Bandwidth Efficient Transmission Through Sparse Channels Using a Parametric Channel Estimation - based DFE

Athanasios A. Rontogiannis [1], **Kostas Berberidis** [2]

[1] Institute of Space Applications and Remote Sensing, National Observatory of Athens., 15236, Athens, Greece

[2] Dept. of Computer Engineering & Informatics / C.T.I-R&D, University of Patras, 26500, Rio, Greece

Emails: tronto@space.noa.gr , berberid@ceid.upatras.gr

Abstract

In this paper a novel equalization scheme based on parametric channel estimation is presented. The proposed scheme is appropriate for channels with long and sparse impulse response, such as encountered in high definition digital video transmission systems. The main trait of the channel estimation part is that instead of seeking the whole channel impulse response (CIR), only the unknown time delays and attenuation factors of the physical channel multipath components are efficiently estimated. The equalization part comprises an adaptive decision feedback equalizer (DFE), whose initialization is properly designed for sparse channels. The information needed in the initialization procedure is provided by the channel estimation part. The DFE converges much faster compared to the conventional DFE and the whole scheme utilizes a very short training sequence and hence a significant saving in bandwidth is achieved.

I. INTRODUCTION

In several high-speed wireless communication systems, the involved multipath channels are characterized by a long CIR consisting of a few dominant components, some of which may have quite large time delays with respect to the main signal. A typical application of the kind is High Definition Television systems (HDTV) [1]. Other applications are Multipoint Microwave Distribution Systems, High Speed Communications over Digital Subscriber Loops, Underwater Digital Communications, to name but a few.

If the information signal is transmitted at high symbol rates through such a multipath channel, then the introduced Intersymbol Interference (ISI) has a span of several tens up to hundreds of symbol intervals and may cause a severe degradation of the overall system's performance. Thus, effective channel equalization is required, in order to mitigate interference and achieve reliable data detection. This in turn implies that quite long adaptive equalizers have to be employed at the receiver's end in order to reduce effectively the ISI component of the received signal. Note that the situation is even more demanding whenever the channel frequency response exhibits deep nulls.

Among the various equalization techniques, the decision feedback equalizer (DFE), which incorporates the least mean squares (LMS) algorithm, has been recognized as an effective method [2],[3], especially for channels with long and sparse impulse response. The main drawback of the LMS-based DFE is its slow convergence, which implies that relatively long training sequences are required. It is well known, however, that the training size is a very important issue in the applications of interest, and its reduction results directly in an improvement of the overall system throughput.

In this paper, we present a new equalization technique, which consists of three distinct parts, e.g., a parametric channel estimator, an efficient DFE initialization procedure appropriate for sparse channels and an adaptive LMS-based DFE. The issues of computational complexity and training size are addressed in the whole scheme, with an effort to be kept as low as possible. In the first part of the algorithm, the time delays and attenuation factors of the channel components are estimated using a limited number of training symbols. The proposed method does not rely on preliminary estimates of the CIR, as in [4],[5],[6],[7], and directly estimates the parameters of the multipath channel components. By assuming knowledge of the pulse shaping function [8], we end up with a least squares (LS) problem which is separable with respect to the unknown parameter sets. Specifically, it is shown that the optimization problem can be separated to two different sub-problems. A sub-problem which is non-linear with respect to the time delays and a sub-problem which is linear with respect to the attenuation parameters. After revealing the special structure of the non-linear problem, a computationally efficient linear search method for the estimation of the unknown time delays is developed. Finally, the attenuation parameters are estimated by solving a low order linear LS problem. The new channel estimation algorithm uses symbol rate sampled signals and is easier to implement compared to other parametric blind or semi-blind techniques [9],[10], which are based on oversampling. It incorporates a much smaller number of training symbols compared to LS-based or correlation-based methods [7] for the same estimation accuracy. The channel estimation task is accomplished at a reasonable computational cost, since a highly costly multidimensional search, which is required in other related parameter estimation algorithms [11], is now avoided.

The estimates obtained from the first part of the algorithm are used to initialize the filters of an adaptive LMS-based DFE. The initialization of the DFE filters is realized through an efficient technique, which is appropriately designed for long and sparse channels [12]. The equalizer operates directly into the decision directed mode and exhibits a much faster convergence compared to the conventional LMS-based DFE. Note that the proposed equalizer relies on the training of the channel estimation part only and thus the size of the overall training sequence is kept low, resulting in a significant saving in bandwidth.

We must emphasize that in this work we do not consider a DFE with “sparse” filters as in [12]. What we actually do is that we initialize “full taps” DFE filters according to approximate expressions derived in [12] for sparse channels. The proposed equalization algorithm could be properly modified to incorporate a DFE with a reduced number of taps. Such an approach would lead to an equalization scheme with reduced computational complexity at the cost of a lower estimation accuracy and is not studied any further here.

The paper is outlined as follows. In section II, the channel model considered in the sequel is briefly described. In section III, the parametric cost function is derived and the new CIR estimation algorithm is developed. In section IV, a DFE initialization technique, appropriate for sparse channels, is presented and the whole equalization scheme is described. Simulation results are provided in section V and concluding remarks are given in section VI.

