Subspace-Based Blind and Semi-Blind Channel Estimation for OFDM Systems

Bertrand Muquet, *Member, IEEE*, Marc de Courville, *Member, IEEE*, and Pierre Duhamel, *Fellow, IEEE*

Abstract—This paper proposes a new blind channel estimation method for orthogonal frequency division multiplexing (OFDM) systems. The algorithm makes use of the redundancy introduced by the cyclic prefix to identify the channel based on a subspace approach. Thus, the proposed method does not require any modification of the transmitter and applies to most existing OFDM systems. Semi-blind procedures taking advantage of training data are also proposed. These can be training symbols or pilot tones, the latter being used for solving the intrinsic indetermination of blind channel estimation. Identifiability results are provided, showing that in the (theoretical) situation where channel zeros are located on subcarriers, the algorithm does not ensure uniqueness of the channel estimation, unless the full noise subspace is considered. Simulations comparing the proposed method with a decision-directed channel estimator finally illustrates the performance of the proposed algorithm.

Index Terms—Blind, channel estimation, HIPERLAN/2, IEEE802.11a, OFDM, semi-blind, subspace.

I. INTRODUCTION

MULTICARRIER systems, and especially orthogonal frequency division multiplexing (OFDM), are considered today to be a reliable choice for high rate transmissions and are now widely adopted and tested in many communication systems. Specifically, OFDM has been chosen for digital audio and video broadcasting (DAB [1], DVB [2]), for high-speed modems over twisted pairs (digital subscriber line: xDSL [3]), and, more recently, for 5-GHz broadband wireless local area networks (HIPERLAN/2, IEEE802.11a and MMAC standards [4]).

OFDM enables very simple equalization of frequency-selective finite impulse response (FIR) channels, thanks to the inverse fast Fourier transform (IFFT) precoding and the insertion of a cyclic prefix (CP) of length larger than the channel memory at the transmitter. Present in each block of transmitted symbols, the CP consists of redundant symbols preceding (and circularly replicated from) the IFFT-precoded nonredundant symbols. At the receiver end, CP is discarded to avoid interblock interference (IBI), and each truncated block is FFT processed. A combination of IFFT and CP at the transmitter with the FFT at the

B. Muquet was with Motorola Labs, Paris, France. He is now with Stepmind, Boulogne-Billancourt, France (e-mail: bertrand.muquet@stepmind.com).

M. de Courville is with Motorola Labs, Paris, France (e-mail: Marc.de.Courville@crm.mot.com).

P. Duhamel was with Ecole Nationale Supérieure des Télécommunications, Paris, France. He is now with CNRS/LSS, Supélec, Gif-Sur-Yvette, France (e-mail: pierre.duhamel@lss.supelec.fr).

receiver converts the frequency-selective channel into parallel flat-faded subchannels, each one corresponding to a different subcarrier. Unless they are zero, flat fades are simply removed by dividing each subchannel output with the channel attenuation at the corresponding subcarrier.

At the same time, the need for high data rates motivated the search for blind identification and equalization methods because they save bandwidth by avoiding the use of training sequences [5]. Hence, numerous blind algorithms have been developed recently (see [6]), where several works have focused specifically on multicarrier systems. A blind equalization criterion has been introduced in [7]; it does not apply to traditional OFDM systems since it relies on a transmitter without CP. Correlation-matching methods based on the transmitted signal cyclostationarity have been presented in [8]–[10]. However, their implementation on existing systems is fairly difficult in practice because the presence of null side carriers (see Section IV-C) seriously complicates the proposed identification results. In [11], a method that could apply to OFDM systems with CP is provided, but it requires the CP length L to be equal to the block size M, which is never the case in practice. Some blind equalizers relying on the information contained in the CP were proposed in [12] and [13], but it may be preferable to first dispose of a channel estimation [for example, in order to shorten the channel impulse response (CIR) [14] or to determine power loading at the transmitter [3]]. Finally, a subspace algorithm that guarantees channel identifiability is proposed in [15] and [16] for the recent OFDM system with zero padding. Obviously, this algorithm does not apply to existing OFDM systems because the transmitter has a different structure and introduces a different kind of redundancy. This paper proposes a new channel estimation method that can be seen as its counterpart for traditional OFDM systems.

Based on a subspace decomposition [17], our algorithm takes advantage of the inherent redundancy introduced by the CP to blindly estimate the channel. It possesses the following attractive properties.

