Pilot-Based Channel Estimation for OFDM Systems by Tracking the Delay-Subspace

Osvaldo Simeone, *Student Member, IEEE*, Yeheskel Bar-Ness, *Fellow, IEEE*, and Umberto Spagnolini, *Senior Member, IEEE*

***Abstract*—In orthogonal frequency division multiplexing (OFDM) systems over fast-varying fading channels, channel estimation and tracking is generally carried out by transmitting known pilot symbols in given positions of the frequency-time grid. The traditional approach consists of two steps. First, the least-squares (LS) estimate is obtained over the pilot subcarriers. Then, this preliminary estimate is interpolated/smoothed over the entire frequency-time grid. In this paper, we propose to add an intermediate step, whose purpose is to increase the accuracy of the estimate over the pilot subcarriers. The presented techniques are based on the observation that the wireless radio channel can be parametrized as a combination of paths, each characterized by a delay and a complex amplitude. The amplitudes show fast temporal variations due to the mobility of terminals while the delays (and their associated delay-subspace) are almost constant over a large number of OFDM symbols. We propose to track the delay-subspace by a subspace tracking algorithm and the amplitudes by the least mean square algorithm (or modifications of the latter). The approach can be extended to multiple input multiple output OFDM or multicarrier code-division multiple-access systems. Analytical results and simulations prove the relevant benefits of the novel structure.**

***Index Terms*—Channel estimation, fading radio channel, multiple input multiple output (MIMO), orthogonal frequency-division multiplexing (OFDM), subspace tracking.**

I. Introduction

APPLICATIONS of orthogonal frequency-division multiplexing (OFDM) to wireless and mobile communications are currently under study. Although multicarrier transmission has several considerable drawbacks (such as high peak to average ratio and strict requirements on carrier synchronization), its advantages in lessening the severe effects of frequency selective fading without complex equalization are very attractive features. In order to obtain the high spectral efficiencies required by future data wireless systems [1], it is necessary to employ multilevel modulation with nonconstant amplitude (e.g., 16QAM [2]). This implies the need for coherent receivers that are capable to track the variations of the fading channel.

Manuscript received May 11, 2002; revised October 3, 2002 and November 22, 2002; accepted November 27, 2002. The editor coordinating the review fo this paper and approving it for publication is G. Leus. This work was supported in part by the National Science Foundation under Grant 9903381.

O. Simeone is with the Dipartimento di Elettronica e Informazione, Politecnico di Milano, I-20133 Milan, Italy, and also with the Center for Communications and Signal Processing Research, New Jersey Institute of Technology, Newark, NJ 07102 USA (e-mail: simeone@elet.polimi.it).

Y. Bar-Ness is with the Center for Communications and Signal Processing Research, New Jersey Institute of Technology, Newark, NJ 07102 USA (e-mail: barness@yegal.njit.edu).

U. Spagnolini is with the Dipartimento di Elettronica e Informazione, Politecnico di Milano, I-20133 Milan, Italy (e-mail: spagnoli@elet.polimi.it).

Digital Object Identifier 10.1109/TWC.2003.819022

The channel estimation (tracking) in OFDM systems is generally based on the use of pilot subcarriers in given positions of the frequency-time grid. For fast-varying channels (e.g., in mobile systems), nonnegligible fluctuations of the channel gains are expected between consecutive OFDM symbols (or even within each symbol) so that, in order to ensure an adequate tracking accuracy, it is advisable to place pilot subcarriers in each OFDM symbol [3], [4]. In particular, in this paper, we consider the comb pilot pattern arrangement [3], [4] shown in Fig. 1, which has been shown to satisfy different criteria of optimality such as mean square error on the channel estimate [5] and capacity [6].

In this framework, the traditional approach to channel estimation, that may be used as an initial estimate in iterative or decision directed receivers [7], [8], consists of two steps. First, the least squares (LS) estimates of the channel gains over the pilot subcarriers are obtained by simply backrotating the received signal according to the knowledge of the pilot symbols. This step can be equivalently seen as the two-dimensional (2-D) (i.e., in frequency and time) sampling of (a noisy version of) the wide sense stationary uncorrelated scattering (WSSUS) process [9] represented by the mobile radio channel. Then, the LS estimates are interpolated/smoothed over the entire frequency-time grid. This task can be accomplished by means of a minimum mean square error (MMSE) filter (2-D or separable) [10]–[12], by simply constraining the support of the temporal response (inverse fast Fourier transform (IFFT)/fast Fourier transform (FFT)-based interpolation) [13] or by a combination of the two approaches [14]. The design of the optimal (MMSE) interpolator requires knowledge of the 2-D correlation function of the channel, i.e., of the power-delay profile and the Doppler spectrum. Since this information is not easily available at the receiver, the design problem becomes that of finding the most robust estimator with respect to a mismatch in the channel correlation [15].

In this paper, we propose a subspace-based technique for channel estimation over the pilot subcarriers that is based on the exploitation of the slowly-varying delay-subspace. From a practical point of view, the method is a preinterpolation channel estimation and it adds an intermediate step between the LS estimator over the pilot subcarriers and the interpolator. The purpose of such a modification of the conventional approach is that of improving the accuracy of the LS estimate of the channel gains over the pilot subcarriers *before* the interpolation [4]. The key observation is that the channel can be parametrized as a sum of contributions, each related to a different multipath component, characterized by a delay and a complex amplitude. The delays present much slower variations in time than the amplitudes, allowing the two types of parameters to be handled

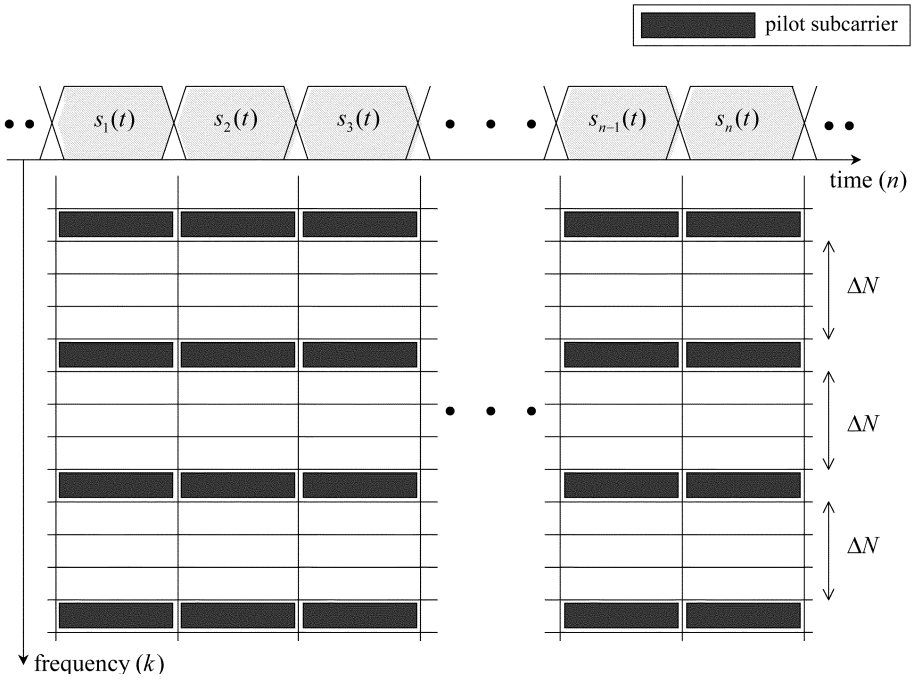


Fig. 1. Comb pilot pattern.

differently. In [16], it is proposed to explicitly estimate the delays by using the ESPRIT algorithm for acquisition and a bank of delay locked loops (DLL) for tracking, whereas the final estimate of the channel is obtained by solving a conventional MMSE problem. The direct computation of the delays is a nonlinear estimation problem and, as such, it suffers from threshold effects at low signal-to-noise ratios (SNRs). Moreover, the computational burden is relevant. Here, we overcome these problems by avoiding the explicit estimate of the delays. Instead, the proposed techniques perform an unstructured tracking of the slow variations of the corresponding delay-subspace through one of the so called subspace tracking algorithms [17]. The amplitudes' fast variations are tracked separately, e.g., by using the least mean square (LMS) algorithm.

Although most of the paper is devoted to the study of a single input single output system, the approach can be generalized to multiple input multiple output (MIMO)-OFDM and multicarrier code-division multiple-access (MC-CDMA) systems.

The structure of the paper is as follows. Sections II and III introduce the baseband model of an OFDM system and of the multipath channel respectively. The proposed channel estimators are discussed in Section IV: the batch algorithm is presented in Section IV-A, the operation of the subspace tracker and the corresponding subspace tracking (ST) estimator in Section IV-B, and the combined subspace-amplitudes tracking (SAT) in Section IV-C. An analytical study of the performance is carried out in Section V. Finally, in Section VI the performance is evaluated through computer simulations demonstrating the benefits of the proposed channel estimation algorithms.

A. Basic Notation

In this paper, bold denotes column vectors or matrices as it will be clear from the context; $(\cdot)^T$ denotes transposition; $(\cdot)^H$ denotes hermitian transposition; if x is a given $M \times 1$ vector, $\mathrm{diag}\,(\mathbf{x})$ is the $M \times M$ diagonal matrix with the entries of x on the main diagonal; the i-th entry of a given $M \times 1$ vector x is denoted as $\mathbf{x}[i]$, $i = 0, .., N-1$; and $\mathbf{I}_M$ is the $M \times M$ identity matrix.

