INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Blind Adaptive Channel Estimation in OFDM Systems

Xenofon G. Doukopoulos — George V. Moustakides

N° 5126

February 2004

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--5126--FR+ENG

Blind Adaptive Channel Estimation in OFDM Systems

Xenofon G. Doukopoulos — George V. Moustakides

Rapport de recherche n° 5126 — February 2004 — 21 pages

Abstract: We consider the problem of blind channel estimation in zero padding OFDM systems. For the first time, blind adaptive algorithms are proposed that identify the impulse response of the multipath channel. In particular, we develop RLS and LMS schemes that exhibit rapid convergence combined with low computational complexity and numerical stability. Both versions are obtained by properly modifying the orthogonal iteration method used in Numerical Analysis for the computation of singular vectors. With a number of simulation experiments we demonstrate the satisfactory performance of our adaptive schemes under diverse signaling conditions.

Key-words: OFDM, Blind channel estimation, Adaptive algorithms, Power method, Orthogonal Iteration, Zero padding.

Unité de recherche INRIA Rennes
IRISA, Campus universitaire de Beaulieu, 35042 Rennes Cedex (France)
Téléphone : +33 2 99 84 71 00 — Télécopie : +33 2 99 84 71 71

Estimation Adaptative Aveugle de Canal dans des Systèmes OFDM

Résumé : On considère le problème de l'estimation aveugle adaptative d'un canal pour le système ZP-OFDM. Pour la première fois des algorithmes adaptatifs aveugles sont proposés, qui identifient la réponse impulsionelle d'un canal à trajets multiples. En particulier, on dévelope des schémas RLS et LMS qui présentent une convergence rapide, ainsi qu'une faible complexité de calcul et une stabilité numérique. Les deux versions sont obtenues en modifiant la méthode d'itération orthogonale utilisée dans l'analyse numérique pour calculer des vecteurs singuliers. À l'aide de simulations nous démontrons la performance satisfaisante de nos deux schémas avec des canaux divers.

Mots-clés : OFDM, Estimation aveugle de canal, Algorithmes adaptatifs, Méthode de la puissance, Itération orthogonale, Remplissage de zéros.

Contents

1 Introduction

Orthogonal Frequency Division Multiplexing (OFDM) constitutes a promising technology for high speed transmission in frequency selective fading environment [1]. OFDM presents several important advantages, some of which are: high spectral efficiency; simple implementation (with IDFT/DFT pairs); mitigation of intersymbol interference (ISI) and robustness to frequency selective fading environments. Inevitably, these desirable characteristics contribute towards a continuously rising interest for OFDM. We should mention that OFDM has been chosen for the European standard of digital audio and video broadcasting (DAB, DVB), digital subscribe line modems (DSL), and wireless local area networks (LANs).

In practice OFDM systems operate over a dispersive channel and therefore a guard interval, no smaller than the anticipated channel spread, is inserted in the transmitted sequence. As far as this guard period is concerned, two alternative schemes have been proposed. The first, known as cyclic prefix (CP), consists in re-transmitting inside the guard interval the initial portion of the transmitted sequence; whereas the second, known as zero padding (ZP), as it is evident from its name, transmits no information during the same interval.

In this work we mainly focus on the ZP-OFDM model. The ZP approach is very appealing [2] and has started gaining popularity mainly because of its simplicity. Its strongest point consists in the *complete elimination* of the inter-block interference (IBI) which allows for a number of interesting detection structures. A detailed comparison between CP-OFDM and ZP-OFDM receivers and several other merits of the ZP model are offered in [2].

In coherent detection and adaptive loading, knowledge of the channel impulse response is imperative. Since the channel impulse response is usually unknown to the receiver, it needs to be efficiently estimated. Channel estimation techniques can be divided into two major categories the *supervised* or *trained* and the *unsupervised* or *blind*. The first requires training/pilot sequences whereas the latter uses only the received data. Due of course to their self-sufficiency in training, blind techniques are considered more attractive than their trained counterparts; they tend however to be heavier from a computational complexity point of view. As far as adaptive implementations is concerned, although one can find numerous trained methods in the literature, this is not the case for blind approaches. Existing blind OFDM channel identification methods are mainly off-line.

The majority of articles dealing with the problem of supervised channel estimation in OFDM systems, uses pilot tones or training sequences [3, 4]. In [5] a comparative study of non-blind methods can be found. The pilot-aided literature is rich, however, since our main interest lies with blind methods we will not pursue its presentation any further. Regarding blind techniques, in [6] channel identification is performed by exploiting the cyclostationarity present in CP-OFDM. In [7] a subspace approach is proposed that takes advantage of the redundancy existing in CP-OFDM. An alternative subspace approach is presented in [8], which extends the previous idea by incorporating virtual carriers inside the OFDM transmitted block. The two latter methods require singular value decomposition (SVD) of the received data autocorrelation matrix and are therefore characterized by high computational cost. It is also known that SVD lacks a repetitive structure that could lead to efficient adaptive implementations and is therefore unsuitable for on-line processing.

In this work we exploit the subspace method in order to develop *adaptive* algorithms for blind channel identification in ZP-OFDM systems. To our knowledge, this is the first time such schemes are proposed for OFDM systems. Specifically we are going to develop RLS and LMS type algorithms that can solve, very efficiently, the blind channel estimation problem. Both versions have significantly lower computational complexity as compared to the direct SVD approaches of [7, 8]. In particular, our LMS version is extremely simple with a computational complexity that is almost two orders of magnitude smaller than a direct SVD approach.

We would also like to stress that our LMS adaptive scheme is based on a novel *adaptive subspace tracking* algorithm introduced here for the first time. Although there exists an abundance of such techniques in the literature, they are mostly focused on estimating the signal subspace and not the noise subspace required here. The algorithm we are going to introduce relies on the *orthogonal iteration method* [9] and is characterized by extreme simplicity and *numerical stability*. The latter characteristic is not enjoyed by other existing algorithms of similar computational complexity.

The rest of the paper is organized as follows. In Section II we introduce the signal model for a ZP-OFDM system. We continue in Section III with the definition of two subspace problems that constitute the heart of the blind channel estimation methodology. Section IV contains the orthogonal iteration and two of its variants that are suitably tuned for the solution of the two subspace problems introduced in Section III. Furthermore in the same section we give adaptive implementations of the two orthogonal iteration variants, which are used in Section V to develop blind adaptive RLS and LMS algorithms for the identification of the channel impulse response. In Section V we also consider the phase and amplitude ambiguity problem, encountered in all blind techniques, and propose a simple remedy for its resolution. Simulation results are offered in Section VI and finally Section VII contains our concluding remarks.

