SPARSE SUBSPACE TRACKING TECHNIQUES FOR ADAPTIVE BLIND CHANNEL IDENTIFICATION IN OFDM SYSTEMS

Christos G. Tsinos, Aris S. Lalos and Kostas Berberidis

Dept. of Computer Engineering & Informatics and CTI/RU8 , University of Patras 26500, Rio-Patras, Greece
Phone: + 30 2610 996970, fax: + 30 2610 996971 , E-mails: {tsinos,lalos,berberid}@ceid.upatras.gr

ABSTRACT

In this paper novel subspace-based blind schemes are proposed and applied to the sparse channel identification problem. Moreover, adaptive sparse subspace tracking methods are proposed so as to provide efficient real-time implementations. The new algorithms exploit the subspace sparsity either via employing ℓ_1-norm relaxation or through greedy-based optimization. The derived schemes have been tested in a Zero-Prefix Orthogonal Frequency Division Multiplexing (ZP-OFDM) system and it turns out that, compared to state-of-art existing schemes, they offer improved performance in terms of convergence rate and steady-state error.

1. INTRODUCTION

The channel estimation task is an important constituent part of the OFDM-based systems. It is often the case that in high-speed wireless communications, the involved multipath channels are typically sparse, i.e., they are characterized by a long Channel Impulse Response (CIR) having only a few dominant components. In the recent literature of system identification, there is a growing interest in exploiting such sparse characteristics. Two major approaches to sparse system identification are ℓ_1-minimization (basis pursuit methods), and greedy algorithms (matching pursuit methods). Basis pursuit methods solve a ℓ_1 constrained convex minimization problem. Greedy algorithms, on the other hand, compute iteratively the signal's support set until a halting condition is met [1].

Traditionally, channel estimation is achieved by sending training sequences through the channel. However, when the channel is varying, even slowly, the training sequence needs to be sent periodically, so as to update the channel estimates. Hence, the transmission efficiency is reduced. The increasing demand for high-bit-rate digital mobile communications makes blind channel identification very attractive. During the past years various blind identification approaches have been proposed either by exploiting the cyclostationarity present in Cyclic Prefix OFDM (CP-OFDM) [2] or by subspace-based estimation techniques [3]-[4].

New channel estimation techniques appeared recently in literature which properly exploit the involved channel sparsity [5], however, the majority of these techniques are training-based. In this paper, we derive sparse channel estimation techniques for ZP-OFDM systems which can operate blindly, i.e., without requiring any pilot tones. Moreover, starting from [4], we develop new subspace tracking methods, that exploit the sparsity of the eigenvectors, and lead to adaptive implementations of the new blind channel estimation techniques. To the best of our knowledge, this is the first time that the adaptive subspace tracking problem is studied in a sparse context.

We follow two different approaches in order to solve the underlying sparse subspace problem and then identify the non-negligible CIR taps. The first approach is based on ℓ_1-norm relaxation while the other one on a greedy optimization strategy. Compared to the non-sparsity aware blind adaptive channel estimation schemes, both sparse approaches exhibit faster convergence and improved tracking capabilities.

The rest of the paper is organized as follows: In Section 2, the problem is formulated and some preliminaries concerning greedy and ℓ_1 relaxation methods are provided. In Section 3, the new adaptive schemes are derived. Simulation results are presented in Section 4. Finally, Section 5 concludes the paper.

2. SYSTEM MODEL - PROBLEM FORMULATION

Let us consider a baseband discrete time ZP-OFDM transmission scheme in which the length N, n-th symbol block $\mathbf{s}_n = [s_n(1), \ \ldots, s_n(N)]$ is modulated by the Inverse Discrete Fourier Transform (IDFT) and then is padded with L zeros. The $(N+L)\times 1$ transmitted block may be written as:

$$\mathbf{x}_n = \begin{bmatrix} \mathbf{F}^H \\ \mathbf{0}_{L\times N} \end{bmatrix} \mathbf{s}_n \tag{1}$$

The transmitted signal propagates through a multipath AWGN channel with CIR $\mathbf{h} = [h_0 \ h_1 \ h_2 \ \ldots h_L]^T$. From the $(L+1)$ CIR coefficients[1] only S are assumed to be non-negligible, located at the positions $k_1, \cdots k_S$.