1) It does not require any modification of the classical OFDM transmitter. Thus, it is compatible with existing standardized OFDM systems.
2) It can be applied to arbitrary signal constellations. Unlike decision directed (DD) algorithms, it does not suffer from performance degradation when the constellation size increases. Moreover, it does not require the knowledge at the receiver of the constellation used for the transmission.
3) It is robust to channel order overdetermination. Furthermore, it guarantees channel identifiability, regardless of
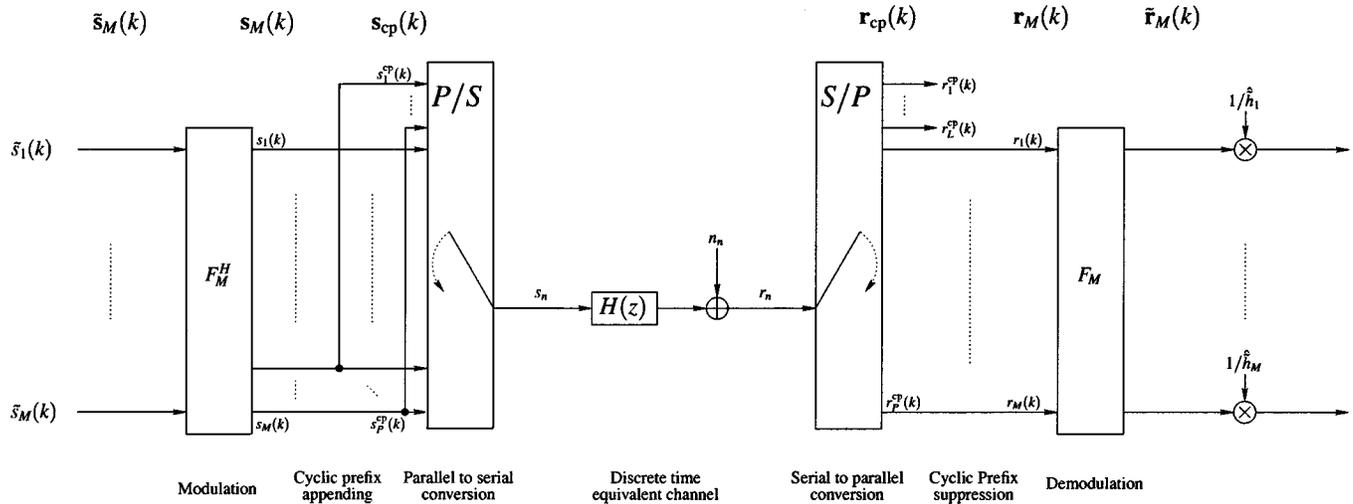
Fig. 1.   OFDM discrete baseband transceiver model.

the channel zeros location when the entire noise subspace is considered.

Actually, this last property is only of theoretical interest since the noise subspace dimension is unknown and can only be lower bounded in practice. Thus, the method effectively used in practice may operate only on a part of the noise subspace. This happens when the channel has nulls on subcarriers, and in this case, channel identifiability is no longer guaranteed. Finally, the theoretical algorithm (and the practical one, when there is no channel zeros on subcarriers) provides a perfect estimation after the observation of a finite number of received symbols in the noiseless case. In the noisy case, the algorithm is proved to be consistent.

Blind methods can also be used in cooperation with training data in order to better track channel variations. In that case, they are referred as semi-blind methods [18]. Usually, known blocks of symbols (usually referred as "pilot symbols") are transmitted at the beginning of each frame for synchronization and initial channel estimation purposes [19], [20]. We propose to use them to avoid the convergence period of the blind subspace algorithm, during which the estimation is unreliable. This semi-blind initialization can be extended to other blind algorithms (e.g., the one proposed in [16]). However, it does not overcome the inherent scalar indetermination of blind methods that still has to be removed. For that purpose, we propose to benefit from "pilot subcarriers" (i.e., subcarriers carrying symbols known to the receiver, which are found in many standards) and resort to a semi-blind least-squares criterion incorporating that knowledge. Finally, we detail some modifications of the original method that are required to comply with standards requirements. That way, we propose a new semi-blind estimator that directly applies to existing systems that *improves the accuracy of the initial channel estimation obtained at the beginning of the frame*. This is confirmed by simulations conducted in the realistic context of the HIPERLAN/2 (HL2), providing some performance comparisons with a DD channel estimator.

The rest of this paper is organized as follows: Section II provides a discrete model of OFDM and defines notations. Section III presents the new channel estimation method. Section IV

details the semi-blind implementation and the modifications required to cope with real systems. Section V presents simulations results, and conclusions are drawn in Section VI.

II. SYSTEM DESCRIPTION AND NOTATIONS

In this document, lower (upper) boldface symbols are used for column vectors (matrices), sometimes with subscripts to emphasize their sizes; argument k is used to index blocks of symbols; $\mathbf{I}_M$ denotes the size M identity matrix; $\otimes$ stands for the Kronecker product, $(.)^H$ denotes Hermitian, and $(.)^T$ transpose.

Fig. 1 depicts a discrete model of an OFDM system. The size M OFDM symbol $\tilde{\mathbf{s}}_M(k) = [\tilde{s}_1(k) \ldots \tilde{s}_M(k)]^T$ is first modulated by the IFFT matrix $\mathbf{F}_M^H$, where $\mathbf{F}_M$ stands for the size M FFT matrix with entries $e^{j2\pi mk/M}/\sqrt{M}$ (the notation tilde $(\tilde{\ })$ is used for frequency domain quantities, i.e., before IFFT precoding). The "time domain" vector $\mathbf{s}_M(k) = [s_1(k) \ldots s_M(k)]^T = \mathbf{F}_M^H \tilde{\mathbf{s}}_M(k)$ is then enlarged by a CP of length L, resulting in a size $P = M + L$ vector $\mathbf{s}_{\mathrm{cp}}(k) = [s_1^{\mathrm{cp}}(k) \ldots s_P^{\mathrm{cp}}(k)]^T$ whose components are finally sent sequentially through the channel. The channel effects are modeled by a linear FIR filter with CIR $\mathbf{h}_M = [h_0, \ldots, h_{M-1}]^T$ and the addition of noise samples n_n. Usually, the system is designed such that the CP is longer than the channel order which involves that $h_i = 0$ for $i > L$. Furthermore, M is usually greater than L, which we assume in the following and we denote as $\mathbf{h} = [h_0, \ldots, h_L]^T$ channel coefficients to be identified. At the receiver end, the CP is simply removed, yielding, after FFT demodulation, to the equivalent frequency domain model of Fig. 2, where $\tilde{\mathbf{h}}_M = [\tilde{h}_0 \ldots \tilde{h}_{M-1}]^T = \sqrt{M}\mathbf{F}_M\mathbf{h}_M$, $\tilde{h}_i$, denoting the channel attenuation on the ith subcarrier (see [21]).