II. The Channel Model

A typical sparse channel (as those used in HDTV tests) is shown in Fig. 1. Although the propagation channel has a large span, it consists of a relatively small number of dominant components [1]. More specifically, if the overall CIR (including transmitter and receiver filters) is assumed to be time-invariant within a small-scale time interval, then it may be written as

$$h(t) = \sum_{k=0}^{p-1} a_k g(t - \tau_k) \tag{1}$$

where a_k and τ_k are the complex attenuation factor and the delay, respectively, of the $k-th$ multipath component. Without loss of generality, it is assumed that $0 < \tau_0 < \tau_1 < \ldots < \tau_{p-1}$. A certain delay τ_i corresponds to the main signal while the remaining delays correspond to undesirable causal and anticausal components. The pulse shaping function $g(t)$ (convolution of transmitter and receiver filters) is assumed to be a raised cosine function with finite support, i.e. $g(t) = 0$ for $t \notin [0, 2L_gT]$, where T is the symbol period. That is, at time L_gT the raised cosine pulse attains its maximum value. We see from (1) that the problem of multipath CIR estimation is reduced to the lower order problem of delay and complex attenuation parameters estimation, provided that the pulse shaping function is known at the receiver.

The sampled overall CIR can be written in vector form as follows

$$\mathbf{h}^T = [h_{-k_1}, h_{-k_1+1}, \cdots, h_0, \cdots, h_{k_2}]$$

where k_1 and k_2 are the number of anticausal and causal channel coefficients, respectively, and $LT = (k_1 + k_2)T$ is the channel span. Note that L is related to the maximum delay as $L = 2L_g + \lceil \frac{\tau_{p-1} - \tau_0}{T} \rceil - 1$. From (1), and using matrix notation, the sampled overall CIR can be re-written in the following form

$$\mathbf{h} = \mathbf{G}(\boldsymbol{\tau})\mathbf{a} \tag{2}$$

where $\boldsymbol{\tau} = [\tau_0 \; \tau_1 \; \ldots \; \tau_{p-1}]^T$, $\mathbf{a} = [a_0 \; a_1 \; \ldots \; a_{p-1}]^T$ and $\mathbf{G}(\boldsymbol{\tau})$ is an $(L+1) \times p$ matrix whose columns are delayed versions of $g(t)$ depending on the unknown parameters τ_k. Specifically this matrix has the form

$$\mathbf{G}(\boldsymbol{\tau}) = \left[\boldsymbol{g}(\tau_0) \; \boldsymbol{g}(\tau_1) \ldots \boldsymbol{g}(\tau_{p-1}) \right]$$

with $\boldsymbol{g}(\tau_k) = [g(-\tau_k) \; g(T - \tau_k) \; \ldots \; g(LT - \tau_k)]^T$.

III. Parametric Channel Estimation

A. Parametric Cost Function

The goal of the new parametric method is the efficient estimation of the unknown parameters' vectors $\boldsymbol{\tau}$ and $\mathbf{a}$, using the smallest possible number of training symbols. This is achieved by taking advantage of the sparse form of the channel and the knowledge of the pulse shaping function. The channel output, which is input to the equalizer, is expressed as follows

$$y(n) = \mathbf{h}^T \mathbf{s}_L(n) + w(n) \tag{3}$$

where $\mathbf{s}_L^T(n) = [s(m), s(m-1), \ldots, s(m-L)]$ is the input data vector with $m = n + k_1$ and $w(n)$ stands for the channel output noise. The input sequence is assumed to be i.i.d. and independent of the noise sequence.

Let us first briefly review conventional LS channel estimation. By assuming that $M+L+1$ training symbols are available, we can form the following system

$$\mathbf{y}_M = \mathbf{S}_{ML}\mathbf{h} + \mathbf{w}_M \tag{4}$$

where

$$\mathbf{y}_M = [y(n)\; y(n+1)\; \cdots\; y(n+M-1)]^T$$

$$\mathbf{S}_{ML} = \begin{bmatrix} s(m) & \cdots & s(m-L) \\ s(m+1) & \cdots & s(m+1-L) \\ \vdots & \vdots & \vdots \\ s(m+M-1) & \cdots & s(m+M-L-1) \end{bmatrix}$$

and $\mathbf{w}_M$ has a form similar to $\mathbf{y}_M$. It is well known that the vector of channel coefficients $\mathbf{h}$ can be obtained from (4) using least squares (LS) estimation as

$$\min_{\mathbf{h}} ||\mathbf{y}_M - \mathbf{S}_{ML}\mathbf{h}||^2, \tag{5}$$

In order to get a unique solution from (5), $\mathbf{S}_{ML}$ must be a tall matrix, i.e. $M > L+1$, and of full column rank.

However, by imposing the channel parametric structure given in (2), the optimization problem takes the following equivalent non-linear LS form

$$\min_{\mathbf{a},\boldsymbol{\tau}} ||\mathbf{y}_M - \boldsymbol{\Phi}(\boldsymbol{\tau})\mathbf{a}||^2, \;\; \boldsymbol{\Phi}(\boldsymbol{\tau}) = \mathbf{S}_{ML}\mathbf{G}(\boldsymbol{\tau}) \tag{6}$$

The above formulation dictates that, first, the p unknown time delays need to be estimated. Then, after computing matrix $\boldsymbol{\Phi}(\boldsymbol{\tau})$, the p unknown attenuation parameters can be estimated by solving the LS problem in (6). This LS problem is considerably reduced with respect to the one in (5) (from $L+1$ to p). As a consequence, M is sufficient to be of the order of p, which means that a much smaller number of training symbols is required compared with the procedure in (5). Moreover, the estimation error is expected to be lower, since a more parsimonious parametrization of the original problem has been adopted. The most critical step in the above procedure is the one related with the estimation of the time delays. To this end, an efficient and accurate technique is proposed in the next section.