In the following, we assume that the receiver is synchronized with the transmitter and also a perfect carrier recovery is achieved, which implies that no intercarrier interference (ICI) is introduced. In case a ZP is employed, the n-th received data block $\mathbf{y}_n$ of length $N+L$ can be expressed as

$$\mathbf{y}_n = \mathbf{H}\mathbf{x}_n + \mathbf{z}_n, \tag{2}$$

where $\mathbf{H}$ is the $(N+L)\times N$ convolution matrix of filter $\mathbf{h}$.

[1] Since a ZP of length L is used, it is assumed that the channel has a finite impulse response of length at most equal to L+1.

2.1. Subspace Based Identification

In a blind identification procedure, the $(L+1)\times 1$ channel vector is solely estimated from the observations of $\mathbf{y}_n$. As in the method of Tong et al. [3] the identification is based on the $(N+L)\times(N+L)$ autocorrelation matrix $\mathbf{R}$ of the received data vector $\mathbf{y}_n$ of eq.(2). Assuming that the elements of $\mathbf{s}_n$ are i.i.d. and of unit norm, and using the orthonormality of the IDFT matrix $\mathbf{F}$ we can easily get that

$$\mathbf{R} = E\left[\mathbf{y}_n\mathbf{y}_n^H\right] = \mathbf{H}\mathbf{H}^H + \sigma^2\mathbf{I}_{N+L} \tag{3}$$

2.1.1. *Subspace Decomposition*

The Singular Value Decomposition (SVD) of matrix $\mathbf{R}$ is

$$\mathbf{R} = [\mathbf{U}_s \quad \mathbf{U}_z]\left[\begin{array}{cc} \mathbf{\Lambda}_s + \sigma^2\mathbf{I}_N & \mathbf{0} \\ \mathbf{0} & \sigma^2\mathbf{I}_L \end{array}\right][\mathbf{U}_s \quad \mathbf{U}_z]^H \tag{4}$$

where $\mathbf{U}_s, \mathbf{U}_z$ form a basis of the signal and noise subspace, respectively, and the $N\times N$ diagonal matrix $\mathbf{\Lambda}_s$ contains the corresponding N largest eigenvalues of $\mathbf{R}$.

2.1.2. *Channel Identification*

Before proceeding note that the columns of $\mathbf{H}$ span the signal subspace. Thus, for every $(N+L)\times 1$ noise subspace vector $\mathbf{v}_i$, $i=1,\ldots,L$ we have $\mathbf{v}_i^H\mathbf{H} = \mathbf{0}_L$ or equivalently

$$\mathbf{h}^T\mathbf{V}_i^* = \mathbf{0}_L, \ \ 1 \le i \le L \tag{5}$$

where $\mathbf{V}_i$ is a $(L+1)\times N$ Hankel matrix, constructed by the elements of vector $\mathbf{v}_i$ as follows:

$$\mathbf{V}_i = \left[\begin{array}{cccc} v_i(1) & v_i(2) & \ldots & v_i(N) \\ v_i(2) & v_i(3) & \ldots & v_i(N+1) \\ \vdots & \vdots & & \vdots \\ v_i(L+1) & v_i(L+2) & & v_i(N+L) \end{array}\right] \tag{6}$$

Under appropriate conditions detailed in [6], the noise subspace related matrix $\mathbf{V}_i$ determines uniquely the channel coefficients up to a multiplicative constant. Thus, the unknown CIR is the solution of the following minimization problem

$$\min_{\mathbf{h}} J_w(\mathbf{h}) \equiv \min_{\mathbf{h}} \mathbf{h}^H\mathbf{W}\mathbf{h}, \ \ \textit{s.t.} \ \ \|\mathbf{h}\|_{l_2} = 1 \tag{7}$$

where $\mathbf{W} = \sum_{i=1}^{L}\mathbf{V}_i\mathbf{V}_i^H$. Alternatively, it can be seen that the unknown channel $\mathbf{h}$ may be estimated via the signal subspace as well. However, here we adopt a noise subspace-based method since, as shown in [4], it may lead to CIR estimation by relying only on one noise eigenvector. On the other hand, a signal subspace-based method requires knowledge of all N associated eigenvectors, hence it is impractical for real-time implementations.