Let $\mathbf{F} = [\mathbf{F}_{\mathrm{cp}}, \mathbf{F}_M]^H$ be the $P \times M$ matrix representing both the IFFT modulation and the CP appending, where $\mathbf{F}_{\mathrm{cp}}$ stands for the $M \times L$ matrix corresponding to the L last columns of $\mathbf{F}_M$. Let $\mathbf{H}_0$ be the $P \times P$ lower triangular Toeplitz matrix with first column $[h_0, \ldots, h_L, 0, \ldots, 0]^T$ and first row $[h_0, 0, \ldots 0]^T$. Let $\mathbf{H}_1$ be the $P \times P$ upper triangular
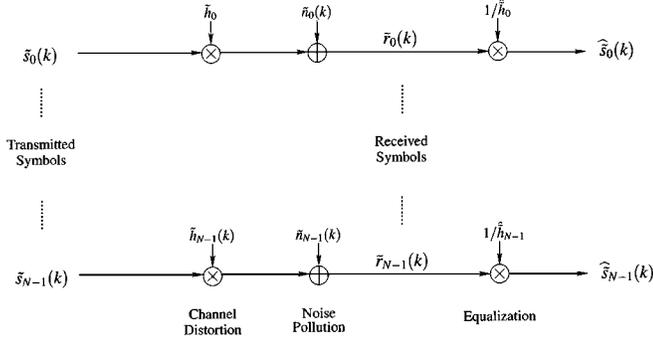
Fig. 2. Parallel carriers equivalent model.

Toeplitz matrix with first column $[0, \ldots, 0]^T$ and first row $[0, \ldots, 0, h_L, \ldots, h_1]$. The transmitted block of symbols is given by $\mathbf{s}_{\mathrm{cp}}(k) := \mathbf{F}\tilde{\mathbf{s}}_M(k)$, and the corresponding received block is

$$
\begin{aligned}
\mathbf{r}_{\mathrm{cp}}(k) &= \mathbf{H}_0\mathbf{s}_{\mathrm{cp}}(k) + \mathbf{H}_1\mathbf{s}_{\mathrm{cp}}(k-1) + \mathbf{n}_P(k) \\
&= \mathbf{H}_0\mathbf{F}\tilde{\mathbf{s}}_M(k) + \mathbf{H}_1\mathbf{F}\tilde{\mathbf{s}}_M(k-1) + \mathbf{n}_P(k)
\end{aligned} \quad (1)
$$

where $\mathbf{n}_P(k) = [n_1(k) \ldots n_P(k)]^T$ represents the noise samples polluting the transmission of $\mathbf{s}_{\mathrm{cp}}(k)$.

III. BLIND SUBSPACE ALGORITHM

We propose to apply a subspace algorithm in order to identify the channel coefficients from the observation of the received signal $\mathbf{r}_{\mathrm{cp}}(k)$. Subspace methods rely on a block formulation of the input–output relationship of the form $\mathbf{r}(k) = \mathbf{H}\tilde{\mathbf{s}}(k)$, where the transfer matrix $\mathbf{H}$ is a tall matrix. It can be observed that (1) does not exactly have the desired structure to directly derive a subspace algorithm due to the IBI term $\mathbf{H}_1\mathbf{F}\tilde{\mathbf{s}}_M(k-1)$. A way to address this problem is to use a particular precoder canceling IBI [15], but this approach requires a change in the transmitter, and we consider here only compatible methods. Another one is to choose $L = M$ to allow IBI suppression by simple block manipulations [11], but this would lead to CP much longer than in existing systems. Instead, this paper proposes a method that directly applies to existing systems.

For the sake of clarity, it is assumed in the following that $M = 4L$, which is a typical value in existing OFDM systems. The extension to any value of M, including noninteger values of the ratio M/L, verifying $M \geq 2L$ is presented in the Appendix. Note that $M \geq 2L$ is not a restrictive constraint since it is always satisfied in practice to limit the amount of introduced redundancy. Since $P = 5L$, $\mathbf{r}_{\mathrm{cp}}(k)$, $\mathbf{s}_{\mathrm{cp}}(k)$, and $\mathbf{n}_P(k)$ can be split into five sub-blocks of equal length L

$$
\begin{aligned}
\mathbf{r}_{\mathrm{cp}}(k) &= [\mathbf{r}_0(k)^T, \ldots, \mathbf{r}_4(k)^T]^T \\
\mathbf{s}_{\mathrm{cp}}(k) &= [\mathbf{s}_0(k)^T, \ldots, \mathbf{s}_4(k)^T]^T \\
\mathbf{n}_P(k) &= [\mathbf{n}_0(k)^T, \ldots, \mathbf{n}_4(k)^T]^T.
\end{aligned} \quad (2)
$$

Since $\mathbf{s}_0(k)$ corresponds to the CP, $\mathbf{s}_0(k) = \mathbf{s}_4(k)$. Intuitively, it is this redundancy that is used to obtain an overdetermined system. Let $\mathbf{C}_0$ be the $L \times L$ Toeplitz matrix with first column $[h_0, \ldots, h_{L-1}]^T$ and first row $[h_0, 0, \ldots 0]^T$. Letting $\mathbf{C}_1$ be the

