrix.

[1] In low mobility wireless applications, one expects only a few active taps which would fade but would not change position as rapidly.

3.2. Transmitter Induced Constraints

The structure of the data transmitted by the receiver can be used to enhance channel estimation and data recovery. This includes the finite alphabet constraint on the data and redundancy in the form of pilots, the cyclic prefix, and coding. We concentrate here on the former three constraints and relegate the discussion on utilizing the code to [4].

Pilots for initial channel estimation: Pilots are necessary to initialize the estimation process. Fortunately, pilots are sent as part of the OFDM packet and are an integral part of several standards as they are needed for time and frequency recovery in addition to channel estimation. Since there are L_s active taps only, they could be pinned down by sending a similar number of pilots. However, by capitalizing on other *natural* constraints, we can reduce the number of pilots necessary to initialize the estimation process. Let $I_p = \{i_1, i_2, \ldots, i_{L_p}\}$ denote the index set of the pilot bins. The pilots, along with the location of the channel active taps, are the only information initially available. Now from (23), we have the subsystem of equations

$$\boldsymbol{\mathcal{Y}}_{i_{I_p}} = (\operatorname{diag}(\boldsymbol{\mathcal{X}}_i))_{I_p} \underline{\boldsymbol{Q}}^*_{I_s} \underline{\boldsymbol{h}}_{I_s} + \boldsymbol{\mathcal{N}}_{i_{I_p}} \tag{24}$$

This can not be solved uniquely for $\underline{\boldsymbol{h}}$ (or $\underline{\boldsymbol{h}}_{I_s}$) when $L_p < L_s$. However, by solving (23) for $\underline{\boldsymbol{h}}_{-1} = \boldsymbol{0}$ and $\boldsymbol{\Pi}^{-1} = \epsilon \boldsymbol{I}$, [2] we obtain the solution $\underline{\boldsymbol{h}}_{I_s}$ of minimum norm. The weight $\boldsymbol{W}$ is set to the noise inverse covariance matrix $\boldsymbol{R}^{-1}_{\mathcal{N}_{I_p}}$.

Finite alphabet constraint: When the channel is exactly known, we can use the MMSE to perform soft detection of the data as done in (16) and (18). The soft values are then rounded to the nearest alphabet point. However, in the case of incomplete channel information, we can use the finite alphabet constraint on the data to enhance the channel estimate. Specifically, given an initial channel estimate $\boldsymbol{\mathcal{H}}$, we obtain the (soft) MMSE estimate of the data (see (16) and (18)) which is then thresholded to the nearest data point (17).

If the SNR and the channel estimate are decent enough, $\hat{\boldsymbol{\mathcal{X}}}_i$ will be almost correct in spite of the channel estimation errors. The estimate $\hat{\boldsymbol{\mathcal{X}}}_i$ can now be considered as a vector of pilots which can be used for improved LS estimation (from (23)). This procedure could be repeated a number of times until it converges. This approach is similar to decision-directed or blind equalization which capitalizes on the finite alphabet constraint on the data to enhance channel estimation. The finite alphabet property was actually used in [7] to perform blind channel identification.

Redundancy of the cyclic prefix: The input exhibits redundancy in the form of a CP. As we have seen, this decomposes the OFDM channel into two: linear and cyclic. This

[2] We make these choices for $\underline{\boldsymbol{h}}_{-1}$ and $\boldsymbol{\Pi}_{-1}$ because we have no further a priori knowledge about $\underline{\boldsymbol{h}}$. The regularization term $\|\underline{\boldsymbol{h}}\|^2_{\Pi_{-1}}$ in (3) now simply ensures that the resulting solution is unique.

transmitter induced constraint was put into use by employing both subchannels for channel estimation (see (23)).

4. SEMI-BLIND ALGORITHM FOR CHANNEL AND DATA RECOVERY

The above represents various elements of what constitutes an algorithm for semi-blind channel estimation and data recovery. Table 1 is a cohesive summary of this algorithm. We stress the iterative nature of the algorithm; an improvement in the quality of the data estimate leads to a corresponding improvement in the channel estimate which, in turn, leads to an improved data estimate, and so on.