B. The new estimation method

It is readily seen that the non-linear LS problem in (6) is separable with respect to the unknown parameters $\boldsymbol{\tau}$ and $\mathbf{a}$. In particular, the LS cost function is nonlinear with respect to the delays $\boldsymbol{\tau}$ and linear with respect to the attenuation factors $\mathbf{a}$. For this type of cost functions the optimization process can be conducted separately with respect to the distinct parameter sets $\boldsymbol{\tau}$ and $\mathbf{a}$ [13, chap. 9]. More specifically

- The delay parameters $\boldsymbol{\tau}$ are obtained from the solution of the following non-linear optimization problem

$$\boldsymbol{\tau}_{opt} = arg \min_{\boldsymbol{\tau}} \{f(\boldsymbol{\tau})\} \tag{7}$$

with

$$f(\boldsymbol{\tau}) = ||(I - \boldsymbol{\Phi}(\boldsymbol{\tau})\boldsymbol{\Phi}^{\dagger}(\boldsymbol{\tau}))\mathbf{y}_M||^2 \tag{8}$$

- The attenuation parameters $\mathbf{a}$ are then determined by the linear LS method as

$$\mathbf{a}_{opt} = \boldsymbol{\Phi}^{\dagger}(\boldsymbol{\tau}_{opt})\mathbf{y}_M \tag{9}$$

where $\dagger$ denotes the pseudoinverse of a matrix. Note, from (7) and (8), that $\boldsymbol{\tau}_{opt}$ is the value of the delay vector which minimizes the projection of vector $\mathbf{y}_M$ to the orthogonal complement of the space spanned by the columns of $\boldsymbol{\Phi}(\boldsymbol{\tau})$. Consequently the same vector maximizes the projection of $\mathbf{y}_M$ to the column space of $\boldsymbol{\Phi}(\boldsymbol{\tau})$ and the optimization problem can be written in the following equivalent form

$$\boldsymbol{\tau}_{opt} = arg \max_{\boldsymbol{\tau}} \{F(\boldsymbol{\tau})\}, \qquad F(\boldsymbol{\tau}) = ||\boldsymbol{\Phi}(\boldsymbol{\tau})\boldsymbol{\Phi}^{\dagger}(\boldsymbol{\tau})\mathbf{y}_M||^2 \tag{10}$$

The non-linear cost function in (10) could be treated either by performing a multidimensional search in the space of parameter set $\boldsymbol{\tau}$ or by applying a non-linear optimization search method, e.g. a Newton type method. In the former case the computational burden may be prohibitive, in the latter the procedure may be trapped in a local minimum, away from the global solution. In the following we will see that, by exploiting the special form of the cost function $F(\cdot)$, an efficient and accurate method for estimating $\boldsymbol{\tau}_{opt}$ can be derived, which overcomes the above mentioned drawbacks. Specifically, the cost function $F(\cdot)$ can be expressed as follows

$$F(\boldsymbol{\tau}) = \mathbf{q}^H(\boldsymbol{\tau})\mathbf{A}^{-1}(\boldsymbol{\tau})\mathbf{q}(\boldsymbol{\tau}) \tag{11}$$

where

$$\mathbf{A}(\boldsymbol{\tau}) = \mathbf{G}^H(\boldsymbol{\tau})\mathbf{S}_{ML}^H\mathbf{S}_{ML}\mathbf{G}(\boldsymbol{\tau}) \tag{12}$$

$$\mathbf{q}^H(\boldsymbol{\tau}) = \boldsymbol{\Phi}^H(\tau)\mathbf{y}_M = [q(\tau_0) \ \ q(\tau_1) \ \ \cdots \ \ q(\tau_{p-1})] \tag{13}$$

and $q(\tau_i) = \mathbf{y}_M^H\mathbf{S}_{ML}\mathbf{g}(\tau_i)$. That is, each element of vector $\mathbf{q}(\boldsymbol{\tau})$ is a function of a single and different time delay parameter. Therefore, we deduce that the cost function $F(\boldsymbol{\tau})$ would be decoupled with respect to the delay parameters, if matrix $\mathbf{A}^{-1}(\boldsymbol{\tau})$ could be approximated by a diagonal matrix.

Due to the i.i.d. property of the input sequence, the law of large numbers dictates that the middle term $\mathbf{S}_{ML}^H\mathbf{S}_{ML}$ of $\mathbf{A}(\boldsymbol{\tau})$, tends to a diagonal matrix with equal elements for sufficiently large M. Moreover, the columns of $\mathbf{G}(\boldsymbol{\tau})$ contain shifted versions of a raised cosine pulse shaping filter. The inner product of two columns of $\mathbf{G}(\boldsymbol{\tau})$ approximates the value of the autocorrelation function of the raised cosine pulse for a lag equal to the difference of the corresponding time delays. Consequently, even for a difference as small as one symbol period the value of the inner product is one order of magnitude smaller than its maximum value, which corresponds to a zero time difference. Thus matrix $\mathbf{G}(\boldsymbol{\tau})$ has a structure very similar to an orthogonal one, especially for the type of channels we consider here. Due to the diagonal form of $\mathbf{S}_{ML}^H\mathbf{S}_{ML}$ and the near-orthogonality of $\mathbf{G}(\boldsymbol{\tau})$, it is easily verified that $\mathbf{A}(\boldsymbol{\tau})$ tends to a diagonal matrix. Using Taylor expansion of $\mathbf{A}(\boldsymbol{\tau})$, it is then straightforward to show that the inverse of this matrix, i.e. $\mathbf{A}^{-1}(\boldsymbol{\tau})$, possesses the same near-diagonal structure.