$L \times L$ Toeplitz matrix with first column $[h_L, \ldots, 0]^T$ and first row $[h_L, \ldots, h_1]$, the block-received signal can be expressed as

$$
\begin{aligned}
\mathbf{r}_{\mathrm{cp}}(k) &= [\mathbf{r}_0(k), \mathbf{r}_1(k), \mathbf{r}_2(k), \mathbf{r}_3(k), \mathbf{r}_4(k)]^T \\
&= \begin{bmatrix} \mathbf{C}_1 & 0 & 0 & 0 & \mathbf{C}_0 \\ 0 & \mathbf{C}_0 & 0 & 0 & \mathbf{C}_1 \\ 0 & \mathbf{C}_1 & \mathbf{C}_0 & 0 & 0 \\ 0 & 0 & \mathbf{C}_1 & \mathbf{C}_0 & 0 \\ 0 & 0 & 0 & \mathbf{C}_1 & \mathbf{C}_0 \end{bmatrix} \begin{bmatrix} \mathbf{s}_4(k-1) \\ \mathbf{s}_1(k) \\ \mathbf{s}_2(k) \\ \mathbf{s}_3(k) \\ \mathbf{s}_4(k) \end{bmatrix} \\
&\quad + [\mathbf{n}_0(k), \mathbf{n}_1(k), \mathbf{n}_2(k), \mathbf{n}_3(k), \mathbf{n}_4(k),]^T.
\end{aligned} \quad (3)
$$

The IBI term no longer appears in (3), but the transfer matrix remains square. However, two successive overlapping received symbols can be considered in order to introduce some overdetermination. Therefore, let $\bar{\mathbf{r}}(k)$, $\bar{\mathbf{s}}(k)$, and $\bar{\mathbf{b}}(k)$ be the vectors defined as

$$
\begin{aligned}
\bar{\mathbf{r}}(k) &= (\mathbf{r}_1(k-1)^T, \ldots, \mathbf{r}_4(k-1)^T, \mathbf{r}_0(k)^T, \ldots, \mathbf{r}_4(k)^T)^T \\
\bar{\mathbf{s}}(k) &= (\mathbf{s}_1(k-1)^T, \ldots, \mathbf{s}_4(k-1)^T, \mathbf{s}_1(k)^T, \ldots, \mathbf{s}_4(k)^T)^T \\
\bar{\mathbf{n}}(k) &= (\mathbf{n}_1(k-1)^T, \ldots, \mathbf{n}_4(k-1)^T, \mathbf{n}_0(k)^T, \ldots, \mathbf{n}_4(k)^T)^T.
\end{aligned} \quad (4)
$$

Vector $\bar{\mathbf{r}}(k)$ is given by $\bar{\mathbf{r}}(k) = \mathbf{H}(\mathbf{h})\bar{\mathbf{s}}(k) + \bar{\mathbf{n}}(k)$, where $\mathbf{H}(\mathbf{h})$ is the $(2M + L) \times 2M$ matrix defined as

$$
\mathbf{H}(\mathbf{h}) = \begin{bmatrix} \mathbf{C}_0 & 0 & 0 & \mathbf{C}_1 & 0 & 0 & 0 & 0 \\ \mathbf{C}_1 & \mathbf{C}_0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{C}_1 & \mathbf{C}_0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{C}_1 & \mathbf{C}_0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{C}_1 & 0 & 0 & 0 & \mathbf{C}_0 \\ 0 & 0 & 0 & 0 & \mathbf{C}_0 & 0 & 0 & \mathbf{C}_1 \\ 0 & 0 & 0 & 0 & \mathbf{C}_1 & \mathbf{C}_0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{C}_1 & \mathbf{C}_0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{C}_1 & \mathbf{C}_0 \end{bmatrix}. \quad (5)
$$

Let $\mathbf{R}_{\bar{\mathbf{r}}\bar{\mathbf{r}}} = E[\bar{\mathbf{r}}(k)\bar{\mathbf{r}}(k)^H]$ and $\mathbf{R}_{\bar{\mathbf{s}}\bar{\mathbf{s}}} = E[\bar{\mathbf{s}}(k)\bar{\mathbf{s}}(k)^H]$ be the autocorrelation matrices of $\bar{\mathbf{r}}(k)$ and $\bar{\mathbf{s}}(k)$, respectively. We assume in this section that matrix $\mathbf{R}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}$ is full rank and discuss the case when this assumption does not hold in Section IV-C. The noise is also assumed to be white with variance σ_n^2, but the method can be extended to colored noise following the lines of [16] and [17]. The subspace method relies on the autocorrelation matrix $\mathbf{R}_{\bar{\mathbf{r}}\bar{\mathbf{r}}}$ defined as $\mathbf{R}_{\bar{\mathbf{r}}\bar{\mathbf{r}}} = \mathbf{H}(\mathbf{h})\mathbf{R}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}\mathbf{H}(\mathbf{h})^H + \sigma_n^2\mathbf{I}_{2M+L}$. We distinguish below two cases that depend on whether some subcarriers are hit by channel nulls or not since this property affects identifiability.