2 System Model

OFDM modulation has the characteristic of multiplexing data symbols over a large number of orthogonal carriers. Consider an OFDM system where the guard interval consists of a zero padded sequence. Fig. 1 depicts the baseband discrete-time block equivalent model of a standard ZP-OFDM transmitter. Let each information block be comprised

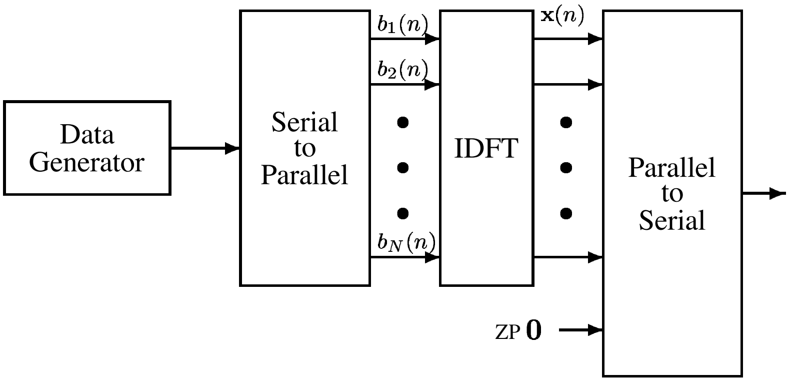

Figure 1: Discrete time block ZP-OFDM transmitter.

of N symbols and denote by L the length of the ZP. The n-th length-N symbol block $\mathbf{b}(n) = [b_1(n) \; \ldots \; b_N(n)]^t$ passes through a serial to parallel converter and is then being modulated by IDFT. Next, a sequence of L zeros (zero padding) is inserted between two consecutive blocks to form the transmitted vector $\mathbf{x}(n)$. The latter is of length $N + L$, and can be put under the following form

$$\mathbf{x}(n) = \begin{bmatrix} \mathbf{F}^H \\ \mathbf{0}_{L \times N} \end{bmatrix} \mathbf{b}(n), \tag{1}$$

where $\mathbf{F}$ stands for the DFT matrix

$$\mathbf{F} = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & W_N & W_N^2 & \cdots & W_N^{N-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & W_N^{N-1} & W_N^{2(N-1)} & \cdots & W_N^{(N-1)(N-1)} \end{bmatrix}, \tag{2}$$

with $W_N = e^{-j\frac{2\pi}{N}}$; superscript "^{H}" denotes conjugate-transpose and $\mathbf{0}_{L\times N}$ is a zero matrix of dimensions $L \times N$. The parallel block $\mathbf{x}(n)$ is finally transformed into a serial sequence in order to be transmitted through the channel.

The transmitted signal propagates through a multipath additive white noise (AWN), *not necessarily Gaussian*, channel with impulse response $\mathbf{h} = [h_0 \ \dots \ h_L]^t$. Here we have assumed that the channel has a finite impulse response of length at most $L+1$ not exceeding the ZP length (plus one). Such an assumption is very common in OFDM systems and constitutes the main reason for introducing the guard interval in the great majority of OFDM models.

Whenever ZP is employed, after assuming synchronization with the transmitted sequence, the n-th received data block $\mathbf{y}(n)$ of length $N+L$ can be expressed as

$$\mathbf{y}(n) = \mathbf{H}\mathbf{F}^H\mathbf{b}(n) + \mathbf{w}(n). \tag{3}$$

In the above relation $\mathbf{H}$ is a convolution matrix of dimensions $(N+L) \times N$ defined as

$$\mathbf{H} = \begin{bmatrix} h_0 & 0 & \cdots & 0 \\ \vdots & h_0 & \ddots & \vdots \\ h_L & \vdots & \ddots & 0 \\ 0 & h_L & \ddots & h_0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_L \end{bmatrix}; \tag{4}$$

where we recall that $\mathbf{b}(n)$ is the n-th block of transmitted symbols and $\mathbf{w}(n)$ is an AWN vector of length $N+L$ with zero-mean independent and identically distributed (i.i.d.) elements that are also independent of the transmitted symbols. From (3) it is possible to verify the very interesting property of the ZP-OFDM model stated in the Introduction, namely its ability to completely eliminate the IBI between consecutive blocks. This is evident from the fact that the received data block $\mathbf{y}(n)$ depends only on $\mathbf{b}(n)$ and not on any other previous or next symbol block. A similar property in order to be enjoyed by the CP-OFDM model, it is necessary, in each received data block of size $N+L$, to *discard* the first L data samples, thus throwing away information that could be useful.

3 Main Idea

In this section we will attempt to solve the channel identification problem after assuming that the received data autocorrelation matrix is available. As it is almost always the case with subspace techniques, the key idea consists in properly decomposing the data into the signal and noise subspace and then defining suitable subspace determination problems that will lead to the final estimate of the channel impulse response.

3.1 A Subspace Approach

Consider the autocorrelation matrix $\mathbf{R}$ of the received data vector $\mathbf{y}(n)$ defined in (3). Assuming that the elements of $\mathbf{b}(n)$ are i.i.d. and of unit norm, using the fact that the DFT matrix $\mathbf{F}$ defined in (2) is orthonormal, we conclude that

$$\mathbf{R} \stackrel{\triangle}{=} \mathbb{E}\{\mathbf{y}(n)\mathbf{y}^H(n)\} = \mathbf{H}\mathbf{H}^H + \sigma^2\mathbf{I}_{N+L}, \tag{5}$$

where σ^2 is the noise power and $\mathbf{I}_K$ denotes the identity matrix of size K. The matrix $\mathbf{H}\mathbf{H}^H$ is Hermitian and nonnegative definite, of dimensions $(N+L)\times(N+L)$. From (3) it is clear that the signal subspace is formed by linearly combining the columns of $\mathbf{H}$; therefore these columns belong to the signal subspace (in fact they span it). Assuming that the channel impulse response $\mathbf{h}$ is not identically zero, since $\mathbf{H}$ is a convolution matrix, it is also of full column rank. This suggests that the signal subspace has rank equal to N and therefore its complement, the noise subspace, rank equal to L.

Taking into account the previous observation, if we apply an SVD on $\mathbf{R}$ we can then write

$$\mathbf{R} = [\mathbf{U}_s \quad \mathbf{U}_w] \begin{bmatrix} \mathbf{\Lambda}_s + \sigma^2\mathbf{I}_N & \mathbf{0} \\ \mathbf{0} & \sigma^2\mathbf{I}_L \end{bmatrix} [\mathbf{U}_s \quad \mathbf{U}_w]^H, \tag{6}$$

where $\mathbf{U}_s, \mathbf{U}_w$ are orthonormal bases for the signal and noise subspace respectively and $\mathbf{\Lambda}_s$ is a diagonal matrix of size N, with positive elements. It is important to point out that $\mathbf{U}_w$ involves the singular vectors of the matrix $\mathbf{R}$ that correspond to its *smallest* singular value (which is equal to σ^2).

Since $\mathbf{U}_w^H\mathbf{U}_s = \mathbf{0}$ and $\mathbf{U}_w$, $\mathbf{U}_s$ are bases for the signal and noise subspace respectively, then any vector in the noise subspace will be orthogonal to any other vector in the signal subspace. Notice that the columns of $\mathbf{H}$ are vectors in the signal subspace, therefore for any vector $\mathbf{v} = [v_1 \ \cdots \ v_{N+L}]^t$ of length $N+L$ in the noise subspace, we have $\mathbf{v}^H\mathbf{H} = 0$. Because of the Toeplitz form of $\mathbf{H}$, depicted in (4), the vector-matrix product $\mathbf{v}^H\mathbf{H}$ can also be written as

$$\mathbf{v}^H\mathbf{H} = \mathbf{h}^t\mathbf{V}^\star = 0, \tag{7}$$

where superscript "$\star$" denotes complex conjugate and $\mathbf{V}$ is a Hankel matrix of dimensions $(L+1) \times N$, made up from the elements of the vector $\mathbf{v}$ as follows

$$\mathbf{V} = \begin{bmatrix} v_1 & v_2 & \cdots & v_N \\ v_2 & v_3 & \cdots & v_{N+1} \\ \vdots & \vdots & & \vdots \\ v_{L+1} & v_{L+2} & \cdots & v_{N+L} \end{bmatrix}. \tag{8}$$

By taking the complex conjugate of the relation in (7) we conclude

$$\mathbf{h}^H\mathbf{V} = 0 = \mathbf{h}^H\mathbf{V}\mathbf{V}^H\mathbf{h}. \tag{9}$$

Since (9) holds for every vector $\mathbf{v}$ in the noise subspace, if $\mathbf{v}_1, \ldots, \mathbf{v}_L$ is a collection of L such vectors, we also have

$$\mathbf{h}^H\mathbf{W}\mathbf{h} = 0, \tag{10}$$

with

$$\mathbf{W} = \sum_{i=1}^{L} \mathbf{V}_i\mathbf{V}_i^H \tag{11}$$

where $\mathbf{V}_i,\ i = 1, \ldots, L$ are the corresponding Hankel matrix versions of the vectors $\mathbf{v}_i,\ i = 1, \ldots, L$, formed according to (8).