From (11), (13) and the analysis concerning $\mathbf{A}^{-1}(\boldsymbol{\tau})$ we deduce that $F(\boldsymbol{\tau})$ can be approximated by a summation of p positive terms, each of which is a function of one delay parameter only. Therefore, the optimization search can be performed separately for each $\tau_i,\ i = 0, 1, \ldots, p-1$ and independently of the other delay parameters. In other words, instead of a p-dimensional search of exponential complexity, p one-dimensional searches are sufficient for the solution of the optimization problem, which dramatically reduces complexity. It is important to note that, as also verified in the simulations, a value of M much lower than L is sufficient to ensure the near-diagonal structure of matrix $\mathbf{A}(\boldsymbol{\tau})$, which leads to the decoupling of the delay parameters. Thus, the number of training symbols is kept low resulting in a significant saving in bandwidth.

After reducing the initial multidimensional problem to p one-dimensional problems, an appropriate search method must be utilized for optimization. In order to avoid trapping of the one-dimensional searches in local optimum points a gradient- or Newton-type search method is avoided. Instead, a simple grid search method seems to work well for the cost function under consideration. In the proposed estimation method a p-dimensional grid is initially defined with a linear step size, say δ. Then $F(\boldsymbol{\tau})$ is evaluated in p lines of the grid, according to (11). A more efficient method is based on an equivalent expression of $F(\boldsymbol{\tau})$ which is described below. Let us first define the following quantities

$$\boldsymbol{\phi}_i = \mathbf{S}_{ML}\boldsymbol{g}(\tau_i), \quad \boldsymbol{\Phi}_{i:j} = \left[\boldsymbol{\phi}_i\ \boldsymbol{\phi}_{i+1}\ \ldots\ \boldsymbol{\phi}_j\right],$$

$i, j = 0, 1, \ldots, p-1,\ j \geq i$. That is, $\boldsymbol{\phi}_i$ is the i-th column of matrix $\boldsymbol{\Phi}(\boldsymbol{\tau})$ and $\boldsymbol{\Phi}_{i:j}$ is its submatrix formed by columns i through j.

Based on these definitions, $\mathbf{A}(\boldsymbol{\tau})$ and $\mathbf{q}(\boldsymbol{\tau})$ given in (12) and (13) can be partitioned as follows

$$\mathbf{A}(\boldsymbol{\tau}) = \begin{bmatrix} \boldsymbol{\Phi}_{0:p-2}^H \boldsymbol{\Phi}_{0:p-2} & \boldsymbol{\Phi}_{0:p-2}^H \boldsymbol{\phi}_{p-1} \\ \boldsymbol{\phi}_{p-1}^H \boldsymbol{\Phi}_{0:p-2} & \boldsymbol{\phi}_{p-1}^H \boldsymbol{\phi}_{p-1} \end{bmatrix} \tag{14}$$

$$\boldsymbol{q}^H(\boldsymbol{\tau}) = \boldsymbol{z}^H \begin{bmatrix} \boldsymbol{\Phi}_{0:p-2} & \boldsymbol{\phi}_{p-1} \end{bmatrix} \tag{15}$$

Matrix $\mathbf{A}^{-1}(\boldsymbol{\tau})$ is easily obtained from (14) by applying the matrix inversion lemma. If we now substitute the expression of $\mathbf{A}^{-1}(\boldsymbol{\tau})$ and (15) in (11), then after some algebra we end up with the following expression for the cost function $F(\boldsymbol{\tau})$

$$F(\boldsymbol{\tau}) = \boldsymbol{y}_M^H \boldsymbol{\Phi}_{0:p-2} (\boldsymbol{\Phi}_{0:p-2}^H \boldsymbol{\Phi}_{0:p-2})^{-1} \boldsymbol{\Phi}_{0:p-2}^H \boldsymbol{y}_M + \frac{||\boldsymbol{\phi}_{p-1}^H (I - \boldsymbol{\Phi}_{0:p-2} \boldsymbol{\Phi}_{0:p-2}^{\dagger}) \boldsymbol{y}_M||^2}{||\boldsymbol{\phi}_{p-1}^H (I - \boldsymbol{\Phi}_{0:p-2} \boldsymbol{\Phi}_{0:p-2}^{\dagger})||^2} \tag{16}$$

We observe that the first term in the right hand side (RHS) of (16) is the cost function of order $p-1$, i.e. the cost function which arises by assuming that the CIR consists of $p-1$ components. By further decomposing this term using the same procedure, $F(\boldsymbol{\tau})$ can be written in the final form

$$F(\boldsymbol{\tau}) = \frac{||\boldsymbol{\phi}_0^H \boldsymbol{y}_M||^2}{||\boldsymbol{\phi}_0||^2} + \sum_{i=1}^{p-1} \frac{||\boldsymbol{\phi}_i^H (I - \boldsymbol{\Phi}_{0:i-1} \boldsymbol{\Phi}_{0:i-1}^{\dagger}) \boldsymbol{y}_M||^2}{||\boldsymbol{\phi}_i^H (I - \boldsymbol{\Phi}_{0:i-1} \boldsymbol{\Phi}_{0:i-1}^{\dagger})||^2} \tag{17}$$

Notice that $F(\boldsymbol{\tau})$ consists of p non-negative terms. The first term is a function of τ_0 only and the i-th term depends on $\tau_0, \ldots, \tau_{i-1}$. Due to the non-negativeness of these terms, their optimum points would coincide with the corresponding components of the global optimum point of $F(\boldsymbol{\tau})$, if they were independent. Although, this is not true, in general, the decoupling of the delay parameters discussed above allows for their efficient estimation based on the form of $F(\boldsymbol{\tau})$ given in (17). The basic steps of the proposed parametric channel estimation algorithm, using the expression of $F(\boldsymbol{\tau})$ from (17), are summarized in Table I.