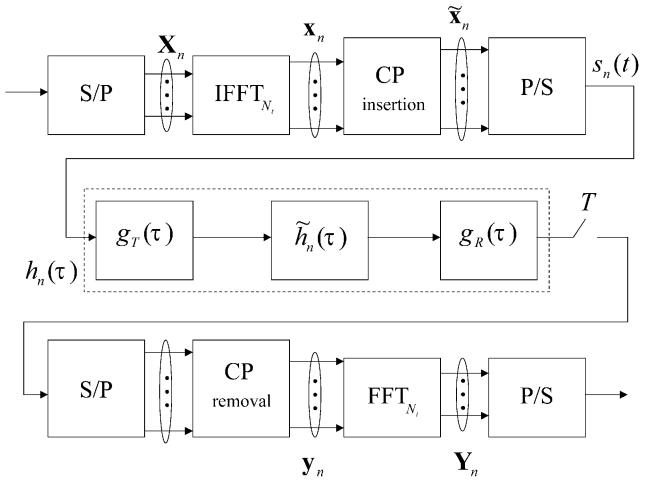


Fig. 2. Baseband model of an OFDM system.

II. OFDM System

The baseband model of an OFDM system is shown in Fig. 2. The bandwidth $B = 1/T$ is divided into N_t equally spaced subcarriers at frequencies $(k - N_t/2)\Delta f$ with $\Delta f = B/N_t$ and $k = 0, .., N_t - 1$. The N central subcarriers $(k - N/2)\Delta f$ with $k = 0, .., N-1$ (without any loss of generality, N_t and N are assumed to be even) are used for transmitting data whereas the remaining subcarriers (i.e., virtual subcarriers [18]) form the guard band. These are necessary to make the transmission filter $g_T(\tau)$ physically realizable. The complex valued data symbols to be transmitted are serial to parallel converted producing a $N\times$

1 vector $\mathbf{X}_n$ every T_b s. The N_t-point IFFT block modulates the entries of $\mathbf{X}_n$ on the N data subcarriers (the virtual subcarriers are zeroed) producing the $N_t \times 1$ vector $\mathbf{x}_n$. The nth OFDM symbol $\tilde{\mathbf{x}}_n$ is completed by adding to $\mathbf{x}_n$ a preamble [cyclic prefix (CP)] made of its last D samples. After parallel to serial conversion, we thus obtain the discrete-time nth OFDM symbol $s_n(t)$, sampled at multiples of T, of duration $T_b = T_s + T_{cp}$, where $T_s = N_t T$ and $T_{cp} = DT$ denote the length of the "useful" part of the OFDM symbol and of the CP respectively. Within the support $t \in (nT_b, (n+1)T_b)$, the OFDM symbol is

$$s_n(t) = \sum_{k=0}^{N-1} \sum_{i=0}^{N_t+D-1} \mathbf{X}_n[k] \exp\left[j2\pi\left(k-\frac{N}{2}\right)\frac{(i-D)}{N_t}\right]\delta(t-iT-nT_b). \tag{1}$$

At the transmitter side, the discrete-time OFDM symbol $s_n(t)$ is (continuous-time) convolved with the pulse-shaping filter $g_T(\tau)$. The radio channel is modeled as a time-varying finite impulse response filter $\tilde{h}_n(\tau)$. The notation $\tilde{h}_n(\tau)$ indicates that the time-varying channel is assumed quasi-static, i.e., constant during the transmission of an OFDM symbol and variable on a symbol-by-symbol basis (see Section III). The additive noise $n(t)$ is assumed to be white and circularly symmetric Gaussian (WG). At the receiver, assuming perfect timing and carrier synchronization, the signal is matched filtered by $g_R(\tau)$ and sampled at $1/T$. The filters $g_T(\tau)$ and $g_R(\tau)$ are square-root Nyquist waveforms with roll-off α. Every T_b s, the serial to parallel converter outputs the $(N_t + D) \times 1$ vector $\tilde{\mathbf{y}}_n$, from which the CP is removed producing the $N_t \times 1$ vector $\mathbf{y}_n$. Then, an N_t-point FFT block generates the $N \times 1$ vector $\mathbf{Y}_n$ where the $(N_t - N)$ samples corresponding to the guard band are discarded.

By assuming that the support of $h_n(\tau) = g(\tau) * \tilde{h}_n(\tau)$ ($g(\tau) = g_T(\tau) * g_R(\tau)$) is smaller than or equal to $T_{cp} + T = (D+1)T$, the received signal can be written in matrix notation as

$$\mathbf{Y}_n = \mathrm{diag}(\mathbf{X}_n)\mathbf{H}_n + \mathbf{N}_n \tag{2}$$

where $\mathbf{N}_n$ is a $N \times 1$ vector of WG noise with $E[\mathbf{N}_n\mathbf{N}_n^H] = \sigma^2\mathbf{I}_N$ and $\mathbf{H}_n$ is a $N \times 1$ vector with $\mathbf{H}_n[k]$ being the Fourier Transform of $h_n(\tau)$ at frequency $(k - N/2)\Delta f$. The channel $\mathbf{H}_n$ and the noise $\mathbf{N}_n$ are statistically independent. Furthermore, without any loss of generality, the fading channel can be assumed to satisfy $E[|\mathbf{H}_n[k]|^2] = 1$. The SNR is defined accordingly as $\mathrm{SNR} = \sigma_x^2/\sigma^2$ ($\sigma_x^2 = E[\mathbf{X}_n[k]]$ assumed to be independent of k). We emphasize that throughout the paper, where not stated otherwise, $E[\cdot]$ denotes the expectation with respect to the distribution of the noise and of the channel (i.e., of delays and amplitudes, see Section III).

III. Time-Varying Multipath Channel

The fading channel $\tilde{h}_n(\tau)$ can be described as composed of L paths, each characterized by a delay $\tau_{n,\ell}$ and an amplitude $\alpha_{n,\ell}$, i.e., $\tilde{h}_n(\tau) = \sum_{\ell=1}^{L} \alpha_{n,\ell}\delta(\tau - \tau_{n,\ell})$[19]. It follows that the equivalent channel reads $h_n(\tau) = \sum_{\ell=1}^{L} \alpha_{n,\ell} g(\tau - \tau_{n,\ell})$ and the channel vector $\mathbf{H}_n$ in the frequency domain is (denoting by $\mathbf{G}[k]$ the Fourier Transform of $g(\tau)$ at frequency $(k-N/2)\Delta f$)

$$\begin{aligned}\mathbf{H}_n[k] &= \mathbf{G}[k]\sum_{\ell=1}^{L}\alpha_{n,\ell}\exp\left[-j\frac{2\pi}{N_t}\left(k-\frac{N}{2}\right)\frac{\tau_{n,\ell}}{T}\right] \\ &= \sum_{\ell=1}^{L}\alpha_{n,\ell}\exp\left[-j\frac{2\pi}{N_t}\left(k-\frac{N}{2}\right)\frac{\tau_{n,\ell}}{T}\right] \\ k &= 0,\ldots,N-1. \end{aligned} \tag{3}$$

The multipath model (3) has been widely used in the context of OFDM systems (see, e.g., [16]). A discussion on its generality in the perspective of the existing literature is in Section III-C. The following further remarks are in order: 1) the equivalent transmission filter $g(\tau)$ does not appear since we are assuming that the guard band is at least αB ($\mathbf{G}[k] = 1$ for $k = 0,\ldots,N-1$); 2) the delays $\tau_{n,\ell} \geq 0$ must be smaller than $T_{cp} + T - T_g$ (where T_g is the duration of the waveform $g(\tau)$, that in practice can be set to $T_g = 4 \div 6T$ with negligible modeling mismatch for practical roll-off factors [20]) in order to satisfy the condition about the temporal support of $h_n(\tau)$ under which we derived (2).

A. Time Variability of the Multipath Parameters

Let us write the combination of L delayed waveforms (3) in the matrix notation

$$\mathbf{H}_n = \mathbf{W}(\boldsymbol{\tau}_n)\boldsymbol{\alpha}_n \tag{4}$$

where the ℓth column of $N \times L$ matrix $\mathbf{W}(\boldsymbol{\tau}_n)$ is

$$\left[\exp\left(-j\frac{2\pi}{N_t}\frac{\tau_{n,\ell}}{T}\left(\frac{-N}{2}\right)\right)\cdots\exp\left(-j\frac{2\pi}{N_t}\frac{\tau_{n,\ell}}{T}\left(\frac{N}{2-1}\right)\right)\right]^T$$

and $\boldsymbol{\alpha}_n = [\alpha_{n,1}\cdots\alpha_{n,L}]^T$. According to the model (4) [or equivalently (3)], the channel is time-varying as a result of the *slow* variations of the set of delays $\boldsymbol{\tau}_n = [\tau_{n,1}\cdots\tau_{n,L}]^T$ and the *fast* variations of the amplitudes $\boldsymbol{\alpha}_n$. In practice, the delays can be considered as constant in M symbols provided that their variations in MT_b s are much smaller than the temporal resolution $1/B$ of the system. To quantify, for a relative radial movement between transmitter and receiver with velocity v, the variation of delay τ in MT_b s is $\Delta\tau = vMT_b/c$ (where $c = 3\times10^8$ m/s) so that the condition $\Delta\tau \ll 1/B$ leads to $M \ll (c/v)/(N_t + D)$. Therefore, even for a fast moving terminal with $v = 300$ km/h, we have $M \ll 3.6 \cdot 10^6/(N_t + D)$, meaning that the delays and the corresponding matrix $\mathbf{W}(\boldsymbol{\tau}_n)$ can be considered as constant for a large number of OFDM symbols (e.g., the condition $\Delta\tau \leq 10^{-2}/B$ with $N_t = 128$ and $D = 16$ yields $M \leq 250$, according to the numerical examples in Section VI).

The variations of the amplitudes across symbols depend on the Doppler shift. In OFDM systems, it is comparable to the spacing between different subcarriers leading to possibly very fast variations along different symbols. On the other hand, assuming constant amplitudes within one symbol entails neglecting the inter-carrier interference due to the Doppler shift. With the bounds derived in [21] it is possible to quantify the error due to this assumption (see Section VI).