Relation (10) constitutes the key equation for recovering the impulse response. Indeed, from (10), we have that $\mathbf{h}$ is the singular vector corresponding to the zero, and therefore the *smallest*, singular value of $\mathbf{W}$. The first step, of course, in estimating $\mathbf{h}$ through the subspace problem defined in (10) is the formation of the matrix $\mathbf{W}$. As we can see from (11), this is possible if we have available a collection of L vectors $\mathbf{v}_i,\ i = 1, \ldots, L$, that lie in the noise subspace. Ideal candidates for these vectors constitute the columns of the matrix $\mathbf{U}_w$ since they span the whole noise subspace and contain no redundant information (because of orthonormality).

If we estimate $\mathbf{h}$ as the singular vector corresponding to the smallest singular value of $\mathbf{W}$ then this will introduce an *amplitude and phase ambiguity* (which is always present in blind techniques). This is because if $\mathbf{h}$ satisfies (10) so does $\alpha\mathbf{h}$ where α any complex number. In order to limit the ambiguity let us consider a normalized version $\mathbf{h}_b$ of the channel impulse response that satisfies $||\mathbf{h}_b|| = 1$. Then the true channel impulse response is related to $\mathbf{h}_b$ through

$$\mathbf{h} = \alpha\mathbf{h}_b, \tag{12}$$

where α a complex scalar. Blind techniques are basically capable of providing estimates for $\mathbf{h}_b$. To be able to resolve the ambiguity due to α, as we are going to see in Section V, it will be necessary to introduce some additional information besides the received data signal vectors $\{\mathbf{y}(n)\}$.

From the preceding discussion we conclude that, in order to estimate the (normalized) channel impulse response we need to solve the following two subspace identification problems:

Subspace Problem 1: The first step in identifying $\mathbf{h}_b$ is the determination of a noise subspace basis $\mathbf{U}_w$. This matrix is of size $(N+L) \times L$ and its columns are singular vectors corresponding to the smallest singular value of the received data autocorrelation matrix $\mathbf{R}$.

Subspace Problem 2: The L columns of the matrix $\mathbf{U}_w$ obtained from the first subspace problem, constitute the collection of L vectors $\mathbf{v}_i$ required to form the matrix $\mathbf{W}$ using (8) and (11). Once $\mathbf{W}$ is computed, the normalized channel impulse response $\mathbf{h}_b$ can be obtained as the singular vector corresponding to the smallest singular value of $\mathbf{W}$.

Both problems involve the determination of subspaces corresponding to the smallest singular value of a matrix. As we can see the proposed method is based solely on the received data process $\{\mathbf{y}(n)\}$, therefore it is clearly blind. Although similar methodology has been developed for channel estimation in CDMA there exists a major difference that distinguishes the current setting from the one used in CDMA. Here we know exactly the noise subspace rank while this is not the case in CDMA where this parameter is variable, depending on the number of users in the channel [10, 11]. Due to this extra knowledge it will be possible to develop algorithms for OFDM that are more powerful than their CDMA counterparts.

3.2 Consistency

Let us now briefly discuss the problem of consistency of the proposed method. We have the following theorem that treats this issue.

Theorem 1 *Let $\mathbf{v}_i,\ i=1,\ldots,L$, be the L columns of a basis $\mathbf{U}_w$ of the noise subspace, with $\mathbf{V}_i$ their corresponding Hankel versions and define $\mathbf{W}$ according to (11). Then the channel impulse response $\mathbf{h}$ is the unique vector (modulo a multiplicative complex scalar ambiguity) that satisfies Equ.(10).*

Proof: The proof is presented in the Appendix. From the proof we can also conclude that consistency is possible even if we use a *single* vector $\mathbf{v}_i$ to form $\mathbf{W}$, provided that the corresponding Hankel matrix $\mathbf{V}_i$ is of full row rank (property that holds with probability one). Although, theoretically, using all vectors $\mathbf{v}_i$ does not contribute to the consistency, it does however ameliorate (considerably) the convergence properties of the adaptive schemes we are going to present in the sequel.

4 Orthogonal Iteration and Variants

The orthogonal iteration [9], is a simple iterative technique that can be used to compute the singular vectors corresponding to the L *largest* singular values of a symmetric nonnegative definite matrix. Let us summarize the method in the following lemma.

Lemma 1 *Consider a symmetric, positive definite matrix* $\mathbf{Q}$ *of size* K *and let* $s_1 \geq \ldots \geq s_L > s_{L+1} \geq \ldots \geq s_K > 0$ *be its singular values and* $\mathbf{f}_1, \ldots, \mathbf{f}_K$ *the corresponding singular vectors. Consider the sequence of matrices* $\{\mathbf{Z}(k)\}$ *of dimensions* $K \times L$*, defined by the iteration*

$$\mathbf{Z}(k) = \text{orthonorm}\{\mathbf{Q}\mathbf{Z}(k-1)\}, \ \ k = 1, 2, \ldots \tag{13}$$

where "orthonorm" stands for orthonormalization using QR decomposition, then

$$\lim_{k \to \infty} \mathbf{Z}(k) = [\mathbf{f}_1 \ \cdots \ \mathbf{f}_L], \tag{14}$$

provided that the matrix $\mathbf{Z}^H(0)[\mathbf{f}_1 \cdots \mathbf{f}_L]$ *is not singular.*

Proof: The proof can be found in [9, Page 354].

A number of remarks are necessary at this point.

Remark 1: If certain of the L largest singular values coincide, then the singular vectors corresponding to the multiple singular values are not unique. In this case the orthogonal iteration converges to a basis in the corresponding subspace.

Remark 2: For the orthogonal iteration to converge, it is imperative that $s_L > s_{L+1}$. In fact one can show [9, Page 354] that the convergence is exponential with rate s_{L+1}/s_L.

Remark 3: If instead of QR we use any other orthonormalization procedure, the sequence $\{\mathbf{Z}(k)\}$ converges to an orthonormal basis in the space spanned by the first L singular vectors. The latter is unimportant in the case where the L largest singular values are all equal (since in this case the singular vectors are not unique).

Remark 4: If $L = 1$ then the orthonormalization process is reduced to a simple vector normalization and the corresponding method is known as *power method* [9].

As we have seen in the previous subsection, in both subspace problems the goal is to find the subspace corresponding to the *smallest* singular value. There are two interesting variants of the orthogonal iteration that can provide such estimates. We present them in the form of a lemma.

Lemma 2 *Let $\mathbf{Q}$ be a symmetric positive definite matrix of size K, with singular values $s_1 \geq s_2 \geq \cdots \geq s_{K-L} > s_{K-L+1} \geq \cdots \geq s_K > 0$ and $\mathbf{f}_i,\ i = 1, \ldots, K$ the corresponding singular vectors. If the sequence $\{\mathbf{Z}(k)\}$ of matrices of dimensions $K \times L$ is defined by either of the two iterations*

$$\mathbf{Z}(k) = \text{orthonorm}\{\mathbf{Q}^{-1}\mathbf{Z}(k-1)\},\ k = 1, 2, \ldots \tag{15}$$

$$\mathbf{Z}(k) = \text{orthonorm}\left\{(\mathbf{I} - \mu\mathbf{Q})\,\mathbf{Z}(k-1)\right\},\ k = 1, 2, \ldots \tag{16}$$

where $0 < \mu < 1/s_1$ and $\mathbf{I}$ the identity matrix, then

$$\lim_{k\to\infty} \mathbf{Z}(k) = [\mathbf{f}_{K-L+1}\ \cdots\ \mathbf{f}_K], \tag{17}$$

provided that the matrix $\mathbf{Z}^H(0)[\mathbf{f}_{K-L+1} \cdots \mathbf{f}_K]$ is not singular.