Since the number of paths p is not known, an over-estimation $\hat{p}$ of this number is made at the beginning of the algorithm. Due to the form of the objective function in (17), it is clear that the maximization of the first p terms of $F(\boldsymbol{\tau})$ (which now consists of $\hat{p}$ terms), leads to the same delay estimates as if the exact

1) Set overestimated values $\hat{L}$, $\hat{p}$ for unknown L, p, respectively and choose a linear step size δ.
2) Maximize the first term of $F(\boldsymbol{\tau})$ in (17) with respect to τ_0 by evaluating it at $\tau_0 = j\delta$, $j = 0, 1, \ldots, \frac{\hat{L}T}{\delta}$. Let $\hat{\tau}_0$ be the optimum point.
3) Substitute τ_0 with $\hat{\tau}_0$ in the second term of $F(\boldsymbol{\tau})$ and maximize this term with respect to τ_1, which gives say $\hat{\tau}_1$ as the optimum point.
4) Repeat step 3 for $i = 2, \ldots, \hat{p} - 1$ by substituting $\tau_0, \ldots, \tau_{i-1}$, with the optimum values obtained in the previous steps and maximize the $i+1$-th term of $F(\boldsymbol{\tau})$ with respect to τ_i.
5) Obtain the attenuation parameters from (9).

TABLE I

SUMMARY OF THE CHANNEL ESTIMATION ALGORITHM

number of paths was known in advance. The maximization of the remaining $\hat{p} - p$ terms only slightly affects the performance of the algorithm (as also verified by extensive simulations), at the cost of some extra computational complexity.

The delay vector estimate, obtained from the procedure described above, can be further improved by applying a few steps of a secondary optimization search, e.g., a Gauss-Newton (G-N) search method [10]. The G-N search is guaranteed to reach the global optimum point under the assumption that the initial delay vector estimate is in the neighborhood of this point. Thus, one can trade off between improved estimation accuracy and increased computational complexity with the incorporation of a secondary G-N optimization procedure in the algorithm. We have observed through simulations that such an approach can lead to performance gains, especially in case there exist closely-spaced multipath components in the channel IR.

IV. Adaptive DFE Initialization

In the equalization part of the algorithm, the LMS-based adaptive DFE is employed, with the DFE filters properly initialized. More specifically, after estimating the multipath channel parameters as described in the previous section, an estimate of the CIR can be directly obtained from (2). Subsequently, the optimum values, in the mean squared error (MSE) sense, of the K-length feedforward (FF) filter and the N-length feedback filter coefficients of the DFE can be computed according to the following equations [12]

$$\mathbf{c}_K = (\mathbf{H}_1\mathbf{H}_1^H + \frac{\sigma_w^2}{\sigma_u^2}\mathbf{I}_K)^{-1}\mathbf{H}_1\mathbf{e}_{K+k_1} \tag{18}$$

$$\mathbf{b}_N = \begin{bmatrix} -\mathbf{H}_2^H\mathbf{c}_K \\ \mathbf{0}_{(N-k_2)\times 1} \end{bmatrix} \tag{19}$$

where $\mathbf{I}_K$ is the $K \times K$ identity matrix, $\mathbf{e}_{K+k_1} = [0 \ \ldots \ 0 \ 1]^T$ is a $(K+k_1) \times 1$ vector, σ_u^2, σ_w^2 are the input and noise variances respectively and the $K \times (k_1 + K)$, $K \times k_2$ matrices $\mathbf{H}_1, \mathbf{H}_2$ are given as follows

$$\mathbf{H}_1 = \left[\begin{array}{ccc|cccc} h_{-k_1} & \cdots & h_{-1} & h_0 & \cdots & \cdots & h_{K-1} \\ 0 & \cdots & h_{-2} & h_{-1} & \cdots & \cdots & h_{K-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & h_{-k_1} & h_{-k_1+1} & \cdots & \cdots & h_{K-k_1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & h_{-k_1} & \cdots & h_0 \end{array}\right] \tag{20}$$

$$\mathbf{H}_2 = \begin{bmatrix} h_K & h_{K+1} & \cdots & h_{k_2} & 0 & \cdots & 0 \\ h_{K-1} & h_K & \cdots & \cdots & h_{k_2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ h_1 & h_2 & \cdots & \cdots & \cdots & \cdots & h_{k_2} \end{bmatrix} \tag{21}$$

It can be shown [12] that for the special class of channels we consider here and under reasonable assumptions concerning the filter sizes, the FF filter given in (18) can be very well approximated by the solution of the much more simpler system of equations

$$\mathbf{H}_{12}^{H}\mathbf{c}_{K} = \mathbf{e}_{K} \tag{22}$$

where $\mathbf{e}_K = [0 \ \ldots \ 0 \ 1]^T$ and $\mathbf{H}_{12}$ is the right $K \times K$ block of $\mathbf{H}_1$. From (22) we recognize that the FF filter of the DFE is approximated by the last column of the inverse of matrix $\mathbf{H}_{12}^H$. Since matrix $\mathbf{H}_{12}^H$ has a Toeplitz structure, its inversion requires $O(K^2)$ operations. Once the FF filter is computed, the FB filter can be obtained from (19). Due to the Toeplitz form of matrix $\mathbf{H}_2^H$, the matrix-by-vector product $\mathbf{H}_2^H\mathbf{c}_K$ can be efficiently computed using FFT. Such an approach reduces in general the computational complexity compared to direct matrix-vector multiplication from $O(Kk_2)$ to $O(k_2 log(k_2))$.