Once $\mathbf{W}$ is computed, the normalized channel impulse response $\mathbf{h}_b$ can be obtained as the singular vector of $\mathbf{W}$'s smallest singular value. The estimated CIR is the unique (up to a scalar factor α) vector $\mathbf{h}$, i.e.,

$$\mathbf{h} = \alpha\mathbf{h}_b \tag{8}$$

2.2. Identification of Sparce CIR's

Recall that $\mathbf{h}\in\mathbb{C}^{L+1}$ is the S-sparse CIR vector, with $S = |supp(\mathbf{h})| << L+1$ the sparsity order of the support set. In order to exploit the CIR filter's sparsity, we formulate the subspace problem of (7) as the constrained optimization problem

$$\min_{\mathbf{h}} J_w(\mathbf{h}) \quad \textit{s. t.} \quad \|\mathbf{h}\|_{\ell_0} \le \delta, \quad \|\mathbf{h}\|_{\ell_2} = 1, \tag{9}$$

meaning that we seek the support set of the CIR (for a predefined tolerance δ) that minimizes eq.(7).

Since finding the optimal solution for this problem is not computationally feasible, two common sub-optimal approaches have been proposed in literature, the l_1-norm relaxation and the greedy optimization strategies. In the first approach the ℓ_0-norm is replaced by the ℓ_1-norm, thus forming a tractable problem with a global minimum, i.e.,

$$\min_{\mathbf{h}} J_w(\mathbf{h}) \quad \textit{s. t.} \quad \|\mathbf{h}\|_{\ell_1} \le \delta, \quad \|\mathbf{h}\|_{\ell_2} = 1. \tag{10}$$

In the second approach we cast a heuristic iterative algorithm so as to find recursively the support set Ω of $\mathbf{h}$ such that

$$\min_{\mathbf{h}_{\Omega(t)}} J_{w_{\Omega(t)}}(\mathbf{h}_{\Omega(t)}), \ \ s.t. \quad \|\mathbf{h}_{\Omega(t)}\|_{\ell_2}=1, \mathbf{h}_{|\Omega^c(t)}=\mathbf{0}_{L-S} \ , \tag{11}$$

where $\Omega^c(t)$ is the complementary to $\Omega(t)$ set, at iteration t.

3. BLIND ADAPTIVE SPARSE CHANNEL IDENTIFICATION

We initially give a brief description of the adaptive subspace estimation of matrices $\mathbf{R}_n$ and $\mathbf{W}_n$ which are the estimates of matrices $\mathbf{R}$ and $\mathbf{W}$, respectively, based on the received samples up to time n. The direct computation of these subspace problems requires the SVD of the involved matrices. This is a very costly task to be implemented in an adaptive fashion. Adaptive subspace-based tracking schemes have been proposed in literature to reduce computational costs. In this paper, we adopt the approach of [4], which is based on the well-known Data Projection Method (DPM) algorithm [7], in order to estimate the noise subspace of matrix $\mathbf{R}_n$. According to the DPM algorithm, the noise subspace $\mathbf{U}_n$ of a matrix $\mathbf{A}_n$ can be tracked via the following updating relation

$$\mathbf{U}_n = orth\Big\{\mathbf{U}_{n-1} - \frac{\mu}{tr(\mathbf{A}_n)}\mathbf{A}_n\mathbf{U}_{n-1}\Big\} \tag{12}$$

where μ is a properly selected step-size parameter and $orth\{\cdot\}$ stands for an orthonormalization procedure (e.g., Gram-Schmidt). Thus, in the adaptive subspace tracking scenario, eq.(12) is firstly applied to compute the noise subspace of matrix $\mathbf{R}_n$ and then the first noise eigenvector of matrix $\mathbf{W}_n$ is computed by the same equation. Note, that the computational complexity of finding the noise subspace of $\mathbf{R}_n$ can be reduced further if $\mathbf{R}_n$ is approximated as $\hat{\mathbf{R}}_n = \mathbf{y}_n\mathbf{y}_n^H$ (i.e., as in stochastic gradient schemes).