B. Delay-Subspace

The properties of the matrix $\mathbf{W}(\boldsymbol{\tau}_n)$ in (4) have been studied in the context of structured delay estimation for known signatures [22]. This matrix is known to be rank-deficient in presence of closely spaced delays with respect to the time resolution $1/B$ of the system. We define the rank of $\mathbf{W}(\boldsymbol{\tau}_n)$ as $r_n = \mathrm{rank}(\mathbf{W}(\boldsymbol{\tau}_n)) \leq \min(N, L)$. As in subspace-based delay estimation, a prominent role in our technique is played by the r_n-dimensional subspace spanned by the columns of the $N \times L$ matrix $\mathbf{W}(\boldsymbol{\tau}_n)$. We refer to this subspace as "delay-subspace" and denote its $N \times r_n$ orthonormal basis as $\mathbf{U}_n$. It is relevant to emphasize again that the method proposed here does not explicitly estimate the multipath delays from the delay-subspace (i.e., by subspace-fitting [22]). Instead, the channel estimate is obtained directly from $\mathbf{U}_n$ as it will be discussed in Section IV.

Introducing the singular value decomposition $\mathbf{W}(\boldsymbol{\tau}_n) = \mathbf{U}_n\boldsymbol{\Lambda}_n\mathbf{V}_n^H$ the model (4) can be equivalently stated as

$$\mathbf{H}_n = \mathbf{U}_n\mathbf{d}_n \tag{5}$$

where $\mathbf{d}_n$ is the $r \times 1$ vector

$$\mathbf{d}_n = \boldsymbol{\Lambda}_n\mathbf{V}_n^H\boldsymbol{\alpha}_n. \tag{6}$$

The new parameterization (5) substitutes the structured matrix $\mathbf{W}(\boldsymbol{\tau}_n)$ with the unstructured basis $\mathbf{U}_n$ and the physical faded amplitudes $\boldsymbol{\alpha}_n$ with the compound amplitudes $\mathbf{d}_n$. It should be noted that while in $\mathbf{W}(\boldsymbol{\tau}_n)$ all the L delays are treated separately, the basis of the delay-subspace $\mathbf{U}_n$ combines all the delays within the resolution of the system in the same basis vector. This reduces the number of delays/basis vectors (and corresponding amplitudes) to be estimated and tracked from L to r. We further recall from the discussion in the previous Section that both the delay-subspace basis $\mathbf{U}_n$ and the compound amplitudes $\mathbf{d}_n$ in (5) depend on the current OFDM symbol n. However, the basis $\mathbf{U}_n$ shows temporal variations two-three orders of magnitude slower than the amplitudes.

C. Time/Frequency Duality in Channel Modeling

The frequency model (3) generalizes the widely used temporal model that considers $\mathbf{H}_n$ as the FFT of a $D \times 1$ (temporal) channel response $\mathbf{h}_n$ with independent identically distributed (i.i.d.) entries (see, e.g., [7]). To elaborate, let us consider (3). The low pass filtering implicitly carried out by neglecting the guard band causes the IFFT of $\mathbf{H}_n$ to be nonzero outside its first D samples (contrary to the temporal model). In other words, we can write

$$\mathbf{T}_{GB}\mathbf{H}_n = \mathbf{FFT}_{N_t}\begin{bmatrix}\mathbf{h}_n \\ \Delta\mathbf{h}_n\end{bmatrix} \tag{7}$$

where $\mathbf{T}_{GB}$ is the $N_t \times N$ matrix that introduces the $N_t - N$ zeros belonging to the guard band (or equivalently the low pass filtering), $\mathbf{h}_n$ a $D \times 1$ vector, and $\Delta\mathbf{h}_n$ is the $(N_t - D) \times 1$ vector that takes into account the truncation effects of the guard band. Now, if there is no guard band ($N_t = N$), we have $\mathbf{T}_{GB} = \mathbf{I}_{N_t}$ and $\Delta\mathbf{h}_n = \mathbf{0}$, so that $\mathbf{H}_n = \mathbf{FFT}_{N_t}[\mathbf{h}_n^T\ \mathbf{0}^T]^T$. In addition, if we further assume that the normalized delays $\tau_{n,\ell}/T$ are integers, the channel vector $\mathbf{h}_n$ has nonzero i.i.d. entries $\mathbf{h}_n[\ell]$ at indexes $\ell = \tau_{n,\ell}/T$. Therefore, it can be concluded that the temporal model follows from the frequency model (3) for the special case of no guard band (for the support of $\mathbf{h}_n$) and T-spaced delays (for the uncorrelation of the taps of $\mathbf{h}_n$). A discussion on the correlation of the channel taps $\mathbf{h}_n[\ell]$ in presence of non-sample-spaced delays can be found in [23, Section IV-C].

IV. Subspace-Based Channel Estimation Methods

The channel estimation methods proposed in this Section rely on the parametrization (5) and differ in the way they handle the two parameters, delay-subspace basis $\mathbf{U}_n$ and compound amplitudes $\mathbf{d}_n$. As we said in Section I, our techniques are presented here for an OFDM system with a comb pilot pattern (see Fig. 1). We denote the number of pilot subcarriers as $\bar{N}$ and upperscore the variables previously defined on the entire set of N available subcarriers when they have to be redefined on the set of $\bar{N}$ pilot subcarriers. For instance, $\bar{\mathbf{Y}}_n$ and $\bar{\mathbf{H}}_n$ are the $\bar{N} \times 1$ vectors that collect, respectively, the received signal and the channel gains on the pilot subcarriers. The input–output relation (2) thus becomes $\bar{\mathbf{Y}}_n = \mathrm{diag}(\bar{\mathbf{X}}_n)\bar{\mathbf{H}}_n + \bar{\mathbf{N}}_n$, where $\bar{\mathbf{H}}_n = \bar{\mathbf{W}}(\boldsymbol{\tau}_n)\boldsymbol{\alpha}_n$. Moreover, everything that has been said in Section III about the $N \times L$ matrix $\mathbf{W}(\boldsymbol{\tau}_n)$ can also be applied to the $\bar{N} \times L$ matrix $\bar{\mathbf{W}}(\boldsymbol{\tau}_n)$. In particular, we adopt the convention to refer to the subspace spanned by $\bar{\mathbf{W}}(\boldsymbol{\tau}_n)$ as "delay-subspace" and we denote the $\bar{N} \times \bar{r}_n$ orthonormal basis of this subspace as $\bar{\mathbf{U}}_n$ so that (5) and (6) modify as

$$\bar{\mathbf{H}}_n = \bar{\mathbf{U}}_n\bar{\mathbf{d}}_n \tag{8}$$

and

$$\bar{\mathbf{d}}_n = \bar{\boldsymbol{\Lambda}}_n\bar{\mathbf{V}}_n^H\boldsymbol{\alpha}_n. \tag{9}$$

Table I presents an overview of the proposed techniques. The second and third columns indicate how each method models and estimates the basis $\bar{\mathbf{U}}_n$. The remaining pair of columns are similarly defined and concern respectively the estimate of the rank $\bar{r}_n$ and the amplitudes $\bar{\mathbf{d}}_n$. At first, in Section IV-A, the batch maximum-likelihood (ML) method is introduced. It is based on the assumption that the delays (i.e., the basis $\bar{\mathbf{U}}_n$ and the rank $\bar{r}_n$) remain constant for M OFDM symbols while the amplitudes vary symbol-by-symbol in an unpredictable manner (i.e., no statistical modeling is applied to the dynamics of the amplitudes). The resulting algorithm processes the OFDM signal in blocks of M symbols. On the other hand, the subspace tracking (ST) and subspace-amplitude tracking (SAT) algorithms proceed in an adaptive symbol-by-symbol fashion, accommodating slow intersymbol variations of the basis and the rank. The ST technique (Section IV-B) is an adaptive implementation of the batch ML method that is based on subspace tracking with rank estimation. Both batch ML and ST estimate the amplitude vector on a symbol-by-symbol basis without any attempt to track its variations. On the contrary, the SAT method (Section IV-C) is built on a subspace tracker (that, as in ST, updates the estimate $\hat{\bar{\mathbf{U}}}_n$) coupled with an amplitude tracker. See Fig. 3 to have a glance of the structure of the algorithms.

As pointed out earlier, the techniques discussed in the following aim to improve the estimation accuracy on the pilot subcarriers. Interpolation is then needed to get the final estimate $\hat{\mathbf{H}}_n$ (see Section IV-D).

TABLE I
OVERVIEW OF THE SUBSPACE-BASED METHODS FOR CHANNEL ESTIMATION PROPOSED IN THIS PAPER

Method	$\bar{\mathbf{U}}_n$	$\hat{\bar{\mathbf{U}}}_n$	$\bar{r}_n$	$\hat{\bar{r}}_n$	$\bar{\mathbf{d}}_n$	$\hat{\bar{\mathbf{d}}}_n$
Batch (Sec. IV-A)	constant for $n = 1, .., M$	eigenvalue decomposition	constant for $n = 1, .., M$	MDL	fast-varying	projection
ST (Sec. IV-B)	slowly-varying	subspace tracking	slowly-varying	adaptive estimation	fast-varying	projection
SAT (Sec. IV-C)	slowly-varying	subspace tracking	slowly-varying	adaptive estimation	fast-varying	Kalman, LMS,..

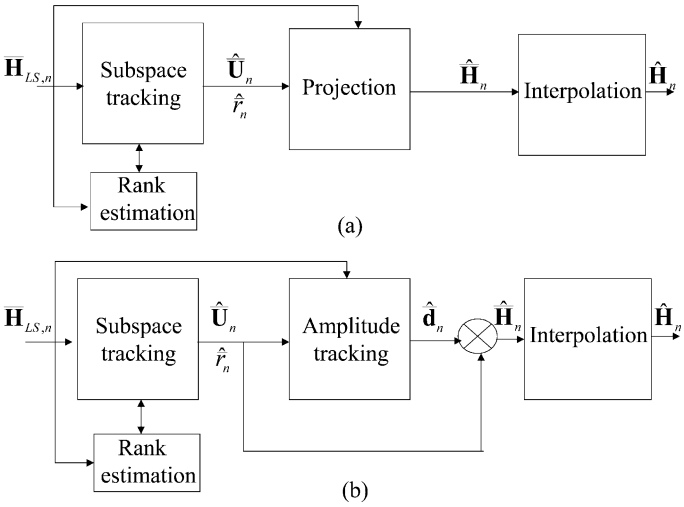


Fig. 3. Block diagrams of the (a) ST and (b) SAT algorithms.