Figure 2 shows the three distinct parts of the proposed channel estimation / equalization scheme. In the beginning of each incoming information frame, the block corresponding to the training portion is buffered and used to estimate the channel as described in section III. Then, the FF and FB coefficients of the DFE are computed as described above and the values obtained are, in fact, used to initialize the conventional LMS-based DFE. The FB filter is initially fed with the last N of the $M + L + 1$ known symbols and starts operating directly into the decision directed mode.

Further reduction in complexity and required bandwidth

It is important to note that in some applications, the delay parameters of the propagation medium change very slowly compared with transmission data rate [1]. On the other hand, the attenuation parameters tend to vary quite rapidly. Therefore, the delay estimation part of the algorithm could be executed in certain frames only, which are separated by a predetermined number of other frames in which the delays are considered invariant. In the remaining frames, the channel estimation task could be restricted to attenuation parameters estimation only. The estimation of the attenuation parameters can be performed in each frame according to (9), after updating matrix $\mathbf{\Phi}(\boldsymbol{\tau})$ of (6) with the new training symbols. Alternatively, vector $\mathbf{a}$ can be estimated blindly in each frame, by adopting a multichannel model and properly transforming the cost function (6) according to a certain blind criterion. Both methods offer further reduction in complexity and more efficient use of bandwidth and are topics under current investigation.

V. SIMULATION RESULTS

In this section we present some indicative simulation results of the new channel estimation/equalization method. We consider as test channel the channel of Fig. 1, whose multipath parameters are $\mathbf{a} = [0.1 + 0.1j, 0.2, 1, 0.08 - 0.17j, 0.15 - 0.2j, -0.16 - 0.19j]^T$ and $\boldsymbol{\tau} = [4.11T, 16.39T, 28, 60.71T, \ 90.95T, 91.71T]^T$. The channel of Fig. 1 has a form similar to that of the HDTV test channels of [1]. The CIR consists of the main echo, two anticausal and three causal components, two of which are closely spaced. The input sequence is taken from a QAM-16 alphabet. In Fig. 3 the proposed estimation method is compared with direct LS channel estimation for different signal-to-noise ratios (SNR), in terms of the root mean squared error (RMSE) given by

$$RMSE = \sqrt{(1/P)\sum_{k=1}^{P} ||\mathbf{h} - \mathbf{h}_{est}^{(k)}||^2}$$

P is the number of independent experiments and $\mathbf{h}_{est}^{(k)}$ is the estimated CIR after the k-th simulation run. In this experiment, P is taken equal to 50. A square root raised cosine pulse shaping filter with 11.5% roll-off is assumed and a step size $\delta = 0.2T$ has been used in the parametric channel estimation method. We see that the proposed method outperforms LS channel estimation, especially for low to medium SNR's, even in the case it utilizes a much lower number of training symbols, i.e. $M = 40$ for the parametric method and $M = 160$ for direct LS estimation. Recall that $M + L + 1$ represents the total number of training symbols used by the channel estimation methods. From Fig. 3 we also observe that over-estimating the number of paths affects only slightly (and rather improves) the performance of the parametric algorithm.

In Fig. 4 the MSE of the proposed DFE is compared with that of the conventional constantly trained DFE for SNR=25dB. The curves result from the average of 200 independent simulation runs. The FF and FB filter lengths were taken equal to 60 and 100 respectively. We see that even though the proposed DFE runs directly in the decision directed mode, (i.e. it relies only on the 140 training symbols of the channel estimation part), it converges almost immediately and remains in steady state during the whole experiment. On the contrary, a number of almost 4000 iterations is required for the classical DFE to converge. Therefore, it is expected that a very high number of training symbols would be necessary for convergence, if the algorithm had been designed to switch once in the decision directed mode. A significant saving in bandwidth is thus obtained, if the proposed method is adopted.

Finally, in Fig. 5 symbol error rate (SER) curves for different SNR's are depicted. In all cases, the DFE of section IV is used, but the initialization of the DFE filters is performed either through conventional LS channel estimation or based on parametric channel estimation, as proposed in this work. The size of the training sequence is either 140 ($M = 40$) or 260 ($M = 160$) and the remaining part of each information frame consists of 2000 QAM-16 symbols. The experiment is performed for 1000 information frames. The results shown in Fig. 5 indicate the importance of a good initial channel estimate in the error rate performance of the adaptive LMS-based DFE. Note that there is an almost 2-dB difference between the proposed method and the conventional approach. Obviously, this difference increases as the information frame size decreases and conversely.

VI. CONCLUSION

A joint channel estimation and equalization scheme appropriate for sparse channels has been developed. The channel estimator utilizes a very short training sequence and estimates efficiently the parameters of the channel multipath components. The DFE involved in the equalization part is initialized using the information provided by the channel estimator, thus improving dramatically its convergence speed. The proposed scheme offers efficient use of the available bandwidth at a reasonable computational cost.

REFERENCES

[1] M. Ghosh, “Blind Decision Feedback Equalization for Terrestrial Television Receivers”, *Proceedings of the IEEE*, vol. 86, no. 10, pp. 2070-2081, Oct. 1998.

[2] J.G. Proakis, “Adaptive Equalization of TDMA Digital Mobile Radio”, *IEEE Trans. on Vehicular Technology*, vol. 40, no. 2, pp. 333-341, May 1991.