Proof: The proof is an immediate application of Lemma 1 and the fact that the matrices $\mathbf{Q}^{-1}$, $\mathbf{I} - \mu\mathbf{Q}$ have singular values $\frac{1}{s_i}$ and $1 - \mu s_i,\ i = 1, \ldots, K$ respectively and exactly the same singular vectors as the matrix $\mathbf{Q}$. In other words $\mathbf{Q}^{-1}$ and $\mathbf{I} - \mu\mathbf{Q}$ constitute two possible ways to map the smallest singular values into the largest ones, without altering the corresponding subspaces, and then apply the orthogonal iteration. The constraint $0 < \mu < 1/s_1$ is required in order for the matrix $\mathbf{I} - \mu\mathbf{Q}$ to be positive definite.

4.1 Adaptive Implementations

We are now interested in the application of the orthogonal iteration, and in particular of its two variants introduced in Lemma 2, under an adaptive setting. Let us therefore assume that $\mathbf{Q}$ is no longer available; instead we have a sequence of random matrices $\{\mathbf{Q}(n)\}$ with expectation equal to $\mathbf{Q}$, that is, $\mathbb{E}\{\mathbf{Q}(n)\} = \mathbf{Q}$. We distinguish two cases.

Case A: $\mathbb{E}\{\|\mathbf{Q}(n) - \mathbf{Q}\|^2\} \ll \|\mathbf{Q}\|^2$. In this case the random matrices $\{\mathbf{Q}(n)\}$ constitute *efficient* estimates of the matrix $\mathbf{Q}$ since the error power is considered significantly smaller than the power of $\mathbf{Q}$. Here we can apply both iterations (15), (16) modified as follows

$$\mathbf{Z}(n) = \text{orthonorm}\{\mathbf{Q}^{-1}(n)\mathbf{Z}(n-1)\}, \tag{18}$$

$$\mathbf{Z}(n) = \text{orthonorm}\{[\mathbf{I} - \mu\mathbf{Q}(n)]\,\mathbf{Z}(n-1)\}, \tag{19}$$

where $0 < \mu < 1/s_1$.

Case B: $\mathbb{E}\{\|\mathbf{Q}(n) - \mathbf{Q}\|^2\} \sim \|\mathbf{Q}\|^2$. In this case the random matrices $\{\mathbf{Q}(n)\}$ constitute *crude* estimates of the matrix $\mathbf{Q}$ because the error power is comparable to the power of the matrix $\mathbf{Q}$. Here we can apply only (16) as it is modified in (19), but with $0 < \mu \ll 1/s_1$.

The reason we distinguish the two cases is because we will propose an adaptive algorithm based on RLS that produces efficient estimates of the matrix we would like to decompose (Case A); and an alternative algorithm of gradient (LMS) type that approximates the desired matrix by instantaneous rank-one vector outer products (Case B). The former will have an excellent performance but at an increased computational cost, whereas the later a slightly inferior performance but with a very interesting computational complexity.

Unfortunately a formal proof of the stochastic convergence capabilities of the adaptive algorithms proposed in (18), (19), requires considerable space and effort (see for example [12]) and is therefore avoided. Instead we are going to give an intuitive explanation as to why these adaptations can work. Case A is rather clear. Indeed if $\mathbf{Q}(n) = \mathbf{Q} + \mathbf{E}(n)$ where $\mathbf{E}(n)$ are small random perturbations, then we can write $\mathbf{Q}^{-1}(n) = \mathbf{Q}^{-1} + \mathbf{E}'(n)$ and $\mathbf{I} - \mu\mathbf{Q}(n) = \mathbf{I} - \mu\mathbf{Q} + \mathbf{E}''(n)$ where $\mathbf{E}'(n)$ and $\mathbf{E}''(n)$ are both small random perturbation matrices. These small perturbations will in turn produce small random perturbations in the adaptation (18) or (19) thus yielding efficient singular vector estimates.

In Case B, on the other hand, the initial perturbations $\mathbf{E}(n)$ are considered important, therefore $\mathbf{E}'(n)$ and $\mathbf{E}''(n)$ will be important as well. This in turn will result in "noisy" singular vector estimates in (18) or in (19) when μ is not small. Therefore both adaptations should be avoided. When however in (19) we select a small step size, then stochastic averaging effects take place and one can show (see [12]) that the mean trajectory of (19) satisfies

$$\begin{aligned} \mathbb{E}\{\mathbf{Z}(n)\} &\approx \text{orthonorm}\{(\mathbf{I} - \mu\mathbb{E}\{\mathbf{Q}(n)\})\,\mathbb{E}\{\mathbf{Z}(n-1)\}\} \\ &\approx \text{orthonorm}\{(\mathbf{I} - \mu\mathbf{Q})\,\mathbb{E}\{\mathbf{Z}(n-1)\}\} \end{aligned} \tag{20}$$

which is the variant in (16). This means that the mean trajectory will converge to the desired singular vectors. Furthermore, at steady state, the estimation error power, as it is always the case in adaptive algorithms with small step size, will be of the order of μ and therefore small. In other words, when $0 < \mu \ll 1/s_1$, (19) will provide efficient estimates of the singular vectors.

5 Blind Adaptive Channel Estimation

In this section our goal is to develop adaptive solutions for the two subspace problems introduced in Section III. We recall that a straightforward (non-adaptive) solution consists in applying an SVD in both problems. The *direct SVD* technique is unfortunately characterized by an excessively high computational cost which is of the order of $O((N+L)^3)$ and is therefore considered unsuitable for on-line implementations. Let us now see how we can use the material presented in the previous section in order to obtain computationally efficient blind adaptive methods.

5.1 Adaptive Solutions for the First Subspace Problem

We are given sequentially the data blocks $\mathbf{y}(n)$ of length $N+L$ and we are interested in estimating a matrix $\mathbf{U}_w$ of size $(N+L) \times L$, containing L singular vectors of the noise subspace. Depending on the estimates we use for the data autocorrelation matrix $\mathbf{R}$ we can obtain alternative adaptations. There exist two interesting choices that we present next.

RLS Adaptation. Let $\mathbf{R}(n)$ be the exponentially windowed sample data autocorrelation matrix defined recursively as follows

$$\mathbf{R}(n) = \lambda\mathbf{R}(n-1) + (1-\lambda)\mathbf{y}(n)\mathbf{y}^H(n), \tag{21}$$

where $0 < \lambda < 1$ a forgetting factor. This case corresponds to an efficient estimate of $\mathbf{R}$ when λ is close to 1, since $\mathbb{E}\{||\mathbf{R}(n) - \mathbf{R}||^2\} = O(1-\lambda)$. Consequently we can apply (18) that involves the inverse matrix $\mathbf{P}(n) = \mathbf{R}^{-1}(n)$. It is well known that RLS computes directly $\mathbf{P}(n)$ with a computational complexity $O((N+L)^2)$. To present the complete adaptation let us assume that at time $n-1$ we have available the inverse $\mathbf{P}(n-1)$ of the data sample autocorrelation matrix and an estimate $\mathbf{U}_w(n-1)$ of the noise subspace basis.