A. Batch ML Estimator

The batch ML estimator is obtained from the observation of M OFDM symbols, where M is chosen so that $\bar{\mathbf{W}}(\boldsymbol{\tau}_n) = \bar{\mathbf{W}}(\boldsymbol{\tau})$ or equivalently $\bar{\mathbf{U}}_n = \bar{\mathbf{U}}$ and $\bar{r}_n = \bar{r}$ for $n = 1, .., M$. Apart from uninteresting constants, the negative log-likelihood function evaluates as

$$\Psi(\bar{\mathbf{U}}, \bar{\mathbf{d}}_1, .., \bar{\mathbf{d}}_M) = \left\| \left[\bar{\mathbf{H}}_{\mathrm{LS},1} \ldots \bar{\mathbf{H}}_{\mathrm{LS},M}\right] - \bar{\mathbf{U}}[\bar{\mathbf{d}}_1 \ldots \bar{\mathbf{d}}_M] \right\|^2 \tag{10}$$

where

$$\bar{\mathbf{H}}_{\mathrm{LS},n} = \mathrm{diag}(\bar{\mathbf{X}}_n)^{-1}\bar{\mathbf{Y}}_n \tag{11}$$

is the LS (or backrotated) estimate. In (10), we have assumed for simplicity that the same power is used on all the pilot subcarriers, i.e., $\mathbf{X}_n\mathbf{X}_n^H = \sigma_x^2\mathbf{I}_{\bar{N}}$. Removing this assumption would just complicate the notation because of the need to weigh the norm in (10) by the powers on different subcarriers. For identifiability, it must be $\mathrm{rank}([\bar{\mathbf{d}}_1 \ldots \bar{\mathbf{d}}_M]) = \bar{r}$. Notice that this condition is not satisfied if $M < \bar{r}$ and/or the degree of fading decorrelation in M symbols is small [24]. In practice, this observation entails the need of a rank detector for the ST algorithm. Maximization of (10) yields

$$\hat{\bar{\mathbf{d}}}_n = \hat{\bar{\mathbf{U}}}^H \bar{\mathbf{H}}_{\mathrm{LS},n} \tag{12}$$

and leads to the batch ML estimator [24]

$$\hat{\bar{\mathbf{H}}}_n = \hat{\bar{\mathbf{U}}}\hat{\bar{\mathbf{d}}}_n = \hat{\bar{\mathbf{U}}}\hat{\bar{\mathbf{U}}}^H \bar{\mathbf{H}}_{\mathrm{LS},n} \tag{13}$$

where $\hat{\bar{\mathbf{U}}}$ denotes the delay-subspace basis estimated as the span of the $\bar{r}$ largest eigenvectors of the sample correlation matrix $\mathbf{R}_h = \sum_{n=1}^{M} \bar{\mathbf{H}}_{\mathrm{LS},n}\bar{\mathbf{H}}_{\mathrm{LS},n}^H$. The batch ML estimator thus reduces to the projection of the LS estimate onto the delay-subspace estimated from $\mathbf{R}_h$. The rank $\bar{r}$ can be estimated from the correlation matrix $\mathbf{R}_h$ by using the traditional minimum description length (MDL) criterion. This algorithm has the following evident flaws:

1) latency in providing the estimate;
2) inability to accommodate intersymbol variations of the basis (the method is based on a quasistatic variation of the basis on M symbols);
3) failure to track the intersymbol variations of the amplitudes in order to improve the estimation accuracy.

Problems 1 and 2 are solved by the ST and SAT algorithm, the SAT technique provides a solution also to 3).

B. ST Channel Estimator

The ST estimator is an adaptive implementation of the batch ML estimator discussed in the previous Section. A block diagram of the algorithm is shown in Fig. 3(a). At each OFDM symbol, say the nth, the LS estimate $\bar{\mathbf{H}}_{\mathrm{LS},n}$ is fed to a subspace tracker that updates the estimate of the basis $\bar{\mathbf{U}}_n$ and of its rank $\bar{r}_n$. We recall that a subspace tracker is an algorithm able to adaptively estimate, in an approximate way, the leading eigenvectors of the correlation matrix of the input vectors, with a reduced computational burden compared to the eigenvalue decomposition (for an overview, see [25]). According to (13), the estimate $\hat{\bar{\mathbf{U}}}_n$ output by the subspace tracker is used to project $\bar{\mathbf{H}}_{\mathrm{LS},n}$ onto the delay-subspace as

$$\hat{\bar{\mathbf{H}}}_n = \hat{\bar{\mathbf{U}}}_n\hat{\bar{\mathbf{d}}}_n = \hat{\bar{\mathbf{U}}}_n\hat{\bar{\mathbf{U}}}_n^H \bar{\mathbf{H}}_{\mathrm{LS},n}. \tag{14}$$

A large number of subspace tracking algorithms (with rank estimation capability) of different computational costs and performances have been proposed in the literature. For the application discussed here, a choice that has been shown to be a good tradeoff between complexity and performance is the fast subspace tracking algorithm proposed in [17] and summarized in Table II using the notation of this paper. The notation $\mathrm{card}\{\cdot\}$

TABLE II
ST ALGORITHM

Initialize: $\bar{r}_{\max}; \mathbf{B}_0 = \begin{bmatrix} \mathbf{I}_{\bar{r}_{\max}} \\ \mathbf{0} \end{bmatrix}; \boldsymbol{\Theta}_0 = \mathbf{I}_{\bar{r}_{\max}}; \mathbf{A}_0 = \mathbf{0}; 0 \le \gamma \le 1; \bar{r}_{\max}$

For each symbol n:

input: $\bar{\mathbf{H}}_{LS,n}$

1. Subspace tracking:

$\mathbf{Z}_n = \mathbf{B}_{n-1}^H \bar{\mathbf{H}}_{LS,n}$

$\mathbf{A}_n = \gamma \mathbf{A}_{n-1} \boldsymbol{\Theta}_{n-1} + \bar{\mathbf{H}}_{LS,n}$

$\mathbf{A}_n = \mathbf{B}_n \mathbf{R}_n$ (QR factorization)

$\boldsymbol{\Theta}_n = \mathbf{B}_{n-1}^H \mathbf{B}_n$

2. Adaptive rank estimation:

$\hat{\lambda}_i = [\mathbf{R}_n]_{ii} \quad i = 1, 2, .., \bar{r}_{\max}$

$p_n = \gamma p_{n-1} + \frac{1}{\bar{N}} \operatorname{tr}\{\bar{\mathbf{H}}_{LS,n} \bar{\mathbf{H}}_{LS,n}^H\}$

$\hat{\sigma}^2 = \frac{\bar{N}}{\bar{N} - \bar{r}_{\max}} p_n - \frac{1}{\bar{N} - \bar{r}_{\max}} \operatorname{tr}\{\mathbf{R}_n\}$

$\widehat{\bar{r}}_n = \operatorname{card}\{\hat{\lambda}_i : \hat{\lambda}_i > \beta \cdot \hat{\sigma}^2\}$

updated basis: $\widehat{\bar{\mathbf{U}}}_n = [\mathbf{B}_n]_{:,1:\widehat{\bar{r}}_n}$

denotes the cardinality of the set under parentheses and the operator $[\cdot]_{:,1:k}$ selects the first k columns of its argument. Referring to [17] for further details, we point out the following.

1) The algorithm requires to set an *a priori* upper bound on $\bar{r}_n$, denoted as $\bar{r}_{\max}$ in Table II. This can be easily obtained by measurement of the radio channel of interest before the deployment of the algorithm.
2) The computational complexity of the subspace tracker (and rank estimation) is $O(\bar{N}\bar{r}_{\max}^2)$. Notice that the QR factorization in Table II can be avoided according to [17] in order to reduce the computational complexity [but retaining the same order $O(\bar{N}\bar{r}_{\max}^2)$].
3) The adaptive rank estimation requires to set the multiplicative coefficient β. A thorough analysis of the optimal selection of β as a function of the system parameters can be found in [26].

C. SAT Channel Estimator

The ST technique estimates the amplitudes separately in each OFDM symbol as $\widehat{\bar{\mathbf{d}}}_n = \widehat{\bar{\mathbf{U}}}_n^H \bar{\mathbf{H}}_{\mathrm{LS},n}$ [see (14)]. Here, we extend the ST estimator in order to include the tracking of the amplitudes variations across different symbols [SAT, see Fig. 3(b)]. In order to decouple the problems of delay-subspace basis and amplitudes estimation, the amplitude tracker is designed as if the estimate of the basis produced by the subspace tracker, $\widehat{\bar{\mathbf{U}}}_n$, was the real basis $\bar{\mathbf{U}}_n$. In Section V, we show that this assumption becomes increasingly accurate for larger n. To elaborate, from (8) and (11), we have that $\bar{\mathbf{H}}_{\mathrm{LS},n} = \bar{\mathbf{U}}_n \bar{\mathbf{d}}_n + \bar{\mathbf{Q}}_n$, where $\bar{\mathbf{Q}}_n$ denotes the additive noise on the LS estimate. Now, assuming that the basis $\bar{\mathbf{U}}_n$ is known (through its estimate $\widehat{\bar{\mathbf{U}}}_n$), the problem is that of tracking the variations of the vector $\bar{\mathbf{d}}_n$ given the observations $\bar{\mathbf{H}}_{\mathrm{LS},n}$ (linear regression).

From (9) and assuming the knowledge of the statistics of the physical amplitudes $\boldsymbol{\alpha}_n$, one could derive the statistics of the compound amplitudes $\bar{\mathbf{d}}_n$. Based on the latter, optimal (MMSE) tracking of the vector $\bar{\mathbf{d}}_n$ could be implemented by means of the classical Kalman filter [27]. Alternatively, the suboptimal techniques proposed in [28] reduce the computational complexity of the Kalman filter with minor performance losses.