A. Subspace Identification When No Channel Zero Is Located on Subcarriers

It is shown in the Appendix that $\mathbf{H}(\mathbf{h})$ is full column rank if no channel zero is located on any subcarrier. In this case, if $\mathbf{R}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}$ is full rank, the matrix $\mathbf{H}(\mathbf{h})\mathbf{R}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}\mathbf{H}(\mathbf{h})^H$ has rank $8L$. Therefore, its left nullspace has dimension L and is spanned by a basis of L vectors $\mathbf{g}_0, \ldots, \mathbf{g}_{L-1}$. A particular basis for this subspace (usually referred as the noise subspace) can be found from the singular value decomposition (SVD) of $\mathbf{R}_{\bar{\mathbf{r}}\bar{\mathbf{r}}}$. It is indeed spanned by the L eigenvectors associated with the L smallest eigenvalues σ_n^2 of the autocorrelation matrix $\mathbf{R}_{\bar{\mathbf{r}}\bar{\mathbf{r}}}$. Moreover, it

is well known that the range space of a matrix is orthogonal to its left null space, and hence, the space spanned by the columns of $\mathbf{H}(\mathbf{h})$ (usually referred as the signal subspace) is orthogonal to the noise subspace and, hence, to each vector $\mathbf{g}_i$ in the basis. Thus, for any vector $\mathbf{g}_i$, the equation $\mathbf{g}_i^H \mathbf{H}(\mathbf{h}) = \mathbf{0}$ holds, and hence, $\mathbf{h}$ must satisfy the system of linear equations given by

$$\mathbf{g}_i^H \mathbf{H}(\mathbf{h}) = \mathbf{0} \text{ for } 0 \leq i \leq L - 1. \tag{6}$$

In order to express (6) in terms of the vector of unknowns $\mathbf{h}$, split any vector $\mathbf{g}_i$ into nine blocks of equal length L: $\mathbf{g}_i = [\mathbf{g}_i^{1T}, \ldots, \mathbf{g}_i^{9T}]^T$, where $\mathbf{g}_i^j = [g_i^j(1) \ldots g_i^j(L)]^T$. In addition, let $\mathbf{A}_i^j$, $\mathbf{B}_i^j$ and $\mathbf{C}_i^j$ be the following size $(L+1) \times L$ matrices:

$$\mathbf{A}_i^j = \begin{bmatrix} g_i^j(1) & \cdots & \cdots & g_i^j(L) \\ \vdots & & \nearrow & 0 \\ \vdots & \nearrow & \nearrow & \vdots \\ g_i^j(L) & \nearrow & & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix} \tag{7}$$

$$\mathbf{B}_i^j = \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ \vdots & & \nearrow & g_i^j(1) \\ \vdots & \nearrow & \nearrow & \vdots \\ 0 & \nearrow & & \vdots \\ g_i^j(1) & \cdots & \cdots & g_i^j(L) \end{bmatrix}. \tag{8}$$

Let $\mathbf{C}_i^j = \mathbf{A}_i^j + \mathbf{B}_i^{j+1}$, and define matrix $\mathbf{G}_i$ as $\mathbf{G}_i = [\mathbf{C}_i^1, \mathbf{C}_i^2, \mathbf{C}_i^3, \mathbf{C}_i^4 + \mathbf{B}_i^1, \mathbf{C}_i^5 + \mathbf{A}_i^9, \mathbf{C}_i^6, \mathbf{C}_i^7, \mathbf{C}_i^8]$. Equation (6) is then equivalent to

$$\mathbf{h}^H \mathbf{G}_i = \mathbf{0} \text{ for } 0 \leq i \leq L - 1 \tag{9}$$

It is shown in the Appendix that this system of equations uniquely determines $\mathbf{h}$ up to a scalar factor, *under the assumption that $\mathbf{H}(\mathbf{h})$ is full column rank*, that is, when no channel null is located on a subcarrier. Note that the method is robust to channel order overestimation since only an upper bound L of the channel order is required.

In practice, $\mathbf{R}_{\overline{\mathbf{r}}\overline{\mathbf{r}}}$ is estimated by an averaging in time over, say, N blocks

$$\widehat{\mathbf{R}}_{\overline{\mathbf{r}}\overline{\mathbf{r}}}^{(N)} = \frac{1}{N} \sum_{k=0}^{N-1} \overline{\mathbf{r}}(k) \overline{\mathbf{r}}(k)^H \tag{10}$$

and hence, (9) has to be solved in the least square sense to estimate $\mathbf{h}$ since only estimates $\widehat{\mathbf{G}}_i$ of matrices $\mathbf{G}_i$ are available. Let us define $\widehat{\mathbf{R}}_{\overline{\mathbf{s}}\overline{\mathbf{s}}}^{(N)}$ as

$$\widehat{\mathbf{R}}_{\overline{\mathbf{s}}\overline{\mathbf{s}}}^{(N)} = \frac{1}{N} \sum_{k=0}^{N-1} \overline{\mathbf{s}}(k) \overline{\mathbf{s}}(k)^H \tag{11}$$

In a noiseless context, $\widehat{\mathbf{R}}_{\overline{\mathbf{r}}\overline{\mathbf{r}}}^{(N)}$ can be expressed as

$$\widehat{\mathbf{R}}_{\overline{\mathbf{r}}\overline{\mathbf{r}}}^{(N)} = \mathbf{H}(\mathbf{h}) \, \widehat{\mathbf{R}}_{\overline{\mathbf{r}}\overline{\mathbf{r}}}^{(N)} \, \mathbf{H}(\mathbf{h})^H \tag{12}$$

and therefore, (9) exactly holds as soon as $\widehat{\mathbf{R}}_{\overline{\mathbf{s}}\overline{\mathbf{s}}}^{(N)}$ is full rank. This ensures a perfect channel estimation after the observation

of a finite number of received symbols $\overline{\mathbf{r}}(k)$, provided that there exists an $N \geq 2M$ such that $\widehat{\mathbf{R}}_{\overline{\mathbf{s}}\overline{\mathbf{s}}}^{(N)}$ is full rank. This condition is usually referred as the persistence of excitation assumption (p.o.e.) [22] and is verified with most constellations for values of $N \approx 2M$. Thus, the method can be applied to arbitrary constellations with performance that hardly depends on the constellation as supported by simulations. In the noisy case, $\widehat{\mathbf{R}}_{\overline{\mathbf{r}}\overline{\mathbf{r}}}^{(N)}$ converges in the mean square sense to the true correlation matrix $\mathbf{R}_{\overline{\mathbf{r}}\overline{\mathbf{r}}}$ as $N \to \infty$ since input $\tilde{\mathbf{s}}(k)$ has finite moments and the channel has finite memory. Hence, $\widehat{\mathbf{h}} \longrightarrow \mathbf{h}$ as $N \to \infty$, and the algorithm is consistent.