In the following subsections we present two different approaches for identifying the non-negligible CIR taps by solving the subspace problem that involves matrix $\mathbf{W}_n$. The first approach is based on the ℓ_1-norm relaxation while the second one on suitable greedy optimization strategies.

3.1. ℓ_1 based sparse subspace tracking

By properly exploiting the results of [8] for the Least-Mean-Squares (LMS) algorithm, we have derived two ℓ_1-sparse subspace tracking algorithms, the so-called Zero-Attracting DPM (ZADPM) and Re-weighted Zero-Attracting DPM (RZADPM) which are briefly described below.

3.1.1. The Zero-Attracting DPM

Let us first define a new cost function $J_1(\mathbf{h}_n)$ as

$$J_1(\mathbf{h}_n) = \mathbf{h}_n^H \mathbf{W}_n \mathbf{h}_n + \gamma \|\mathbf{h_n}\|_{\ell_1}, \tag{13}$$

Thus, the constraint on the ℓ_1 norm of eq.(10) is embodied to the cost function and the equivalent problem is given by

$$\min_{\mathbf{h}_n} J_1(\mathbf{h}_n)\ s.t.\ \|\mathbf{h}_n\|_{\ell_2} = 1, \tag{14}$$

where parameter γ controls the solution's sparsity order. In order to derive a gradient search procedure for this minimization problem, we first need to compute the gradient of the $J_1(\mathbf{h}_n)$ cost function. It can be shown that the resulting updating step of the zero-attracting DPM (ZADPM) is given by

$$\mathbf{h}_n = orth\left\{\mathbf{h}_{n-1} - \frac{\mu}{tr(\mathbf{W}_n)}\mathbf{W}_n\mathbf{h}_{n-1} - \frac{\mu\gamma}{tr(\mathbf{W}_n)} sgn(\mathbf{h}_{n-1})\right\}, \tag{15}$$

where $sgn(.)$ is the component-wise sign function.

3.1.2. The Reweighted Zero-Attracting DPM algorithm

We may redefine the cost function of eq.(13) by including a log-sum penalty instead of the ℓ_1 one of ZADPM, i.e.,

$$J_2(\mathbf{h}_n) = \mathbf{h}_n^H \mathbf{W}_n \mathbf{h}_n + \gamma' \sum_{n=1}^{N} log\left(1 + \frac{\left|h_n^{(i)}\right|}{\epsilon}\right), \tag{16}$$

where parameters γ' and ϵ control the sparsity order of the solution. The above log-sum penalty behaves more similarly to the ℓ_0 norm that the ℓ_1 penalty of the ZADPM algorithm. We compute again the gradient and thus derive in a similar manner the update step of the RZADPM, which is given by

$$\mathbf{h}_n = orth\left\{\mathbf{h}_{n-1} - \frac{\mu}{tr(\mathbf{W}_n)}\mathbf{W}\mathbf{h}_{n-1} - \frac{\mu\gamma'}{tr(\mathbf{W}_n)} \frac{sgn(\mathbf{h}_{n-1})}{1+\epsilon|\mathbf{h}_n|}\right\} \tag{17}$$

3.2. Greedy Approaches for Sparse Subspace Tracking

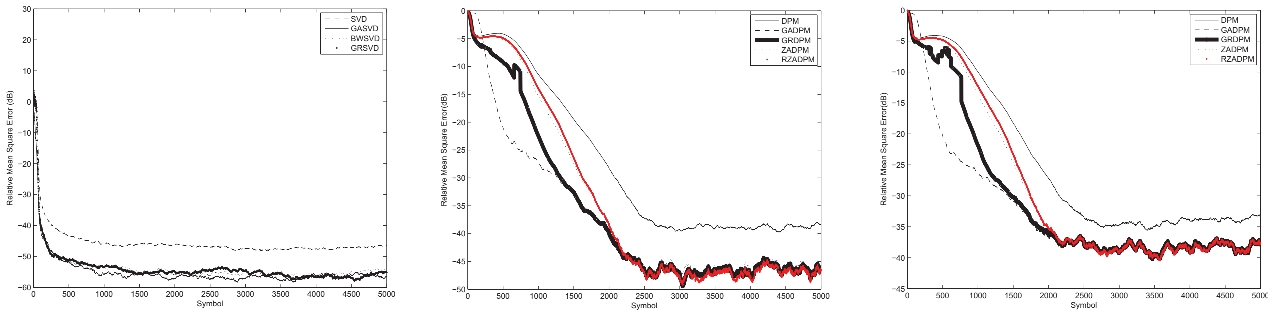
In this section a solution to the problem of eq.(9) is sought by using greedy methods. In [9] greedy search and branch-and-bound methods were proposed for maximizing a quadratic function of a matrix $\mathbf{A}$