For simplicity, here we limit our analysis to the LMS algorithm. The derivation is based on the subspace tracker proposed in [17] and summarized in Table II. Let $\mathbf{B}_n$ be the $\bar{N} \times \bar{r}_{\max}$ orthonormal matrix produced by the delay-subspace tracker (see Table II, notice that the first $\widehat{\bar{r}}_n$ columns of $\mathbf{B}_n$ give $\widehat{\bar{\mathbf{U}}}_n$), the estimate of the amplitudes $\widehat{\bar{\mathbf{d}}}_n$ can be, thus, updated as follows ($0 < \mu < 2$ for stability since $E[\mathbf{B}_n^H \mathbf{B}_n] = \mathbf{I}$ [27])

$$\boldsymbol{\varepsilon}_n = \bar{\mathbf{H}}_{\mathrm{LS},n} - \mathbf{B}_n \mathbf{b}_{n-1} \tag{15a}$$

$$\mathbf{b}_n = \mathbf{b}_{n-1} + \mu \mathbf{B}_n^H \boldsymbol{\varepsilon}_n \tag{15b}$$

$$\widehat{\bar{\mathbf{d}}}_n = [\mathbf{b}_n]_{1:\widehat{\bar{r}}_n} \tag{15c}$$

where $[\cdot]_{1:r}$ denotes the $r \times 1$ vector containing the first r entries of its argument. The estimate of the channel vector is then obtained as

$$\widehat{\bar{\mathbf{H}}}_n = \widehat{\bar{\mathbf{U}}}_n \widehat{\bar{\mathbf{d}}}_n. \tag{16}$$

Notice that because of the structure of the subspace tracker employed here, $\bar{r}_{\max}$ amplitudes are tracked even if only the first $\widehat{\bar{r}}_n$ are needed for the channel estimate (16). The initial condition for the amplitudes can be set to zero: $\mathbf{b}_0 = 0$.

D. Remarks on Channel Interpolation

Once we have obtained the estimate of the channel on the $\bar{N}$ pilot subcarriers $\widehat{\bar{\mathbf{H}}}_n$ by any algorithm (LS, ST, or SAT), it is necessary to perform interpolation (and possibly smoothing) in frequency and time to obtain the final $N \times 1$ estimate $\hat{\mathbf{H}}_n$. In this paper, we consider for simplicity the case of a comb pilot pattern (Fig. 1) so that we can limit the analysis to a one-dimensional (1-D) (or frequency only) interpolator. Any (linear) interpolator/smoother can be described by a $N \times \bar{N}$ matrix $\mathbf{T}$

$$\hat{\mathbf{H}}_n = \mathbf{T} \widehat{\bar{\mathbf{H}}}_n. \tag{17}$$

The optimum solution for $\mathbf{T}$ in the MMSE sense is the Wiener filter $\mathbf{T} = \mathbf{R}_{\mathbf{H}\bar{\mathbf{H}}} (\mathbf{R}_{\bar{\mathbf{H}}\bar{\mathbf{H}}} + 1/\mathrm{SNR} \cdot \mathbf{I}_{\bar{N}})^{-1}$, where $\mathbf{R}_{\mathbf{H}\bar{\mathbf{H}}} = E[\mathbf{H}_n \bar{\mathbf{H}}_n^H]$ and $\mathbf{R}_{\bar{\mathbf{H}}\bar{\mathbf{H}}} = E[\bar{\mathbf{H}}_n \bar{\mathbf{H}}_n^H]$. The design of $\mathbf{T}$ requires the knowledge of both the *SNR* and the frequency-time correlation function of the channel. When this information is missing, or it is not reliable enough, the most adequate (robust) design choice is to consider $\mathrm{SNR} \to \infty$ and a uniform power-delay profile in [0, T_{cp}] [15]. Alternatively, and almost equivalently, we can use efficient interpolation techniques based on the pair IFFT/FFT that constrain the energy of the temporal response to be supported within D samples [13], [14].

E. Extension to MIMO Systems

The algorithms described in the previous sections can be easily extended to a radio link characterized by multiple antennas at both the transmitter and the receiver (MIMO system). In fact, by employing the pilot pattern described in [13], the task of channel estimation can be performed separately for each pair transmit–receive antenna. Since a quasi-synchronous

MC-CDMA system can be regarded as a MIMO-OFDM system in which each transmit antenna is an independent user [29], the subspace-based channel estimation methods are applicable to this modulation scheme as well. As opposed to generic MIMO-OFDM systems, in MC-CDMA the data carriers contain the superposition of the information streams spread in frequency by each user.

V. MSE Performance Analysis

The channel on the entire set of N used subcarriers, i.e., the $N \times 1$ vector $\hat{\mathbf{H}}_n$, is obtained by interpolating the $\bar{N}$ estimates from the pilots $\widehat{\bar{\mathbf{H}}}_n$ according to (17). Since the channel estimate over the pilots is affected by the error $\bar{\mathbf{Q}}_n$

$$\widehat{\bar{\mathbf{H}}}_n = \bar{\mathbf{H}}_n + \bar{\mathbf{Q}}_n \tag{18}$$

the mean square error (MSE) of the overall estimate can be evaluated as

$$\mathrm{MSE} = E[\|\mathbf{H}_n - \hat{\mathbf{H}}_n\|^2] = \mathrm{MSE}_I + \mathrm{MSE}_Q \tag{19}$$

where the terms

$$\mathrm{MSE}_I = E[\|\mathbf{H}_n - \mathbf{T}\bar{\mathbf{H}}_n\|^2] \tag{20a}$$

$$\mathrm{MSE}_Q = E[\|\mathbf{T}\bar{\mathbf{Q}}_n\|^2] - 2\mathrm{Re}[\mathrm{tr}(\mathbf{T}E[\bar{\mathbf{Q}}_n\mathbf{H}_n^H])] - 2\mathrm{Re}[\mathrm{tr}(\mathbf{T}E[\bar{\mathbf{Q}}_n\bar{\mathbf{H}}_n^H]\mathbf{T}^H)] \tag{20b}$$

depend on the interpolation error and on the inaccuracy of the estimate over the pilots, respectively. The factorization (19) can be easily proved by substitution. The MSE_I (20a) accounts for the bias on the estimate due to interpolation (e.g., aliasing in the temporal domain [14]) and it sets an error floor on the *MSE* that can not be reduced by improving the *SNR* but only by appropriately designing $\mathbf{T}$ or by increasing the number of pilot subcarriers $\bar{N}$. The literature on channel estimation for OFDM systems has mainly focused on reducing MSE_I while in this paper, we are concerned with the problem of reducing MSE_Q.

Let us now compare MSE_Q for the LS estimate and the ST algorithm. We emphasize that in the case of a static delay-pattern within the observation interval, that is the scenario of interest of this section, the performance of the ST technique coincides with that of the batch ML algorithm. For the LS estimate the estimate error in (18) is

$$\bar{\mathbf{Q}}_{n,\mathrm{LS}} = \mathrm{diag}(\bar{\mathbf{X}}_n)^{-1}\bar{\mathbf{N}}_n \tag{21}$$

so that

$$\mathrm{MSE}_{Q,\mathrm{LS}} = \frac{\|\mathbf{T}\|^2}{\mathrm{SNR}} \tag{22}$$

where we used the fact that $E[\bar{\mathbf{N}}_n\mathbf{H}_n^H] = \mathbf{0}$ and $E[\bar{\mathbf{Q}}_n\bar{\mathbf{H}}_n^H] = \mathbf{0}$. In order to evaluate MSE_Q for the ST technique, we assume that the delays are stationary over a very large number of symbols ($\bar{\mathbf{U}}_n = \bar{\mathbf{U}}$ and $\bar{r}_n = \bar{r}$) so that the asymptotic performance for $n \to \infty$ can be evaluated. Moreover, the rank is assumed to be correctly estimated, $\hat{\bar{r}}_n = \bar{r}$. The estimator of the basis is consistent [24] and, thus, $\hat{\bar{\mathbf{U}}}_n \to \bar{\mathbf{U}}$ for $n \to \infty$. The symbol-by-symbol estimate (13) is obtained by projecting the LS estimate onto the delay-subspace so that asymptotically ($n \to \infty$)

$$\hat{\bar{\mathbf{H}}}_n \to \bar{\mathbf{U}}\bar{\mathbf{U}}^H\bar{\mathbf{H}}_{\mathrm{LS},n} = \bar{\mathbf{U}}(\bar{\mathbf{d}}_n + \bar{\mathbf{U}}^H\bar{\mathbf{Q}}_{n,\mathrm{LS}}) \tag{23}$$

the error

$$\bar{\mathbf{Q}}_{n,\mathrm{ST}} \to \bar{\mathbf{U}}\bar{\mathbf{U}}^H\bar{\mathbf{Q}}_{n,\mathrm{LS}} \tag{24}$$

is reduced with respect to $\bar{\mathbf{Q}}_{n,\mathrm{LS}}$ by the projection onto a reduced-size subspace, as it will be discussed in the following. Furthermore, from (23) and (24), it is clear that the error $\bar{\mathbf{Q}}_{n,\mathrm{ST}}$ asymptotically depends only on the error relative to the estimate of the amplitudes $\bar{\mathbf{U}}^H\bar{\mathbf{Q}}_{n,\mathrm{LS}}$. We can conclude that asymptotically ($n \to \infty$)

$$\mathrm{MSE}_{Q,\mathrm{ST}} \to \frac{E[\|\mathbf{T}\bar{\mathbf{U}}\|^2]}{\mathrm{SNR}} = \frac{\mathrm{tr}(\mathbf{T}E[\bar{\mathbf{U}}\bar{\mathbf{U}}^H]\mathbf{T}^H)}{\mathrm{SNR}} \tag{25}$$

where we recall that the operator $E[\cdot]$ performs the average over the distribution (of the noise, not relevant here, and) of the channel. In particular, (24) describes the performance of the ST algorithm averaged over the distribution of the delays, since the basis $\bar{\mathbf{U}}$ is independent of the amplitudes. Inspection of (22) and (25) suggests that the ST algorithm reduces MSE_Q through the projection onto the delay-subspace. In fact, $E[\|\mathbf{T}\bar{\mathbf{U}}\|^2]$ depends only on the fraction $\bar{r}/\bar{N}$ of the nonzero eigenvalues of $\mathbf{T}\mathbf{T}^H$ that lie in the delay-subspace. Therefore, even though it is not easy to explicitly derive this gain due the interaction in (25) of the interpolator matrix $\mathbf{T}$ and the delay-subspace basis $\bar{\mathbf{U}}$, we expect an MSE improvement in the order of $\bar{N}/\bar{r}$ (see Section VI). In the following, we show that it is exactly the asymptotic gain that we get by considering the mean square error on the pilot subcarriers. Before addressing this subject, let us mention that the SAT algorithm is able to further lower the value of MSE_Q by improving the accuracy of the estimate of the amplitudes. The expected gain depends on the Doppler shift as slower amplitude variations (i.e., smaller Doppler shifts) allow better tracking performance. An analytic study is beyond the scope of this paper and it will not be pursued here.