Actually, solving for (9) usually requires a huge amount of memory and computations. Hence, the storage of all matrices $\widehat{\mathbf{G}}_i$ is usually avoided by computing on the fly the quadratic form matrix $\widehat{\mathbf{Q}} = \sum_{i=0}^{L-1} \widehat{\mathbf{G}}_i \widehat{\mathbf{G}}_i^H$ and by solving the quadratic equation $q(\mathbf{h}) = 0$, where $q(\mathbf{h}) = \mathbf{h}^H \widehat{\mathbf{Q}} \mathbf{h}$. At this point, $\mathbf{h}$ can be obtained (up to a scalar coefficient) by minimizing $q(\mathbf{h})$ subject to a properly chosen constraint, avoiding the trivial solution $\mathbf{h} = \mathbf{0}$. In practice, a power control is performed at the receiver to ensure that the received power is approximately constant. Therefore, a natural choice is to minimize $q(\mathbf{h})$ subject to the constraint $\|\mathbf{h}\| = 1$. In this case, the channel estimate $\widehat{\mathbf{h}}$ is the unit-norm eigenvector associated with the smallest eigenvalue of $\widehat{\mathbf{Q}}$. This way, the channel is identified up to a phase factor that has to be removed, and Section IV A proposes a mechanism for that purpose.

B. Subspace Identification When Channels Zeros Are Located on Some Subcarriers

It is shown in the Appendix that $\mathbf{H}(\mathbf{h})$ is no longer full column rank if some channel zeros are located on subcarriers. Specifically, $\mathbf{H}(\mathbf{h})$ is shown to have rank $2M - Z$, where Z is the number of subcarriers hit by channel nulls. Hence, the noise subspace has dimension $L + Z$, and the system of (9) makes use of only L independent vectors of it. In this case, the subspace algorithm may fail as shown in the Appendix, but it is possible to enlarge the system (9) to consider a full basis of the noise subspace that will guarantee that the channel is uniquely identified, as shown in the Appendix. Since it is possible to find the number Z of zeros asymptotically (it is equal to the multiplicity of the smallest eigenvalue), identifiability using the (theoretical) subspace algorithm is thus guaranteed. However, this result is only of theoretical interest because it is impossible in practice to observe an infinite number of symbols. Thus, it is difficult to evaluate the number Z of channel zeros, and practically, we will never be able to use the full noise subspace, and hence, channel identifiability cannot be guaranteed in practice.

IV. SEMI-BLIND IMPLEMENTATION IN REAL CONTEXTS

This section explains how to take advantage of the training data (that are usually specified in standards) to raise the scalar indetermination (Section IV-A) and improve the performance of the algorithm (Section IV-B). It also details the modifications that are necessary to apply the algorithm to existing systems (Section IV-C).

A. A Mechanism to Remove the Phase Indetermination Inherent to Blind Methods

Blind methods always identify the channel up to one scalar indetermination, which has to be removed to allow the received symbols to be equalized. This appears clearly in (9), which admits $\alpha\mathbf{h}$ for any α as a solution. Actually, standards often specify some pilot subcarriers carrying known symbols for phase tracking and channel estimation refinements purposes. We propose to exploit them to remove the blind scalar indetermination and, thus, to allow the received symbols to be equalized as detailed in what follows.

Let $\hat{\mathbf{h}}_{\mathrm{sub}}$ be the channel estimation provided by the blind subspace algorithm. The problem is to find the scalar coefficient α such that $\hat{\mathbf{h}} = \hat{\mathbf{h}}_{\mathrm{sub}}/\alpha$ is close in the quadratic sense to the true channel vector $\mathbf{h}$. Denote by F the number of pilot subcarriers on which some known symbols are transmitted. These carriers can be equispaced (as in HL2), or not, and at fixed or various positions, but it is not necessary for what follows to explicitly state their exact position. Let $\tilde{s}_{\mathrm{pil}}(1), \ldots, \tilde{s}_{\mathrm{pil}}(F)$ be the known symbols transmitted on the pilot subcarriers and $\tilde{r}_{\mathrm{pil}}(1), \ldots, \tilde{r}_{\mathrm{pil}}(F)$ be the corresponding FFT-processed received symbols. An estimation of the channel attenuations at the corresponding frequencies is provided by

$$\widehat{\mathbf{h}}_{\mathrm{pil}} = \left[\frac{\tilde{r}_{\mathrm{pil}}(1)}{\tilde{s}_{\mathrm{pil}}(1)}, \ldots, \frac{\tilde{r}_{\mathrm{pil}}(F)}{\tilde{s}_{\mathrm{pil}}(F)} \right]^T. \tag{13}$$

Let $\mathbf{F}_{\mathrm{pil}}$ be the $F \times L$ matrix obtained from the L first columns of matrix $\sqrt{M}\mathbf{F}_M$ by selecting the rows corresponding to the pilot carriers and by removing the other ones. Another estimation of these coefficients can be inferred from the subspace identification up to α

$$\widehat{\mathbf{h}}_{\mathrm{sub}} = \mathbf{F}_{\mathrm{pil}}\hat{\mathbf{h}}_{\mathrm{sub}} = \alpha\mathbf{F}_{\mathrm{pil}}\hat{\mathbf{h}} = \alpha\widehat{\mathbf{h}}_{\mathrm{pil}}. \tag{14}$$