When the new data block $\mathbf{y}(n)$ is available we apply

$$\mathbf{K}(n) = \mathbf{P}(n-1)\mathbf{y}(n) \tag{22}$$

$$\gamma(n) = \frac{1-\lambda}{\lambda + (1-\lambda)\mathbf{y}^H(n)\mathbf{K}(n)} \tag{23}$$

$$\mathbf{P}(n) = \frac{1}{\lambda}\left(\mathbf{P}(n-1) - \gamma(n)\mathbf{K}(n)\mathbf{K}^H(n)\right) \tag{24}$$

$$\mathbf{U}_w(n) = \text{orthonorm}\{\mathbf{P}(n)\mathbf{U}_w(n-1)\}. \tag{25}$$

In the first three equations we recognize the RLS algorithm, while in the last the variant of the orthogonal adaptation proposed in (18). The computational complexity of the scheme is $O((N+L)^2)$ for RLS; $O((N+L)^2L)$ to form the product $\mathbf{P}(n)\mathbf{U}_w(n-1)$ and $O((N+L)L^2)$ for the orthonormalization part (see [13]). Thus the leading complexity is $O((N+L)^2L)$ which is almost an order of magnitude smaller than the complexity of the direct SVD approach.

LMS Adaptation. Here we propose a crude estimate for $\mathbf{R}$, namely $\mathbf{R}(n) = \mathbf{y}(n)\mathbf{y}^H(n)$; we therefore need to apply (19) with a small step size μ. Since the size of μ is relative to the largest singular value s_1 of the matrix $\mathbf{R}$, we propose the use of a normalized step size of the form $\mu = \bar{\mu}/\text{trace}\{\mathbf{R}(n)\}$. We know that $\text{trace}\{\mathbf{R}\} \geq s_1$, however most of the time we have $\text{trace}\{\mathbf{R}\} \gg s_1$, therefore selecting $\bar{\mu}$ even close to unity results in $\mu \ll 1/s_1$. Since here $\text{trace}\{\mathbf{R}(n)\} = \|\mathbf{y}(n)\|^2$, the corresponding algorithm takes the following form

$$\mathbf{s}(n) = \mathbf{U}_w^H(n-1)\mathbf{y}(n) \tag{26}$$

$$\mathbf{T}(n) = \mathbf{U}_w(n-1) - \frac{\bar{\mu}}{\|\mathbf{y}(n)\|^2}\mathbf{y}(n)\mathbf{s}^H(n) \tag{27}$$

$$\mathbf{U}_w(n) = \text{orthonorm}\{\mathbf{T}(n)\}. \tag{28}$$

The first two relations have computational complexity $O((N+L)L)$ and the last, as in the RLS algorithm, $O((N+L)L^2)$. The latter is also the leading complexity in this LMS version.

Both algorithmic schemes first appeared in [14] as a means to perform *adaptive subspace tracking*. We should mention that the subspace tracking literature is particularly rich offering numerous algorithms for adaptively estimating (and tracking) subspaces. In fact there even exist versions with complexity (translated to our terminology) $O((N+L)L)$, which is smaller than the one enjoyed by our LMS scheme. We would like however to point out that these low complexity algorithms, are primarily applied for estimating subspaces corresponding to the largest singular values. There exist very few schemes providing estimates for the smallest singular values and can be found in [15, 16, 17]. Unfortunately, as it is reported in [17], all of them exhibit numerical instability.

It turns out that for the special algorithm proposed in (26), (27) we can develop an orthonormalization procedure with complexity $O((N+L)L)$ thus reducing the overall complexity to this level. Specifically Equ. (28) must be replaced with the following set of equations

$$\mathbf{a}(n) = \mathbf{s}(n) - \|\mathbf{s}(n)\|\mathbf{e}_1 \tag{29}$$

$$\hat{\mathbf{T}}(n) = \mathbf{T}(n) - \frac{1}{\mathbf{s}^H(n)\mathbf{a}(n)}[\mathbf{T}(n)\mathbf{a}(n)]\mathbf{a}^H(n) \tag{30}$$

$$\mathbf{U}_w(n) = \mathrm{norm}\{\hat{\mathbf{T}}(n)\}, \tag{31}$$

where $\mathbf{e}_1 = [1\ 0 \ldots 0]^t$ and "norm" stands for normalization of the columns of the matrix $\hat{\mathbf{T}}(n)$. The corresponding complexity is $O((N+L)L)$ since the normalization of a vector of length $N+L$ requires $O(N+L)$ operations. Compared to the complexity $O((N+L)^2L)$ of RLS we have gained an order of magnitude.

The strong point of our algorithm is its numerical stability, that is, even if orthonormality is lost, the adaptation converges rapidly to an orthonormal matrix. Furthermore the algorithm is simple having only a single parameter (the step size $\bar{\mu}$) to be specified. The numerical stability, as well as, the analysis of the transient and steady state behavior of the algorithm will be detailed in an upcoming article. In the same article the algorithm will also be compared against all existing subspace tracking schemes of similar complexity.

Regarding now the initialization of the two versions, we propose the following common scheme. We apply a QR decomposition on the matrix $\sum_{n=1}^{N+L} \mathbf{y}(n)\mathbf{y}^H(n)$ and use the L last orthonormal vectors to initialize $\mathbf{U}_w(n)$.

5.2 Adaptive Solution of the Second Subspace Problem

Once we have available the estimates $\mathbf{U}_w(n)$ of the noise subspace basis, either through the RLS: (22)-(25) or the LMS: (26),(27), (29)-(31) adaptation, we can then proceed with the estimate of the matrix $\mathbf{W}$. If $[\mathbf{v}_1(n) \cdots \mathbf{v}_L(n)]$ are the L columns of the matrix $\mathbf{U}_w(n)$ we then transform each column $\mathbf{v}_i(n)$ into the corresponding Hankel version $\mathbf{V}_i(n)$ according to (8) and finally compute $\mathbf{W}(n)$ according to (11).

Notice that the vectors $\mathbf{v}_i(n)$ constitute efficient estimates of the singular vectors of the noise subspace, therefore $\mathbf{W}(n)$ is also an efficient estimate of $\mathbf{W}$. Because of this fact the singular vector corresponding to the smallest singular value of $\mathbf{W}$ can be estimated using either (18) or (19). We propose the use of (19) since it has complexity $O(L^2)$ as opposed to $O(L^3)$ for (18) (due to the matrix inversion). Adopting, as in the first problem, a normalized step size, that is, $\mu = \bar{\mu}/\text{trace}\{\mathbf{W}(n)\}$, with $0 < \bar{\mu} \leq 1$, we propose $\bar{\mu} = 1$. Thus the final channel adaptation becomes

$$\mathbf{h}_b(n) = \mathbf{h}_b(n-1) - \frac{1}{\text{trace}\{\mathbf{W}(n)\}}\mathbf{W}(n)\mathbf{h}_b(n-1). \tag{32}$$

The computational complexity of the second subspace problem is as follows. For the computation of the matrix $\mathbf{W}(n)$ we need $O(NL^2)$. This complexity is attainable if we carefully exploit the Hankel structure of the matrices $\mathbf{V}_i(n)$. Finally, as it was pointed out, (32) requires $O(L^2)$ operations. Thus the leading complexity is $O(NL^2)$.

5.3 Phase and Amplitude Ambiguity Removal

The true channel impulse response $\mathbf{h}$ is related to the normalized version $\mathbf{h}_b$ through Equ. (12), where the complex parameter α expresses the phase and amplitude ambiguity. It is possible to recover α by inserting *pilot* symbols in the symbol blocks $\mathbf{b}(n)$. Even a single pilot symbol (in every symbol block) is sufficient to eliminate this ambiguity. We should mention that pilot symbols are included in all current standards.