The mean square error evaluated only on the pilot carriers $\overline{\mathrm{MSE}} = E[\|\bar{\mathbf{H}}_n - \widehat{\bar{\mathbf{H}}}_n\|^2]$ can be compared for the LS and the ST technique. By noting that $\overline{\mathrm{MSE}} = E[\mathrm{tr}(\bar{\mathbf{Q}}_n\bar{\mathbf{Q}}_n^H)]$ and recalling (21) and (24), we can easily conclude that

$$\overline{\mathrm{MSE}}_{\mathrm{LS}} = E[\mathrm{tr}(\bar{\mathbf{Q}}_{n,\mathrm{LS}}\bar{\mathbf{Q}}_{n,\mathrm{LS}}^H)] = \frac{\bar{N}}{\mathrm{SNR}} \tag{26}$$

and

$$\overline{\mathrm{MSE}}_{\mathrm{ST}} = E[\mathrm{tr}(\bar{\mathbf{Q}}_{n,\mathrm{ST}}\bar{\mathbf{Q}}_{n,\mathrm{ST}}^H)] \to \frac{\bar{r}}{\mathrm{SNR}}. \tag{27}$$

Therefore, the asymptotic gain of the ST with respect to the LS algorithm in terms of $\overline{\mathrm{MSE}}_Q$ is $\bar{N}/\bar{r}$. A bound on $\overline{\mathrm{MSE}}_{\mathrm{ST}}$ can be derived for any number of symbols n (see the Appendix).

VI. Simulation Results

The performance of the proposed subspace channel estimators is evaluated for an OFDM system having a bandwidth $B =$ 800 KHz, $N_t = 128$ subcarriers, $N = 120$, $D = 16$, $\bar{N} =$ 15 equispaced pilot subcarriers ($k = N_t/D \cdot m$ with $m = 1, 2, .., D-1$, notice that $k = 0$ falls in the guard band) and $\sigma_x^2 = 1$. The Doppler is modeled according to the Clarke's model, the maximum Doppler shift is assumed to be $f_D = 500$ Hz so that, by using the theoretical results of [21], we can conclude that the signal-to-intercarrier interference ratio due to the Doppler effect

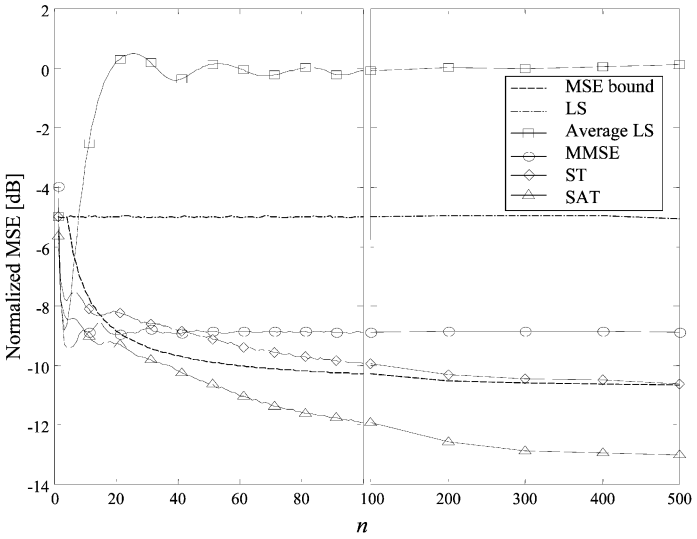


Fig. 4. Normalized $\overline{\mathrm{MSE}}$ versus the number of blocks n of LS (dashed lines), multiblock average of LS estimates, MMSE-FIR filtering of the LS estimates, ST, and SAT for $f_D = 200$ Hz. Also shown as a reference is the MSE bound for ST computed in the Appendix . The left side shows the simulation results with adaptive estimation of the rank $\bar{r}_n$ whereas for the right side the rank is set to $\hat{\bar{r}}_n = \bar{r} = 4$.

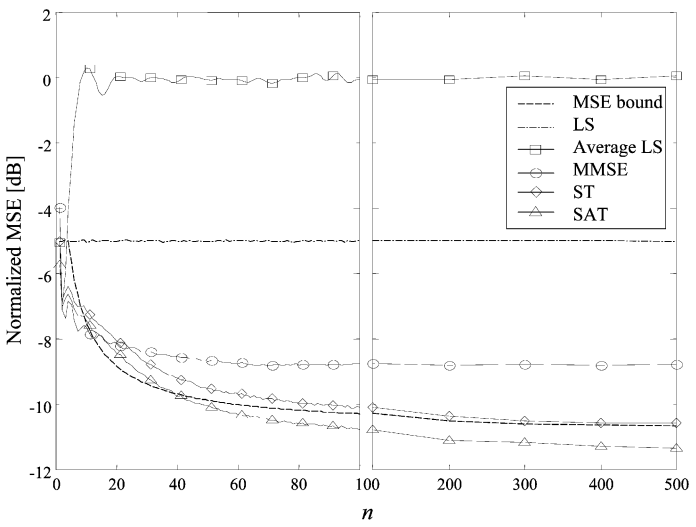


Fig. 5. Normalized $\overline{\mathrm{MSE}}$ versus the number of blocks n of LS (dashed lines), multiblock average of LS estimates, MMSE-FIR filtering of the LS estimates, ST, and SAT for $f_D = 500$ Hz. Also shown as a reference is the MSE bound for ST computed in the Appendix . The left side shows the simulation results with adaptive estimation of the rank $\bar{r}_n$, whereas for the right side, the rank is set to $\hat{\bar{r}}_n = \bar{r} = 4$.

is (tightly) lower bounded by 19 dB. It follows that the assumption of quasistatic variations of the amplitudes from symbol to symbol that we used to derive the model (2) holds true if the *SNR* is sufficiently smaller that 19 dB. The channel has four i.i.d. delays with exponential distribution $p(\tau) = 1/\tau_o \exp(-\tau/\tau_o)$ with $\tau_o = 4.4$ μs (this guarantees that $\tau_{n,\ell} \leq T_{cp} + T - T_g$ with probability close to one). At each simulation run, the delays are selected independently and the dimension of the delay-subspace is $\bar{r} = 4$ (with probability close to one). The parameters of the subspace tracking algorithm (Table II) have been set to $\gamma = 0.999$, $\bar{r}_{\max} = 8$ and $\beta = 1$; furthermore, $\mu = 0.6$ in the LMS update (15b). Notice that the parameters have been selected without any attempt of optimization as a function of the system (see [26] for a discussion on the optimal selection of β).

The influence of the number of symbols for a static delay-pattern is considered in Figs. 4–6 by evaluating the $\overline{\mathrm{MSE}}$ over the pilot subcarriers (Figs. 4 and 5) and the estimated rank $\hat{\bar{r}}_n$ (Fig. 6) (averaged over 10^4 simulation runs) versus the number of processed symbols n for $\mathrm{SNR} = 5$ dB and the Doppler shifts $f_D = 200$ (Fig. 4) and 500 Hz (Fig. 5). The performance of the ST and SAT algorithms are compared with the LS estimate $\bar{\mathbf{H}}_{\mathrm{LS},n}$ and with two alternative multisymbol preprocessing of the LS estimate: the average $1/n \sum_{i=1}^{n} \bar{\mathbf{H}}_{\mathrm{LS},n}$ and a 64-taps MMSE-FIR filtering carried out according to [15] with parameter $K_o = 15$ ([15, Section IV-B]). Notice that the latter algorithm has the same order of computational complexity as the proposed techniques. As a reference, the MSE bound for the ST technique derived in Appendix is also shown in Figs. 4 and 5.

Let us first consider the left part of Figs. 4 and 5. Compared to the LS estimate, ST guarantees a gain of 5 dB for $n = 100$ irrespective of the Doppler frequency while the SAT improves this margin by 2 dB for $f_D = 200$ Hz and 0.5 dB for $f_D = 500$ Hz. As expected, the tracking of the amplitudes becomes less advantageous for increasing Doppler shifts (at least if we assume lack of knowledge of the fading statistics). The right parts of Figs. 4 and 5 are

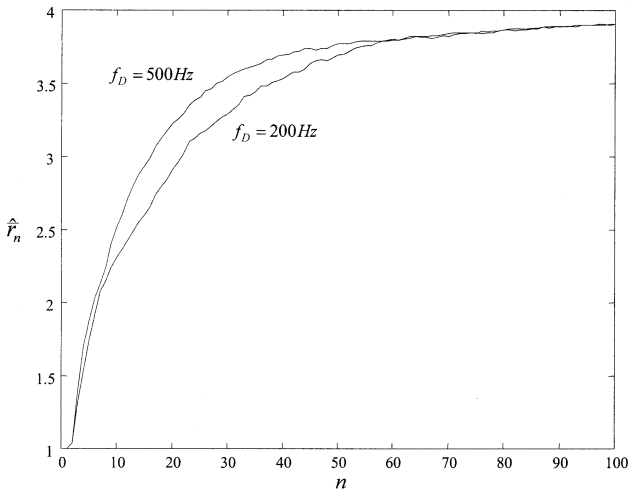


Fig. 6. Adaptive estimate of the rank $\hat{\bar{r}}_n$ for $f_D = 200$ and $f_D = 500$ Hz.

not relevant from the standpoint of practical implementation (in practice a memory of 20–30 symbols is enough for the proposed methods) but it has been added in order to show the convergence of the $\overline{\mathrm{MSE}}_{\mathrm{ST}}$ to the MSE bound computed in the Appendix . Notice that for $n \geq 100$ the rank has been fixed to $\hat{\bar{r}}_n = \bar{r} = 4$ since, without an adjustment of the factor β according to [26], the adaptive rank estimation tends to overestimate the rank for very large n with a slight loss of performance (i.e., below 0.2 *dB* for the parameters chosen here). Fig. 6 shows the estimated rank $\hat{\bar{r}}_n$ corresponding to the simulations of Figs. 4 and 5. It can be seen that higher Doppler shifts entail the need of a larger rank.