From (13) and (14), α can be determined by solving the linear system $\widehat{\mathbf{h}}_{\mathrm{sub}} = \alpha\widehat{\mathbf{h}}_{\mathrm{pil}}$ in the least square sense. However, if the channel estimation $\hat{\mathbf{h}}_{\mathrm{sub}} = \alpha\hat{\mathbf{h}}$ obtained using the subspace algorithm is far from the true CIR $\mathbf{h}$, the final channel estimation remains inaccurate, even if α is estimated such that $\|\mathbf{h} - \hat{\mathbf{h}}_{\mathrm{sub}}/\alpha\|$ is minimal. Somehow, no benefit is taken from the knowledge of the channel attenuations on the pilot carriers for the subspace algorithm. This can be overcome by considering the following system of equations:

$$\begin{cases} \mathbf{G}_i^H\mathbf{h} = \mathbf{0}, & \text{for } 0 \le i \le L-1 \\ \mathbf{F}_{\mathrm{pil}}\mathbf{h} = \widetilde{\mathbf{h}}_{\mathrm{pil}}. \end{cases} \tag{15}$$

Since this system of equations holds only approximately, it has to be solved in the least squares sense similarly to (9), which amounts to minimizing the quadratic criterion $q'(\mathbf{h})$, which is defined as

$$q'(\mathbf{h}) = \left(\sum_{i=0}^{L-1} \|\hat{\mathbf{G}}_i^H\mathbf{h}\|^2 + \|\mathbf{F}_{\mathrm{pil}}\mathbf{h} - \widehat{\mathbf{h}}_{\mathrm{pil}}\|^2 \right). \tag{16}$$

It is possible to increase the robustness of this approach by changing the confidence degree in the pilot carriers using a weighting factor β (see [23] for optimally choosing β). In this case, $(\sum_{i=0}^{L-1} \|\hat{\mathbf{G}}_i^H\mathbf{h}\|^2 + \beta\|\mathbf{F}_{\mathrm{pil}}\mathbf{h} - \widehat{\mathbf{h}}_{\mathrm{pil}}\|^2)$ has to be minimized to estimate the channel. This criterion is close to the one proposed in another context in [24]; the difference here is that the training symbols alone are not sufficient to estimate the channel. Denote by $\hat{\mathbf{Q}}_{\mathrm{pil}} = \sum_{i=0}^{L-1} \hat{\mathbf{G}}_i\hat{\mathbf{G}}_i^H + \beta\mathbf{F}_{\mathrm{pil}}^H\mathbf{F}_{\mathrm{pil}}$ the channel estimation, which is finally given by $\hat{\mathbf{h}} = (\mathbf{Q}_{\mathrm{pil}})^{-1}\beta\mathbf{F}_{\mathrm{pil}}^H\widehat{\mathbf{h}}_{\mathrm{pil}}$. Thus, this semi-blind subspace algorithm can be seen as a channel-dependent interpolator between the pilot subcarriers.

Furthermore, it is shown in the Appendix that considering the composite system of (15) that accounts directly for the pilot symbols reduces the lack of identifiability of the blind algorithm. Indeed, channel identifiability is guaranteed in this case if the number Z of zeros located on subcarriers is smaller than the number F of pilot subcarriers (which is a situation that still becomes less likely).

B. Training Symbol-Based Initialization of the Blind Algorithm

An inherent problem to blind methods is their rather slow convergence rate [6], which often prevents their use in realistic contexts where methods based on training sequences are preferred. Thus, standards usually specify the transmission at the beginning of each frame of known blocks of symbols for synchronization and initial channel estimation at the receiver. We propose to use these pilot symbols to initialize the estimation of the autocorrelation matrix. This enables us to skip the long convergence period of the blind algorithm and to obtain immediately the same accuracy as the pilot-based estimation. The steps of the proposed algorithm are detailed as follows.

```
1) Obtain an initial channel estimation:
ĥ⁽⁰⁾ through the pilot symbols (see, for
instance, [19] and [25]).
2) From ĥ⁽⁰⁾, generate an estimation of ma-
trix R̂r̄r̄ as R̂r̄r̄⁽⁰⁾ = H(ĥ⁽⁰⁾)Rs̄s̄H(ĥ⁽⁰⁾)ᴴ.
3) Refine iteratively the autocorrelation
matrix estimation each time a new block
symbol r̄(k) is received using an adaptive
process of forgetting factor λ ∈ [0,1]
```

$$\widehat{\mathbf{R}}_{\overline{\mathbf{rr}}}^{(N)} := \lambda\widehat{\mathbf{R}}_{\overline{\mathbf{rr}}}^{(N-1)} + (1-\lambda)\overline{\mathbf{r}}(N)\overline{\mathbf{r}}(N)^H \tag{17}$$

```
or a sliding window of length W in (18),
shown at the bottom of the next page.
4) Perform the subspace algorithm based on
R̂r̄r̄⁽ᴺ⁾.
```

Note that Step 2 does not require the noise variance σ_n^2 to be known because the signal and the noise subspaces of matrix $\mathbf{R}_{\overline{\mathbf{rr}}}$ do not depend on it. However, the autocorrelation matrix $\mathbf{R}_{\overline{\mathbf{ss}}}$ has to be known, which is not the case with the blind algorithm.