$$\max_{\mathbf{h}} J_{\mathbf{A}}(\mathbf{h}) \quad s.\ t. \quad \|\mathbf{h}\|_{\ell_0} \leq \delta, \quad \|\mathbf{h}\|_{\ell_2} = 1\ . \tag{18}$$

It turns out that the above maximization problem is equivalent to the search of the submatrix of matrix $\mathbf{A}$ with the largest leading eigenvalue. The iterative suboptimal methods suggested in [9] compute, at iteration step n, the $(n-1)\times(n-1)$ (backward method) or the $(n+1)\times(n+1)$ (forward method) sub-matrix of the $n\times n$ input matrix $\mathbf{A}$, with the largest eigenvalue out of all candidate sub-matrices. By properly exploiting these results we may derive an adaptive greedy subspace tracking technique for the problem at hand. Unfortunately, the $O(L^4)$ complexity of the previous methods seems to be prohibitive for adaptive implementations. Moreover, the methods cannot be directly applied to our case, since we search for the sparsity pattern of the minimum eigenvector. A solution to this problem would be to apply these methods either on matrix $\mathbf{W}_n^{-1}$ or on $\mathbf{I} - \mu\mathbf{W}_n$. Such a selection is based on the observation that matrices $\mathbf{W}_n^{-1}, \mathbf{I} - \mu\mathbf{W}_n$ have singular values in reverse order with respect to those of $\mathbf{W}_n$. An alternative solution to the problem of (9) would be to estimate the minimum eigenvector of $\mathbf{W}_n$ via eq. (12) and form a sparse vector by setting to zero the $N - S$ coefficients with the smallest magnitude (i.e. $\hat{\mathbf{h}}_{|\Lambda_S}(n) = \mathbf{0}_{(N-S)\times 1}\}$ where $\Lambda_S = min\left(\left|\hat{\mathbf{h}}(n)\right|, N - S\right)$). Surprisingly, it seems that, for the proposed adaptive greedy subspace algorithms, such a simple thresholding method is sufficient. In the simulations' section it is clearly shown that the performance of the backward search strategy coincides with that of simple thresholding for the case of our interest.

4. SIMULATIONS

The performance of the derived techniques for a CIR of length $L + 1 = 25$ was tested in a stationary and in a slow-fading environment. The non-zero channel taps were placed at positions $\Omega = \{0, 10, 15, 19, 24\}$ for both channels. The stationary channel's taps were generated as i.i.d. complex Gaussian variables of zero mean and variance equal to 1. The fading channel taps were generated by $h_i(n) = J_0(2\pi \cdot 10^{-2.5})h_i(n-1) + \sqrt{1-|a|^2}\zeta(n)$, where J_0 is the zero order Bessel function of the first kind and $\zeta(n)$ is a white noise random variable of unit variance. In order to get a variation in the sparsity pattern, we force the last two non-zero taps to slowly fade towards zero in the fading channel's case. The system used $N = 64$ sub-carriers, the received data were BPSK symbols corrupted by i.i.d. complex Gaussian noise variables such that the Signal-to-Noise (SNR) to be equal to 20dB. It should be mentioned that in the simulation results given below we used the approach of [4] to resolve the scalar ambiguity of (8) by using 4 pilot symbols placed on the positions $\{0, 16, 32, 64\}$ of the OFDM block symbol. Even though the scheme uses these few pilots, a comparison with a training-based scheme is meaningless, since the performance of the latter depends linearly on the number of the used pilots symbols. The SVD-based schemes