Let us first estimate α assuming that a normalized channel impulse response $\mathbf{h}_b$ and the statistics of the processes involved are available. We can verify from (3) that if $\omega_i = i\frac{2\pi}{N}$, $i = 0, \ldots, N-1$, is the i-th subcarrier frequency, then

$$\begin{aligned}
[1\ e^{-j\omega_i} \cdots e^{-j(N+L-1)\omega_i}]\mathbf{y}(n) &= \mathcal{H}(\omega_i)[1\ e^{-j\omega_i} \cdots e^{-j(N-1)\omega_i}]\mathbf{F}^H\mathbf{b}(n) + w_i(n) && (33)\\
&= \mathcal{H}(\omega_i)[1\ W_N^i \cdots W_N^{i(N-1)}]\mathbf{F}^H\mathbf{b}(n) + w_i(n) && (34)\\
&= \sqrt{N}\,\mathcal{H}(\omega_i)b_i(n) + w_i(n) && (35)
\end{aligned}$$

where the last equality is due to the orthonormality of $\mathbf{F}$; we have that

$$\mathcal{H}(\omega_i) = \sum_{k=0}^{L} e^{-jk\omega_i}h_k = [1\ e^{-j\omega_i} \cdots e^{-jL\omega_i}]\mathbf{h} \tag{36}$$

is the channel frequency response at ω_i; $b_i(n)$ is the i-th symbol in the $\mathbf{b}(n)$ symbol block and finally $w_i(n)$ is a noise term.

If p pilot symbols are available at the $i_1, i_2, \ldots, i_p$ positions of the block $\mathbf{b}(n)$, let us consider the following matrix of dimensions $p \times K$

$$\mathbf{\Phi}_K = \begin{bmatrix} 1 & e^{-j\omega_{i_1}} & \cdots & e^{-j(K-1)\omega_{i_1}} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & e^{-j\omega_{i_p}} & \cdots & e^{-j(K-1)\omega_{i_p}} \end{bmatrix}. \tag{37}$$

Because of (35), (36) and (12) we can then write

$$\begin{aligned}\mathbf{\Phi}_{N+L}\mathbf{y}(n) &= \sqrt{N}\,\mathrm{diag}\{\mathbf{b}_p(n)\}[\mathcal{H}(\omega_{i_1})\cdots\mathcal{H}(\omega_{i_p})]^t + \mathbf{w}_p(n) &(38)\\ &= \sqrt{N}\,\mathrm{diag}\{\mathbf{b}_p(n)\}\mathbf{\Phi}_{L+1}\mathbf{h} + \mathbf{w}_p(n) &(39)\\ &= \alpha\sqrt{N}\,\mathrm{diag}\{\mathbf{b}_p(n)\}\mathbf{\Phi}_{L+1}\mathbf{h}_b + \mathbf{w}_p(n), &(40)\end{aligned}$$

where $\mathbf{b}_p(n) = [b_{i_1}(n)\cdots b_{i_p}(n)]^t$ is a vector containing the pilot symbols and $\mathbf{w}_p(n)$ is a noise vector. Since the symbols $b_i(n)$ are of unit norm and also independent from the noise term $\mathbf{w}_p(n)$ we conclude that

$$\alpha = \frac{\mathbf{h}_b^H\mathbf{\Phi}_{L+1}^H\mathbb{E}\left\{\mathrm{diag}\{\mathbf{b}_p^\star(n)\}\mathbf{\Phi}_{N+L}\mathbf{y}(n)\right\}}{\sqrt{N}\,\mathbf{h}_b^H\mathbf{\Phi}_{L+1}^H\mathbf{\Phi}_{L+1}\mathbf{h}_b}. \tag{41}$$

This suggests the following simple adaptation for the scalar parameter α

$$\alpha(n) = \nu\,\alpha(n-1) + (1-\nu)\frac{\mathbf{h}_b^H(n)\mathbf{\Phi}_{L+1}^H\mathrm{diag}\{\mathbf{b}_p^*(n)\}\mathbf{\Phi}_{N+L}\mathbf{y}(n)}{\sqrt{N}\,\mathbf{h}_b^H(n)\mathbf{\Phi}_{L+1}^H\mathbf{\Phi}_{L+1}\mathbf{h}_b(n)}, \tag{42}$$

where $0 < \nu < 1$ is a forgetting factor and $\mathbf{h}_b(n)$ is available from the blind subspace part, that is, adaptation (32). Notice that (42) involves only known quantities.

6 Simulations

Let us now present several simulation examples. Following the HYPERLAN2 standard, we consider $N = 64$ with a zero padding of length $L = 16$. Inside each symbol block there are $p = 4$ pilot symbols at the positions $i_1 = 0, i_2 = 16, i_3 = 32$ and $i_4 = 48$. For the RLS algorithm we select $\lambda = 0.997$ when SNR=20 dB and $\lambda = 0.9985$ when SNR=10 db. For LMS we select $\bar{\mu} = 1$ when SNR=20 dB and $\bar{\mu} = 0.8$ when SNR=10 dB. Finally for the adaptation in (42) we select $\nu = 0.99$. The reason we change our parameters with SNR is to be able to come up with graphs that allow for *fair comparisons*. This is possible when we match either convergence rates or steady state behaviors. Finally the type of additive noise used in all simulations is Gaussian.

In addition to the RLS and LMS version we also simulate the direct SVD approach. This consists in applying, at each time step, an SVD on the matrix $\mathbf{P}(n)$ provided by RLS, to obtain $\mathbf{U}_w(n)$. We then form $\mathbf{W}(n)$ and then apply an SVD on this matrix to obtain $\mathbf{h}_b(n)$. Once $\mathbf{h}_b(n)$ is available we adapt $\alpha(n)$ using (42).

In all figures we plot the ratio $\mathbb{E}\{\|\mathbf{h}-\alpha(n)\mathbf{h}_b(n)\|^2\}/\|\mathbf{h}\|^{\mathbf{2}}$, in dB, which corresponds to the relative channel estimation error power. Expectation is approximated using average of 100 independent runs. Figures 2 & 3 depict the performance of the algorithms under a non-fading channel. We start with a channel that has impulse response $\mathbf{h}^t$=[0.555 0.160 0.141 0.316]+ j[0.214 0.636 0.290 -0.114] and at time 5000 we abruptly switch to $\mathbf{h}^t$=[-0.189 -0.284 0.127 -0.045]+j[0.427 0.698 0.432 0.091]. The main characteristic of both channels

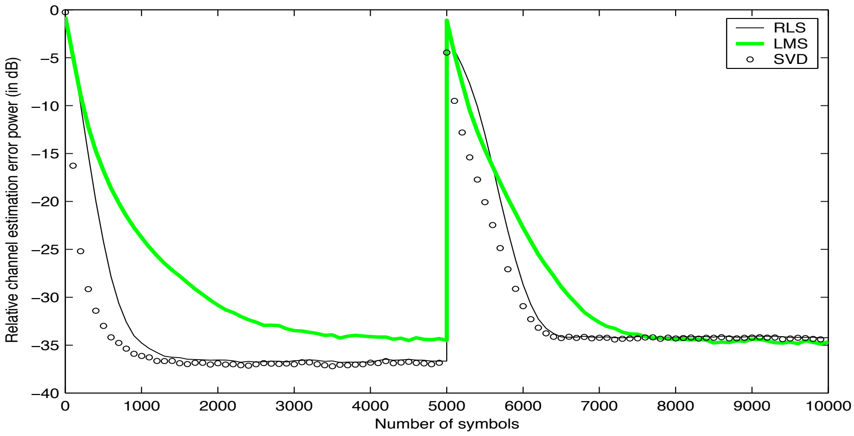

Figure 2: Performance of RLS, LMS and Direct SVD for SNR=20 dB; no fading.