[3] P.K. Shukla, L.F. Turner, “Examination of an Adaptive DFE and MLSE / near MLSE for Fading Multipath Radio Channels”, *IEE Proceedings I Communications, Speech and Vision*”, vol. 139, no. 4, pp. 418-428, Aug. 1992.

[4] M.C. Vanderveen, A. Paulraj, “Improved Blind Channel Identification Using a Parametric Approach”, *IEEE Communication Letters*, vol.2, no.8, pp. 226-228, Aug. 1998.

[5] M.C. Vanderveen, A. van der Veen, A. Paulraj, “Estimation of Multipath Channel Parameters in Wireless Communications”, *IEEE Trans. on Signal Processing*, vol. 46, no. 3, pp. 682-690, Mar. 1998.

[6] J-T. Chen, A. Paulraj, U. Reddy, “Multichannel Maximum-Likelihood Sequence Estimation (MLSE) Equalizer for GSM Using a Parametric Channel Model”, *IEEE Trans. on Communications*, vol. 47, no. 1, pp. 53-63, Jan. 1999.

[7] S. Özen, M.D. Zoltowski,“Time-of-Arrival (TOA) Estimation Based Structured Sparse Channel Estimation Algorithm, with Applications to Digital TV-Receivers”, In Proc. of *ICASSP*, vol. IV, pp. 481-484, Hong Kong, Apr. 2003.

[8] Z. Ding, “Multipath Channel Identification Based on Partial System Information”, *IEEE Trans. on Signal Processing*, vol. 45, pp. 235-240, Jan. 1997.

[9] I. Kang, M.P. Fitz, S.B. Gelfand, “Blind Estimation of Multipath Channel Parameters: A Modal Analysis Approach”, *IEEE Trans. on Communications*, vol. 47, no.8, pp. 1140-1150, Aug. 1999.

[10] A.A. Rontogiannis, A. Marava, K. Berberidis, J. Palicot, “Efficient Multipath Channel Estimation Using a Semi-Blind Parametric Technique”, In Proc. of *ICASSP*, vol. IV, pp. 477-480, Hong Kong, Apr. 2003.

[11] L. Perros-Meilhac, E. Moulines, K. Abed-Meraim, P. Chevalier, P. Duhamel, “Blind Identification of Multipath Channels: A Parametric Subspace Approach”, *IEEE Trans. on Signal Processing*, July 2001, pp. 1468-80.

[12] A.A. Rontogiannis, K. Berberidis, “Efficient Decision Feedback Equalization for Sparse Wireless Channels”, *IEEE Trans. on Wireless Communications*, vol.2, Issue 3, pp. 570-581, May 2003.

[13] A. Bjorck, *Numerical Methods for Least Squares Problems,* Soc. for Ind. and Applied Math. (SIAM), 1996.

Figures

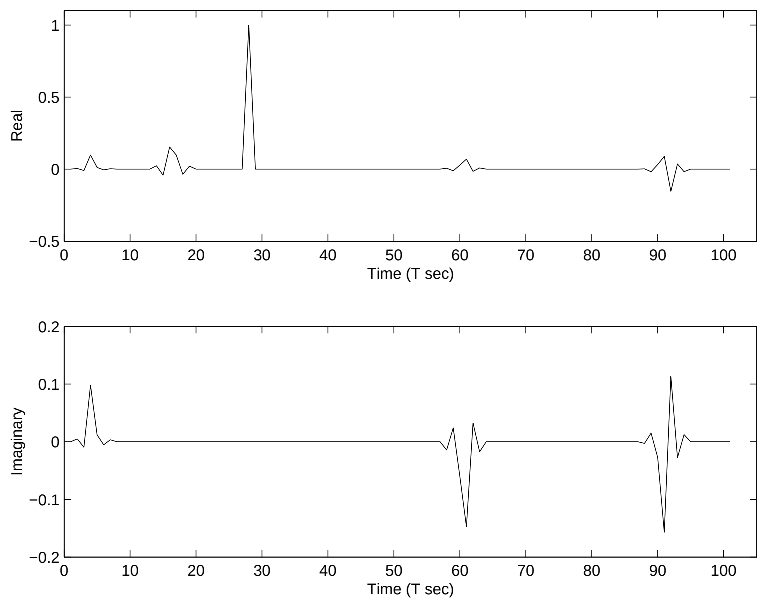

Real
1
0.5
0
−0.5
0 10 20 30 40 50 60 70 80 90 100
Time (T sec)
Imaginary
0.2
0.1
0
−0.1
−0.2
0 10 20 30 40 50 60 70 80 90 100
Time (T sec)

Fig. 1.

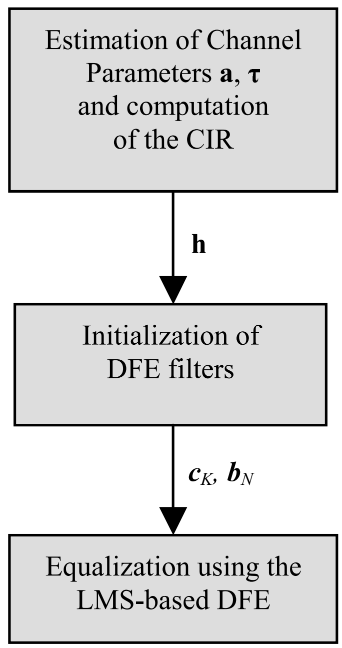

Estimation of Channel
Parameters **a**, **τ**
and computation
of the CIR
h
Initialization of
DFE filters
$\boldsymbol{c}_K$, $\boldsymbol{b}_N$
Equalization using the
LMS-based DFE

Fig. 2.