According to the discussion in Section V, here the combined effect on the *MSE* over the entire set of N subcarriers of the interpolation error MSE_I (20a) and the inaccuracy of the estimate on the pilot subcarriers MSE_Q (20b) is analyzed numerically. Let us consider a MMSE interpolator based on a uniform power delay profile in [0, T_{cp}] according to the "robust" design of [15], a Doppler shift $f_D = 200$ Hz and let the delay-pattern be

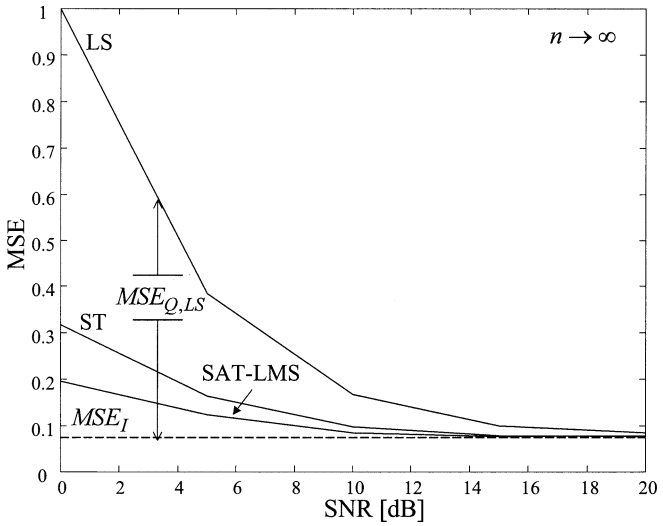


Fig. 7. *MSE* versus *SNR* for different estimators compared to the error floor set by the interpolation error.

stationary over a very large number of OFDM symbols so that we can rely on the asymptotic analysis for $n \to \infty$ (Section V). Fig. 7 compares the *MSE* of different estimators versus *SNR*. We notice that for high values of *SNR* ($\mathrm{SNR} > 10$ dB) the error floor due to interpolation becomes the predominant term, i.e., $\mathrm{MSE} \simeq \mathrm{MSE}_I$. In this case, although the preprocessing carried out by the ST and the SAT still guarantees the same asymptotic gain in terms of MSE_Q, that is approximately $\bar{N}/\bar{r}$ for ST, the benefits of deploying an intermediate step between the LS estimate and the interpolator become negligible. Of course, the results shown in Fig. 7 depend on the interpolator and any improvement in the interpolation step would extend the *SNR* interval over which the our subspace methods are beneficial.

VII. Conclusion

Novel subspace-based channel estimation techniques for OFDM systems over fast-fading channels have been proposed. The methods are designed for the comb pilot pattern (i.e., each OFDM symbol carries equispaced pilot subcarriers) and perform a multisymbol preprocessing of the LS estimate over the pilot subcarriers. The key is to capitalize on the different varying rate of the parameters of the multipath channel (delays and amplitudes) by performing an unstructured tracking of the slow variations of the delays through a subspace tracking technique. The performance has been evaluated in terms of the MSE on the channel estimate both analytically and through simulations, demonstrating the considerable benefits of the proposed algorithms.

Appendix

Here, we compute a lower bound on the mean square error of the ST estimate over the pilot subcarriers, i.e., on $\overline{\mathrm{MSE}}_{\mathrm{ST}} = E[\|\bar{\mathbf{H}}_n - \hat{\bar{\mathbf{H}}}_n\|^2]$. The bound is obtained by computing the Cramer–Rao lower bound (CRLB) on the variance of the estimate of an unbiased estimator [30], i.e., on $E_{\mathbf{N}}[\|\bar{\mathbf{H}}_n - \hat{\bar{\mathbf{H}}}_n\|^2]$ (where $E_{\mathbf{N}}[\cdot]$ denotes the average over the distribution of noise only) and then averaging it with respect to the statistics of the channel (see Section II). To make the analysis feasible, we assume that the delays are constant within the observation interval, i.e., $\bar{\mathbf{W}}(\boldsymbol{\tau}_n) = \bar{\mathbf{W}}(\boldsymbol{\tau})$ for $\forall n = 1, .., M$ (or equivalently $\bar{\mathbf{U}}_n = \bar{\mathbf{U}}$ and $\bar{r}_n = \bar{r}$). As we pointed out in Section V, under this condition the performance of ST coincides with that of the batch algorithm presented in Section IV-A.

From the expression of the likelihood function (10), we infer that the collection of the LS estimates relative to M OFDM symbols, $\bar{\mathbf{H}}_{\mathrm{LS}}(M) = [\bar{\mathbf{H}}_{\mathrm{LS},1}^T \dots \bar{\mathbf{H}}_{\mathrm{LS},M}^T]^T$, represents a sufficient statistic for the estimate of the channel. In addition, we recall that the ST method is based a parametrization of the channel that reduces the parameters to be estimated to the vector $\boldsymbol{\theta} = [\mathrm{vec}(\bar{\mathbf{U}})^T \ \mathrm{vec}(\bar{\mathbf{D}})^T]^T$, where $\bar{\mathbf{D}} = [\bar{\mathbf{d}}_1 \dots \bar{\mathbf{d}}_M]$. Now, since $\bar{\mathbf{H}}_{\mathrm{LS}}(M)$ is circularly symmetric Gaussian with mean $\mathrm{vec}(\bar{\mathbf{U}}\bar{\mathbf{D}})$ and covariance $1/\mathrm{SNR} \cdot \mathbf{I}_{MN_p}$, the CRLB on the estimate of $\bar{\mathbf{H}}_n$ is [31] (we average over the M symbols)

$$E_{\mathbf{N}}[\|\bar{\mathbf{H}}_n - \hat{\bar{\mathbf{H}}}_n\|^2] \geq \frac{1}{M \cdot \mathrm{SNR}} \mathrm{tr}\{[\mathbf{C}(\mathbf{C}^H\mathbf{C})^{-1}\mathbf{C}^H]\} \quad (28)$$

where

$$\mathbf{C} = \left\{ \frac{\partial \mathrm{vec}(\bar{\mathbf{U}}\bar{\mathbf{D}})}{\partial \boldsymbol{\theta}^T} \right\}_{\boldsymbol{\theta}=\boldsymbol{\theta}_o} \quad (29)$$

with the subscript "o" denoting hereafter the "true" value of the corresponding parameter. The lower bound on $\overline{\mathrm{MSE}}_{\mathrm{ST}}$ is thus computed as

$$\begin{aligned} \overline{\mathrm{MSE}}_{\mathrm{ST}} &= E_{\boldsymbol{\theta}_o}[E_{\mathbf{N}}[\|\bar{\mathbf{H}}_n - \hat{\bar{\mathbf{H}}}_n\|^2]] \\ &\geq \frac{1}{M \cdot \mathrm{SNR}} \cdot E_{\boldsymbol{\theta}_o}[\mathrm{tr}\{[\mathbf{C}(\mathbf{C}^H\mathbf{C})^{-1}\mathbf{C}^H]\}]. \end{aligned} \quad (30)$$

Notice that in (30), we exploited the fact that the expectation with respect to the distribution of the channel can be equivalently obtained by averaging with respect to the parameter vector $\boldsymbol{\theta}_o$ [recall (8)]. The sensitivity matrix $\mathbf{C}$ is easily computed

$$\mathbf{C} = [\bar{\mathbf{D}}_o^T \otimes \mathbf{I}_N | \mathbf{I}_M \otimes \bar{\mathbf{U}}_o]. \quad (31)$$

Now, notice that it is

$$\begin{aligned} \mathrm{tr}\{[\mathbf{C}(\mathbf{C}^H\mathbf{C})^{-1}\mathbf{C}^H]\} &= \mathrm{rank}\{[\mathbf{C}(\mathbf{C}^H\mathbf{C})^{-1}\mathbf{C}^H]\} \\ &= \mathrm{rank}\{\mathbf{C}^H\mathbf{C}\} \end{aligned} \quad (32)$$

where the first equality stems from the observation that $\mathbf{C}(\mathbf{C}^H\mathbf{C})^{-1}\mathbf{C}^H$ is a projection matrix. Furthermore, it is

$$\mathbf{C}^H\mathbf{C} = \begin{bmatrix} \bar{\mathbf{D}}_o^* \bar{\mathbf{D}}_o^T \otimes \mathbf{I}_N & \bar{\mathbf{D}}_o^* \otimes \bar{\mathbf{U}}_o \\ \bar{\mathbf{D}}_o^T \otimes \bar{\mathbf{U}}_o^H & \mathbf{I}_M \otimes \bar{\mathbf{U}}_o^H \bar{\mathbf{U}}_o \end{bmatrix} \quad (33)$$

so that by using the standard result on the rank of a partitioned matrix [32], it can be shown that $\mathrm{rank}(\mathbf{C}^H\mathbf{C}) = \bar{r}(\bar{N} + M) - \bar{r}^2$, independent of $\boldsymbol{\theta}_o$ (for a known value of $\bar{r}$). From (30), we conclude that

$$\overline{\mathrm{MSE}}_{\mathrm{ST}} \geq \frac{1}{M \cdot \mathrm{SNR}} \cdot (\bar{r}(\bar{N} + M) - \bar{r}^2). \quad (34)$$

Notice that for $M \to \infty$ we obtain the result in (27).