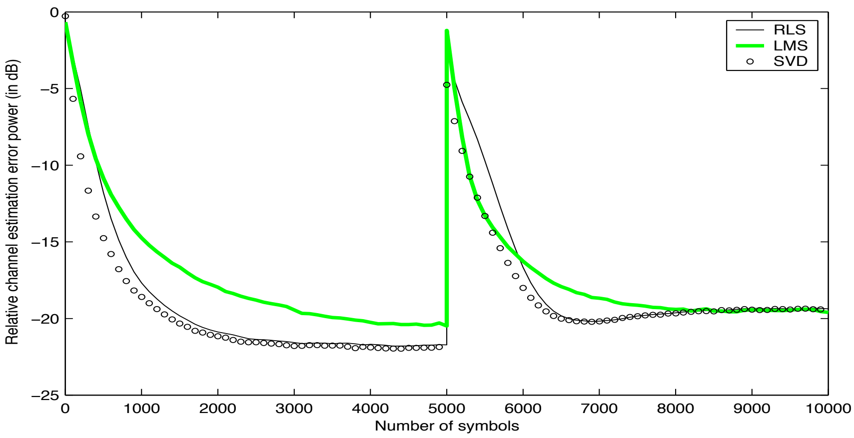

Figure 3: Performance of RLS, LMS and Direct SVD for SNR=10 dB; no fading.

is that they strongly attenuate certain frequency regions (for details see [18]). Although both channels are of length 4 we estimate them as being of maximum length $L+1=17$. Fig. 2 presents the results for SNR=20 dB and Fig. 3 for SNR=10 dB. In both cases RLS is very close to the direct SVD approach but at a computational level almost an order of magnitude smaller. The LMS version, on the other hand, has performance that compares very favorably with the other two algorithms but with a very appealing computational complexity.

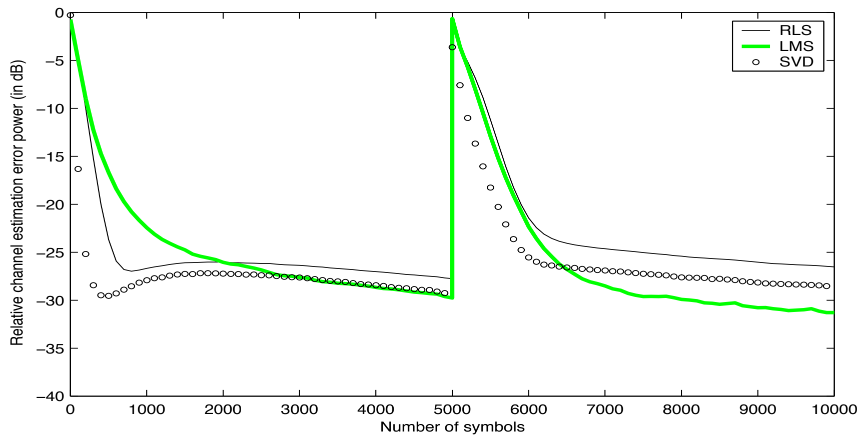

Figure 4: Performance of RLS, LMS and Direct SVD for SNR=20 dB; fading channel.

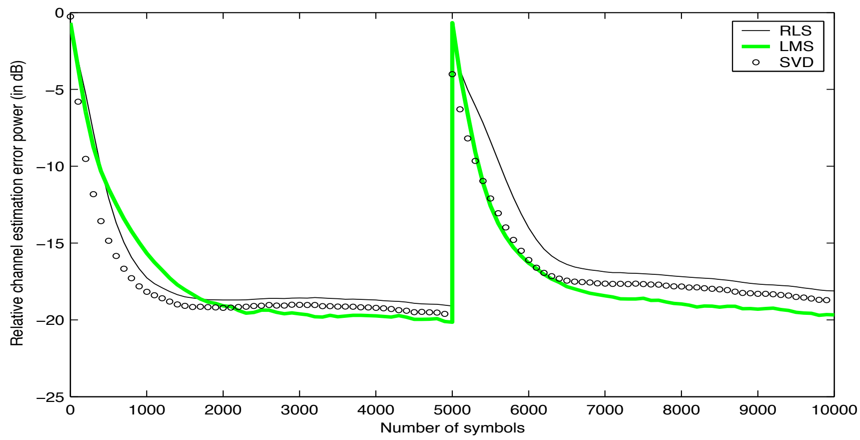

Figure 5: Performance of RLS, LMS and Direct SVD for SNR=10 dB; fading channel.

In Figures 4 & 5 we present the performance of the three algorithms under a fading channel for SNR=20 dB and SNR=10 dB respectively. We use a Jakes-like model, proposed in [19], to simulate fading. The parameters of the model are: communication frequency carrier at 5Ghz; data rate at 2Mbits/sec; receiver speed 3m/sec and 15 scatterers per channel coefficient (scatterers for different channel taps are independent). As we can see RLS continues to follow closely the direct SVD approach. However what is remarkable here is that LMS can outperform both algorithms. Even though this property might seem extraordinary we should point out that a similar performance of LMS has already been observed in conventional adaptive system identification [20, Page 651].

7 Conclusion

In this article, we have considered the problem of blind adaptive channel estimation in ZP-OFDM systems. By defining two subspace problems we were able to determine the channel impulse response modulo a phase and amplitude ambiguity. Motivated by the orthogonal iteration method, known from Numerical Analysis for the computation of singular vectors, RLS and LMS schemes were developed capable of providing blind adaptive channel estimates. As far as the LMS version is concerned it was based on a novel, low complexity and numerically stable subspace tracking algorithm proposed here for the first time. Both versions were also extended to take into account the existence of pilot symbols in order to eliminate the ambiguity which is intrinsic in all blind techniques. The proposed algorithms were tested under diverse signaling conditions involving medium and high SNR levels in stationary and slowly fading channels that also exhibit abrupt changes. In all cases convergence was rapid matching the performance of the non adaptive and computationally intense, direct SVD approach.

Appendix

Proof of Theorem 1: Let $\mathbf{h} = [h_0 \cdots h_L]^t$ be the true channels coefficients. Define the polynomial $h(z) = h_0 + zh_1 + \cdots + z^L h_L$ and suppose for simplicity that $h(z)$ has L distinct roots z_i, $i = 1, \ldots, L$. Consider now the vectors $\mathbf{z}_i = [1\; z_i^\star \cdots (z_i^\star)^{N+L-1}]^t$, $i = 1, \ldots, L$. These L vectors span the noise subspace because, as we can verify, $\mathbf{H}\mathbf{H}^H \mathbf{z}_i = 0$ and they are linearly independent. From this we conclude that any vector $\mathbf{v}$ in the noise subspace can be written as a linear combination of the vectors $\mathbf{z}_i$. Because $\mathbf{U}_w$ spans also the noise subspace we can write

$$\mathbf{U}_w = [\mathbf{v}_1 \cdots \mathbf{v}_L] = [\mathbf{z}_1 \cdots \mathbf{z}_L]\mathbf{A}, \tag{43}$$

where $\mathbf{A} = [\mathbf{a}_1 \cdots \mathbf{a}_L]$ is a full rank (thus invertible) matrix of dimensions $L \times L$.