(a) *SVD-based methods* (b) *Adaptive methods-Stationary Channel* (c) *Adaptive methods-Fading Channel*

Fig. 1. *Performance of the techniques at SNR=20dB*

track the correlation matrix by employing an exponential window update with a forgetting factor equal to 0.998. The step size variable is set to $\mu = 1$ for all the DPM-based techniques and the parameters of the ZADPM and RZADPM algorithms are set to $\gamma = 10^{-4}$, $\gamma' = 10^{-2}$ and $\epsilon = 10$, respectively. The relative mean square error of each algorithm is depicted after averaging the results of 100 independent runs. In Fig.1(a) the performance of the SVD-based schemes is depicted. For these schemes, the required noise subspaces are computed by directly computing the SVD of the involved matrices. More specifically we plot the performance of: the simple SVD ("SVD"), the genie-aided SVD ("GASVD"), the greedy-based SVDs using the thresholding approach ("GRSVD") and the backward method ("BWSVD"). Note that in "GASVD" it is assumed that the sparsity pattern is known and the required eigenvectors are computed from the direct SVD of the sub-matrix that consists of the Ω rows and columns of $\mathbf{W}$. The sparse schemes exhibit significant performance improvement compared to the non-sparse ones. It is also noteworthy that the two greedy techniques ("GRSVD" and "BWSVD") attain similar performance, thus justifying our choice to use the first one in the corresponding adaptive technique.

In Fig.2(b)-(c) the performance of the adaptive sparse DPM-based techniques is presented for a stationary and a slow-fading environment respectively. The $\ell_1 - based$ approaches are denoted as "ZADPM" and "RZADPM", respectively, and the greedy approach that employs the thresholding technique is denoted as "GRDPM". For comparison purposes we also plot in the same figure the DPM and the genie-aided DPM ("GADPM") techniques. We observe that the sparse approaches exhibit faster convergence and significantly better steady-state performance in both channel cases.

5. CONCLUSION

In this paper blind methods were proposed for sparse channel estimation. Moreover, adaptive sparse subspace tracking techniques were proposed so as to provide efficient on-line implementations of the proposed estimation schemes. The presented techniques were applied in a ZP-OFDM system and it was shown by typical simulations that they achieve significantly improved performance compared to known non-sparse approaches. Unfortunately due to space limitations, we were unable to provide some initial theoretical results concerning the convergence analysis and the identifiability properties of the proposed schemes. The above issues are still under study and the derived results will be presented in a next paper.

6. ACKNOWLEDGMENT

This work was supported by EU and national funds via the National Strategic Reference Framework (NSRF) - Research Funding Program: Heracleitus II and also by University of Patras.

7. REFERENCES

[1] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Transactions on Information Theory*, vol. 55, no. 5, pp. 2230–2249, 2009.

[2] J. Heath, R.W. and G. Giannakis, "Exploiting input cyclostationarity for blind channel identification in OFDM systems," *IEEE Transactions on Signal Processing*, vol. 47, no. 3, pp. 848 –856, mar 1999.

[3] L. Tong, G. Xu, and T. Kailath, "A new approach to blind identification and equalization of multipath channels," in *Asilomar Conference on Signals, Systems and Computers*, Nov. 1991, pp. 856 –860 vol.2.

[4] X. Doukopoulos and G. Moustakides, "Blind adaptive channel estimation in OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 5, no. 7, pp. 1716 –1725, July 2006.

[5] M. Raghavendra and K. Giridhar, "Improving channel estimation in OFDM systems for sparse multipath channels," *IEEE Signal Processing Letters*, vol. 12, no. 1, pp. 52 – 55, Jan. 2005.

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

[7] J.-F. Yang and M. Kaveh, "Adaptive eigensubspace algorithms for direction or frequency estimation and tracking," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 36, pp. 241 –251, Feb. 1988.

[8] C. Yilun, Y. Gu, and A. O. Hero, "Sparse LMS for system identification," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009, pp. 3125–3128.

[9] B. Moghaddam, Y. Weiss, and S. Avidan, "Spectral bounds for sparse pca: Exact and greedy algorithms," in *Advances in Neural Information Processing Systems*. MIT Press, 2006, pp. 915–922.