Table 1: *Semi-blind algorithm for channel identification and equalization*

A priori information

- Noise covariance $\boldsymbol{R}_n$, pilot locations (indexed by I_p), and sparsity information (indexed by I_s)

Initial channel estimation

- Solve (24) in the LS sense to obtain an initial channel estimate $\hat{\underline{\boldsymbol{h}}}_0$ (4) and the error covariance $\boldsymbol{\Pi}_0^{-1}$ (5)
  - $\underline{\boldsymbol{h}}_{-1} = \boldsymbol{0}, \quad \boldsymbol{\Pi}_{-1} = \epsilon \boldsymbol{I}, \quad \boldsymbol{W} = \boldsymbol{R}_n^{-1}$

Iterative refinement of estimates: As long as the data keeps changing, perform:

- MMSE data detection
  - Obtain $\hat{\boldsymbol{\mathcal{H}}}$ (14) and its covariance $\text{Cov}_{\mathcal{H}}$ (19)
  - Perform MMSE detection $\hat{\boldsymbol{\mathcal{X}}}_i^{\text{MMSE}}$ (use (16) or (18))
  - Round to the nearest alphabet point $\hat{\boldsymbol{\mathcal{X}}}_i$ (17)
- LS channel estimation
  - Solve (23) in the LS sense to obtain $\underline{\boldsymbol{h}}_0$ (4) and the error covariance $\boldsymbol{\Pi}_0^{-1}$ (5). For this set $\underline{\boldsymbol{h}}_{-1} = \boldsymbol{0}, \quad \boldsymbol{\Pi}_{-1} = \boldsymbol{0}, \quad \boldsymbol{W} = \boldsymbol{R}_n^{-1}$

5. SIMULATIONS

We consider an OFDM binary transmission with $N = 128$ and $L = 16$ through an FIR channel with 15 taps. The output SNR is 11 dB. The packet contains 8 pilots only, which are used to identify a channel with no a priori sparsity information in Fig. 1 (convergence is achieved at the 5th iteration). In Fig. 2, we use the pilots to identify a channel with 8 active taps only whose location is known a priori. The algorithm converges, within three iterations only, to a lower steady-state error.

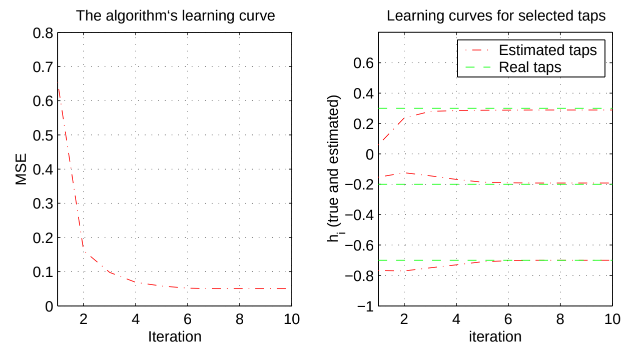

Figure 1: Learning curves for the MSE and selected taps (no sparsity information)

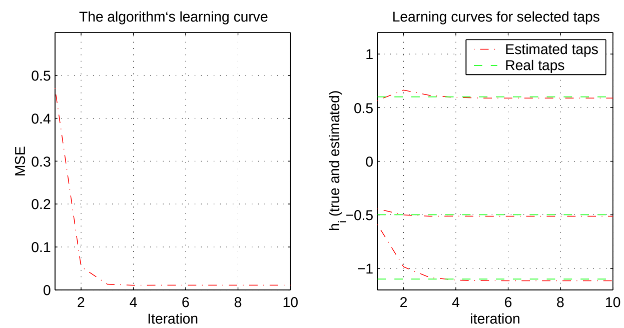

Figure 2: Learning curves for the MSE and selected taps (with a priori sparsity information).

6. REFERENCES

[1] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications: Where Fourier meets Shannon," *IEEE Signal Process. Mag.,* vol. 17, No. 3, pp. 29–48, May 2000.

[2] R. W. Heath Jr. and G. B. Giannakis, "Exploiting input cyclostationarity for blind channel identification in OFDM systems," *IEEE Trans. Signal Process.,* vol 47, pp. 774–780, Mar. 1999.

[3] B. Muquet, M. de Courville, P. Duhamel, and V. Buzenae, "A subspace based blind and semi-blind channel identification method for OFDM systems," *Proc. IEEE Workshop for Signal Process. Advances in Wireless Commun.,* Annapolis, MD, pp. 243–246. May 1999.

[4] G. Al-Rawi, T. Y. Al-Nafouri, A. Bahai, and J. Cioffi, " Iterative joint decoding and blind channel estimation for OFDM systems," *submitted to ICC 2002.*

[5] T. Kailath, A. Sayed, and B. Hassibi, *Linear Estimation,* Prentice Hall, 2000.

[6] X. Wang and R. Liu, "Adaptive channel estimation using cyclic prefix in multicarrier modulation system," *IEEE Commun. Lett.,* vol. 3, no. 10, pp. 291–293, Oct. 1999.

[7] S. Zhou and G. B. Giannakis, " Finite-alphabet based channel estimation for OFDM and related multi-Carrier systems," *IEEE Trans. Commun.,* vol. 49, no. 8, pp. 1402–1414, Aug. 2001.