If $\tilde{\mathbf{h}}$ is a vector satisfying (9) then, due to the nonegative definitness of the terms $\mathbf{V}_i\mathbf{V}_i^H$ that compose $\mathbf{W}$, we also have

$$\mathbf{V}_i^H\tilde{\mathbf{h}} = 0,\ i = 1, \ldots, L. \tag{44}$$

Now we recall that $\mathbf{V}_i$ is the Hankel version of $\mathbf{v}_i$ which is the i-th column of $\mathbf{U}_w$. From (43) we have that $\mathbf{v}_i = [\mathbf{z}_1 \cdots \mathbf{z}_L]\mathbf{a}_i = \sum_{l=1}^{L} a_{il}\mathbf{z}_i$ where $\mathbf{a}_i = [a_{i1} \cdots a_{iL}]^t$. This means that $\mathbf{V}_i = \sum_{l=1}^{L} a_{il}\mathbf{Z}_i$, where $\mathbf{Z}_i$ is the Hankel version of $\mathbf{z}_i$. The latter, due to the special form of $\mathbf{z}_i$ can be written as $\mathbf{Z}_i = [1\ z_i^\star \cdots (z_i^\star)^L]^t[1\ z_i^\star \cdots (z_i^\star)^{N-1}]$. Due to this property, if we define the matrices

$$\mathbf{B} = \begin{bmatrix} 1 & \cdots & 1 \\ z_1 & \cdots & z_L \\ \vdots & \vdots & \vdots \\ z_1^{N-1} & \cdots & z_L^{N-1} \end{bmatrix},\ \mathbf{C} = \begin{bmatrix} 1 & z_1 & \cdots & z_1^L \\ \vdots & \vdots & \vdots & \\ 1 & z_L & \cdots & z_L^L \end{bmatrix}, \tag{45}$$

we can then see that we can write $\mathbf{V}_i$ more compactly as follows

$$\mathbf{V}_i^H = \mathbf{B}\,\mathrm{diag}\{\mathbf{a}_i\}\mathbf{C}. \tag{46}$$

Because $\mathbf{B}$ is Vandermonde and the z_i are distinct, when $N \geq L$ then $\mathbf{B}$ is of full column rank. This, using (44), allows us to write

$$\mathbf{B}\,\mathrm{diag}\{\mathbf{a}_i\}\mathbf{C}\tilde{\mathbf{h}} = \mathbf{0} = \mathrm{diag}\{\mathbf{a}_i\}\mathbf{C}\tilde{\mathbf{h}} = \mathrm{diag}\{\mathbf{C}\tilde{\mathbf{h}}\}\mathbf{a}_i,\ i = 1, \ldots, L. \tag{47}$$

We can now combine the L equations $\mathrm{diag}\{\mathbf{C}\tilde{\mathbf{h}}\}\mathbf{a}_i = \mathbf{0},\ i = 1, \ldots, L$ into $\mathrm{diag}\{\mathbf{C}\tilde{\mathbf{h}}\}\mathbf{A} = \mathbf{0}$, from which we obtain $\mathrm{diag}\{\mathbf{C}\tilde{\mathbf{h}}\} = \mathbf{0}$, thanks to the invertibility of $\mathbf{A}$. The latter is also equivalent to

$$\mathbf{C}\tilde{\mathbf{h}} = \mathbf{0}. \tag{48}$$

We should note that the same conclusion can be drawn directly from (47) and in particular from $\mathrm{diag}\{\mathbf{a}_i\}\mathbf{C}\tilde{\mathbf{h}} = \mathbf{0}$. Indeed if the vector $\mathbf{a}_i$ of a *single* $\mathbf{v}_i$ has *all* its elements different than zero then we can also conclude that $\mathbf{C}\tilde{\mathbf{h}} = \mathbf{0}$. From (46) we have that the vector $\mathbf{a}_i$ has nonzero elements iff the corresponding $\mathbf{V}_i$ is of full row rank. The interesting point is that if we select a vector $\mathbf{v}$ in the noise subspace then the probability that this vector will have a Hankel version $\mathbf{V}$ which is not of full row rank is *zero*.

From (48) and because of the special form of $\mathbf{C}$, depicted in (45), we deduce that the z_i, $i = 1, \ldots, L$, are the L roots of the polynomial that has as coefficients the elements of $\tilde{\mathbf{h}}$. But this polynomial is uniquely defined (modulo a multiplicative parameter) through its roots. Therefore $\tilde{\mathbf{h}} = \alpha\mathbf{h}$, and this concludes the proof.

References

[1] J.A.C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Comm. Mag.*, vol. 28, pp. 5-14, May 1990.

[2] B. Muquet, Z. Wang, G.B. Giannakis, M. de Courville, and P. Duhamel, "Cyclic prefixing or zero padding for wireless multicarrier transmissions?" *IEEE Trans. Comm.*, vol. 50, pp. 2136-2148, Dec. 2002.

[3] L. Deneire, P. Vandenameele, L. Perre, B.Gyselinckx, and M. Engels, "A low-complexity ML channel estimator for OFDM," *IEEE Trans. Comm.*, vol. 51, pp. 135-140, Feb. 2003.

[4] O. Edfors, M. Sandel, J.J.V. Beek, S.K. Wilson, and P.O. Börjesson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Comm.*, vol. 46, pp. 931-939, July 1998.

[5] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems," *IEEE Trans. Sign. Proc.*, vol. 49, pp. 3065-3073, Dec. 2001.

[6] R.W. Heath and G.B. Giannakis, "Exploiting input cyclostationarity for blind channel identification in OFDM systems," *IEEE Trans. Sign. Proc.*, vol. 37, no. 3, pp. 848-856, March 1999.

[7] B. Muquet, M. de Courville, and P. Duhamel, "Subspace-based blind and semi-blind channel estimation for OFDM systems," *IEEE Trans. Sign. Proc.*, vol. 50, pp. 1699-1712, July 2002.

[8] C. Li and S. Roy, "Subspace-based blind channel estimation for OFDM by exploiting virtual carriers," *IEEE Trans. Wir. Comm.*, vol. 2, pp. 141-150, Jan 2003.

[9] G.H. Golub and C.F. Van Loan, *Matrix Computations*, 2nd edn, The John Hopkins University Press, 1990.

[10] X.G. Doukopoulos and G.V. Moustakides, "Power techniques for blind adaptive channel estimation in CDMA systems," IEEE Global Communications Conference, GLOBECOM'2003, San Francisco, Dec. 2003.

[11] Z. Xu, P. Liu and X. Wang, "Blind multiuser detection: From MOE to subspace methods," *IEEE Trans. on Signal Processing*, accepted for publication, to appear.

[12] G.V. Moustakides, "Exponential convergence of products of random matrices, application to the study of adaptive algorithms," *International Journal of Adaptive Control and Signal Processing*, vol. 2, no. 12, pp. 579-597, Dec. 1998.

[13] P. Comon and G.H. Golub, "Tracking a few extreme singular values and vectors in signal processing," *Proceedings of the IEEE*, vol. 78, no. 8, pp. 1327-1343, Aug. 1990.

[14] J.F. Yang and M. Kaveh, "Adaptive eigenspace algorithms for direction or frequency estimation and tracking," *IEEE Trans. ASSP*, vol. 36, no. 2, pp. 241-251, Feb. 1988.

[15] E. Oja, "Principal components, minor components, and linear neural networks," *Neural Networks*, vol. 5, pp. 927-935, Nov./Dec. 1992.

[16] B. Yang, "Projection approximation subspace tracking", *IEEE Trans. Sign. Proc.*, vol. 43, pp. 95-107, Jan. 1995.

[17] S. Attallah and K. Abed-Meraim, "Fast algorithms for subspace tracking," *IEEE Signal Processing Letters*, vol. 8, no. 7, pp 203-206, July 2001.

[18] S. Roy and C. Li, "A subspace blind channel estimation method for OFDM systems without cyclic prefix," *IEEE Trans. Wir. Comm.*, vol. 1, pp. 572-579, Oct. 2002.

[19] J.K. Cavers, *Mobile Channel Characteristics*, Kluwer Avademic Publishers, 2000.

[20] S. Haykin, *Adaptive Filter Theory*, 4-th Edition, Prentice-Hall, Upper Saddle River, NJ, 2002.

Unité de recherche INRIA Rennes
IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifi que
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot-St-Martin (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399