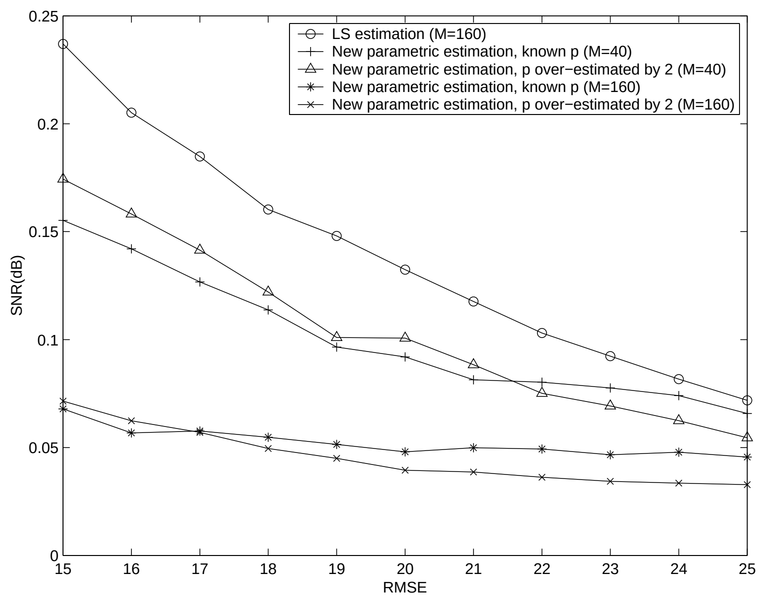
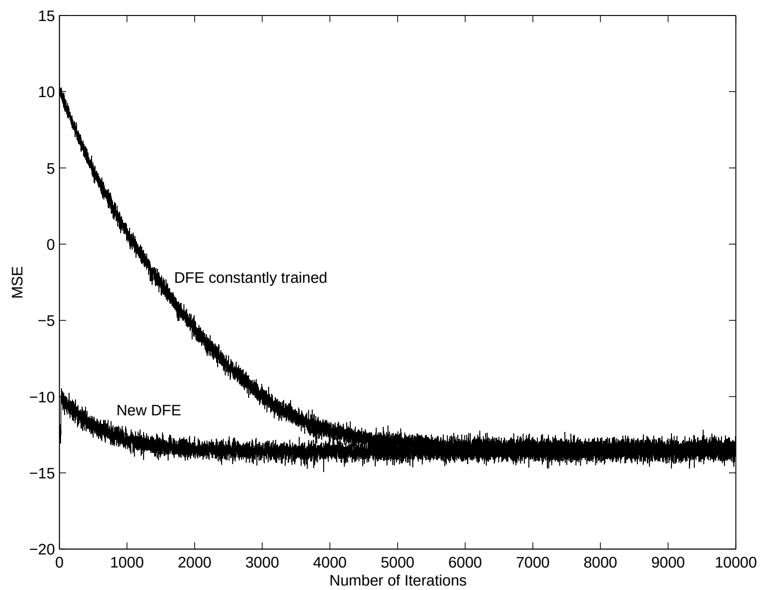

LS estimation (M=160)
New parametric estimation, known p (M=40)
New parametric estimation, p over−estimated by 2 (M=40)
New parametric estimation, known p (M=160)
New parametric estimation, p over−estimated by 2 (M=160)
SNR(dB)
RMSE
0.25
0.2
0.15
0.1
0.05
0
15
16
17
18
19
20
21
22
23
24
25

Fig. 3.



15
10
5
0
−5
−10
−15
−20
MSE
DFE constantly trained
New DFE
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
10000
Number of Iterations

Fig. 4.

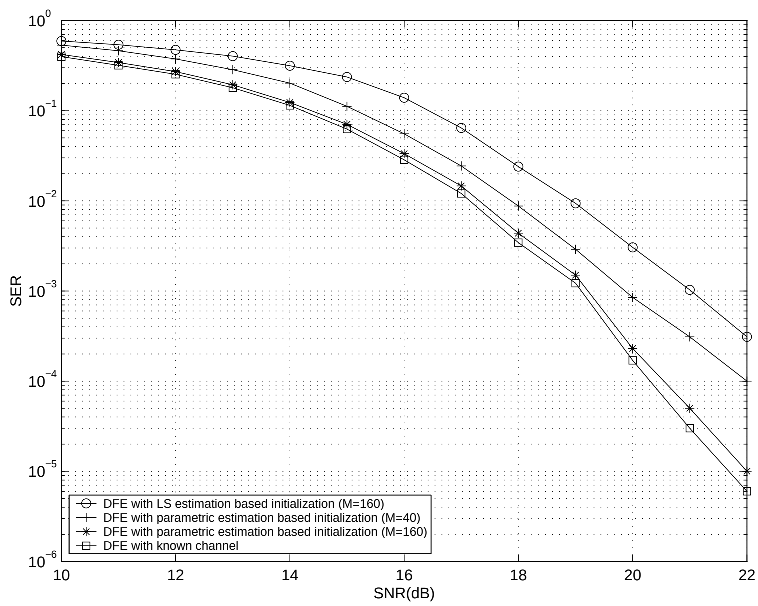

SER
SNR(dB)
10^0
10^−1
10^−2
10^−3
10^−4
10^−5
10^−6
10
12
14
16
18
20
22
DFE with LS estimation based initialization (M=160)
DFE with parametric estimation based initialization (M=40)
DFE with parametric estimation based initialization (M=160)
DFE with known channel

Fig. 5.