REFERENCES

[1] D. J. Goodman, J. Borras, and N. B. Mandayam, "Infostations: A new system model for data and messaging services," in *Proc. IEEE VTC'97*, vol. 2, May 1997, pp. 969–973.

[2] *Part11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: High-Speed Physical Layer in the 5 GHz Band*, IEEE Std 802.11a-1999.

[3] M. Hsieh and C. Wei, "Channel estimation for OFDM systems based on comb-type pilot arrangement in frequency selective fading channels," *IEEE Trans. Consumer Electron.*, vol. 44, pp. 217–225, Feb. 1998.

[4] S. Coleri, M. Ergen, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," *IEEE Trans. Broadcast.*, vol. 48, pp. 223–229, Sept. 2002.

[5] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM system," *IEEE Trans. Consumer Electron.*, vol. 44, pp. 1122–1128, Aug. 1998.

[6] S. Adireddy, L. Tong, and H. Viswanathan, "Optimal placement of training for frequency-selective block-fading channels," *IEEE Trans. Inform. Theory*, vol. 48, pp. 2338–2353, Aug. 2002.

[7] B. Lu, X. Wang, and K. R. Narayanan, "LDPC-based space-time coded OFDM systems over correlated fading channels: Performance analysis and receiver design," *IEEE Trans. Commun.*, vol. 50, pp. 74–88, Jan. 2002.

[8] E. Jaffrot and M. Siala, "Turbo channel estimation for OFDM systems on highly time and frequency selectiv e channels," in *Proc. Int. Conf. Acoustics Speech and Signal Processing (ICASSP)*, 2000, pp. 2977–2980.

[9] P. A. Bello, "Characterization of randomly time-variant linear channel," *IEEE Trans. Circuits Syst.*, vol. CAS-11, pp. 360–393, Dec. 1963.

[10] P. Hoeher, S. Kaiser, and I. Robertson, "Two-dimensional pilot-symbol-aided channel estimation by wiener filtering," in *Proc. Int. Conf. Acoustics Speech and Signal Processing (ICASSP)*, 1997, pp. 1845–1848.

[11] J.-J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Borjesson, "On channel estimation in OFDM systems," in *Proc. IEEE VTC'96*, Nov. 1996, pp. 815–819.

[12] O. Edfors, M. Sandell, J. van de Beek, S. K. Wilson, and P. O. Borjesson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Commun.*, vol. 46, pp. 931–939, July 1998.

[13] V. K. Jones and G. G. Raileigh, "Channel estimation for wireless OFDM systems," in *Proc. IEEE GLOBECOM'98*, 1998, pp. 980–985.

[14] B. Yang, Z. Cao, and K. B. Letaief, "Analysis of low-complexity windowed DFT-based MMSE channel estimator for OFDM systems," *IEEE Trans. Commun.*, vol. 49, pp. 1977–1987, Nov. 2001.

[15] Y. Li, L. J. Cimini, and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading," *IEEE Trans. Commun.*, vol. 46, pp. 902–915, July 1998.

[16] B. Yang, K. B. Letaief, R. S. Cheng, and Z. Cao, "Channel estimation for OFDM transmission in multipath fading channels based on parametric channel modeling," *IEEE Trans. Commun.*, vol. 49, pp. 467–478, Mar. 2001.

[17] P. Strobach, "Low-rank adaptive filters," *IEEE Trans. Signal Processing*, vol. 44, pp. 2932–2947, Dec. 1996.

[18] H. Sari, G. Karam, and I. Jeanclaude, "Transmission techniques for digital terrestrial TV broadcasting," *IEEE Commun. Mag.*, vol. 33, pp. 100–109, Feb. 1995.

[19] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2619–2692, Oct. 1998.

[20] A. Van der Veen, M. C. Vanderveen, and A. Paulraj, "Joint angle and delay estimation using shift-invariance techniques," *IEEE Trans. Signal Processing*, vol. 46, pp. 405–418, Feb. 1998.

[21] Y. Li and L. J. Cimini, "Bounds on the interchannel interference of OFDM in time-varying impairments," *IEEE Trans. Commun.*, vol. 49, pp. 401–404, Mar. 2001.

[22] A. L. Swindlehurst, "Time delay and spatial signature estimation using known asynchronous signals," *IEEE Trans. Signal Processing*, vol. 46, pp. 449–462, Feb. 1998.

[23] C. Komninakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, "Multi-input multi-output fading channel tracking and equalization using Kalman estimation," *IEEE Trans. Signal Processing*, vol. 50, pp. 1065–1076, May 2002.

[24] O. Simeone and U. Spagnolini, "Multi-slot estimation of space-time channel," in *Proc. IEEE Int. Conf. Communications (ICC)*, May 2002, pp. 802–806.

[25] P. Comon and G. H. Golub, "Tracking a few extreme singular values and vectors in signal processing," *Proc. IEEE*, vol. 78, pp. 1327–1343, Aug. 1990.

[26] A. Kavcic and B. Yang, "Adaptive rank estimation for spherical subspace tracker," *IEEE Trans. Signal Processing*, vol. 44, pp. 1573–1579, June 1996.

[27] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1986.

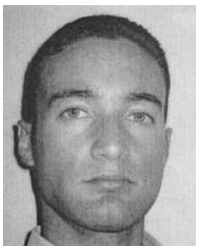
[28] L. Lindbom, M. Sternad, and A. Ahlén, "Tracking of the time-varying mobile radio channels—Part I: The Wiener LMS algorithm," *IEEE Trans. Commun.*, vol. 49, pp. 2207–2217, Dec. 2001.

[29] S. Hara and R. Prasad, "Design and performance of multicarrier CDMA system in frequency-selective Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 48, pp. 1584–1595, Sept. 1999.

[30] S. M. Kay, *Fundamentals of Statistical Signal Processing, Estimation Theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1993.

[31] P. Stoica and T. L. Marzetta, "Parameter estimation problems with singular information matrices," *IEEE Trans. Signal Processing*, vol. 49, pp. 87–90, Jan. 2001.

[32] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: The Johns Hopkins Univ. Press, 1996.



Osvaldo Simeone (S'02) received the M.Sc. degree (with honors) from the Politecnico di Milano, Milan, Italy, in 2001, where he is currently working toward the Ph.D. degree.

From February to September 2002, he was a Visiting Researcher at the Center for Communications and Signal Processing Research, New Jersey Institute of Technology, Newark. He holds a patent on the work developed for his M.S. thesis. His current research interests lie in the field of signal processing for digital communications, with emphasis on MIMO systems, multicarrier modulation, and channel estimation.

Yeheskel Bar-Ness (M'69–SM'78–F'89) received the B.Sc. and the M.Sc. degrees in electrical engineering from the Technion, Haifa, Israel, and the Ph.D. degree in applied mathematics from Brown University, Providence, RI.

In 1973, he joined the School of Engineering, Tel-Aviv University, Tel-Aviv, Israel, where he held the position of Associate Professor of Control and Communications. Between September 1978 and September 1979, he was a Visiting Professor with the Department of Applied Mathematics, Brown University. He was on leave with the University of Pennsylvania, Philadelphia, and Drexel University, Philadelphia, PA. He came to the New Jersey Institute of Technology (NJIT), Newark, from AT&T Bell Laboratories, Holmdel, NJ, in 1985. Between September 1993 and August 1994, he was on sabbatical with the Telecommunications and Traffic Control Systems Group, Faculty of Electrical Engineering, Delft University of Technology, Delft, The Netherlands. Between September 2000 and August 2001, he was on sabbatical at Stanford University, Stanford, CA. Currently, he is a Distinguished Professor of Electrical and Computer Engineering and Foundation Chair of Communication and Signal Processing Research at NJIT. He is also the Executive Director of the Center for Communication and Signal Processing Research (CCSPR), NJIT. He worked for Rafael Armament Development Authority, Haifa, Israel, in the field of communications and control and for the Nuclear Medicine Department, Elscint Ltd., Haifa, as a Chief Engineer in the field of control, image, and data processing. His current research interests include adaptive multiuser detection, array processing and interference cancellation, and wireless mobile and personal communications. He published numerous papers in these areas. Currently, he is an Editor for *Wireless Personal Communications*. He also was an Editor for *Adaptive Processing Systems*.

Dr. Bar-Ness was the Founder and Editor-in-Chief for the IEEE COMMUNICATIONS LETTERS and an Area Editor for IEEE TRANSACTIONS ON COMMUNICATIONS (Transmission Systems). He was a Chairman of the Communication Systems Committee, and the Vice Chair of the Communications Theory Committee of the IEEE Communications Society. He served as the General Chair of the 1994 and 1999 Communication Theory Mini-Conference. He was also the Technical Chair for the IEEE Sixth International Symposium on Spread Spectrum Techniques and Applications (ISSSTA 2000). He was a recipient of the Kaplan Prize, which is awarded annually by the government of Israel to the ten best technical contributors, in 1973.

Umberto Spagnolini (M'99–SM'03) received the Dott. Ing. Elett. degree (*cum laude*) in telecommunications from the Politenico di Milano, Milan, Italy, in 1988.

Since 1988, he has been with the Dipartimento di Elettronica e Informazione, Politenico di Milano, where he has held the position of Associate Professor of digital signal processing since 1998. His research interests include array processing and wavefield interpolation (mobile communication and geophysics), inverse problems (ground penetrating radar), parameter estimation (2-D phase unwrapping for SAR), and non-Gaussian EMI reduction.

Dr. Spagnolini is an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and a member of the Society of Exploration Geophysicists (SEG) and European Association of Geoscientists and Engineers (EAGE). He was awarded the Associazione Elettronica Italiana (AEI) Award and the Van Weelden Award of EAGE, both in 1991, and received the Best Paper Award from EAGE in 1998.