C. Subspace Estimation With Null Side Carriers

In all standardized OFDM systems, null carriers (zeros) are appended on each side of the modulator (cf. Fig. 3). These carriers are specified to provide frequency guard bands in order to avoid interference between adjacent systems. This particularity is often ignored in the literature but has to be considered for the
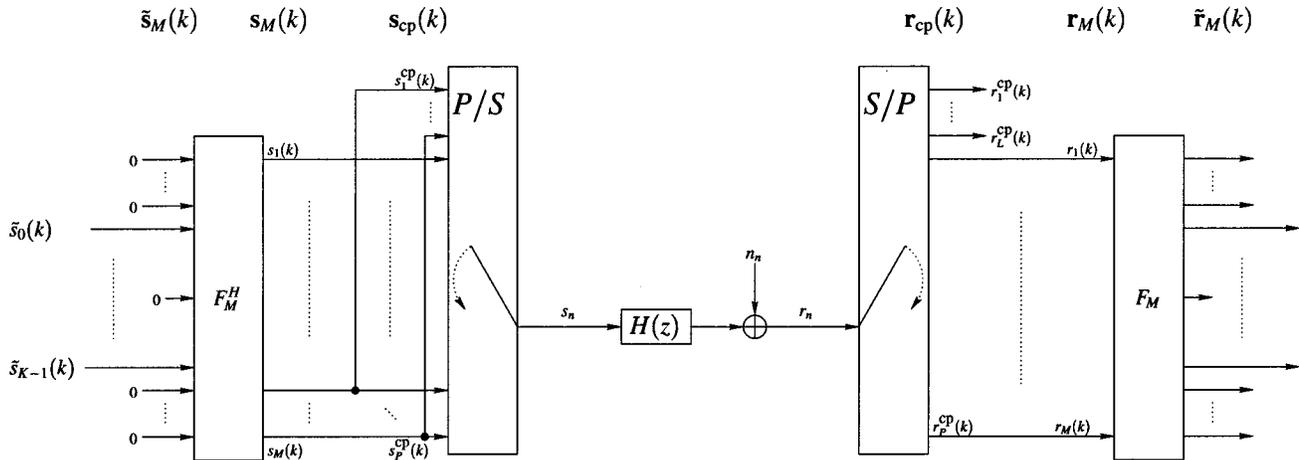
Fig. 3.   Oversampled OFDM discrete model.

subspace algorithm since $\mathbf{R}_{\overline{\mathbf{ss}}}$ is no longer full rank: a fact that modifies the signal and noise subspaces of $\mathbf{R}_{\overline{\mathbf{rr}}}$. We detail, in what follows, some modifications of the algorithm that are necessary to cope with this practical situation.

Define K as the number of nonzero subcarriers, and assume without loss of generality that the side carriers are located at the bottom of the FFT input: $\tilde{s}_i(k) = 0$ for $i > K$. Let $\tilde{\mathbf{s}}_K(k) = [\tilde{s}_1(k), \ldots, \tilde{s}_K(k)]^T$ be the vectors of nonzero entries and $\tilde{\mathbf{s}}_{2K}(k) = [\tilde{\mathbf{s}}_K(k)^T, \tilde{\mathbf{s}}_K(k-1)^T]^T$. Let $\mathbf{F}_{M \times K}$ be the $M \times K$ obtained from $\mathbf{F}_M^H$ by removing its last $M - K$ columns, and define $\mathbf{F}_{2M \times 2K}$ as $\mathbf{F}_{2M \times 2K} = \mathbf{I}_2 \otimes \mathbf{F}_{M \times K}$. Due to the zeros at the FFT input, $\overline{\mathbf{r}}(k) = \mathbf{H}(\mathbf{h})\overline{\mathbf{s}}(k) + \overline{\mathbf{n}}(k)$ can also be expressed as

$$\overline{\mathbf{r}}(k) = \mathbf{H}(\mathbf{h})\mathbf{F}_{2M \times 2K}\tilde{\mathbf{s}}_{2K}(k) + \overline{\mathbf{n}}(k). \tag{19}$$

Thus, it is possible to proceed as in Section III to obtain the channel coefficients from the autocorrelation matrix $\mathbf{R}_{\overline{\mathbf{rr}}}$ of the received signal

$$\mathbf{R}_{\overline{\mathbf{rr}}} = \mathbf{H}(\mathbf{h})\mathbf{F}_{2M \times 2K}\mathbf{R}_{\tilde{\mathbf{s}}_{2K}\tilde{\mathbf{s}}_{2K}}\mathbf{F}_{2M \times 2K}^H + \sigma_n^2 \mathbf{I}_{2M+L}\mathbf{H}(\mathbf{h})^H \tag{20}$$

where $\mathbf{R}_{\tilde{\mathbf{s}}_{2K}\tilde{\mathbf{s}}_{2K}} = E[\tilde{\mathbf{s}}_{2K}(k)\tilde{\mathbf{s}}_{2K}(k)^H]$. Consistently with the p.o.e. assumption, matrix $\mathbf{R}_{\tilde{\mathbf{s}}_{2K}\tilde{\mathbf{s}}_{2K}}$ can be assumed to have full rank, and hence, the noise subspace of $\mathbf{R}_{\overline{\mathbf{rr}}}$ has dimension $P + M - 2K$. Thus, it is spanned by a set of $P+M-2K$ independent vectors (which can be obtained as previously by SVD), and it can be shown in a similar way that $\mathbf{h}$ has to satisfy

$$\mathbf{F}_{2M \times 2K}^T \mathbf{G}_i^H \mathbf{h} = \mathbf{0} \text{ for } 0 \leq i \leq P + M - 2K - 1. \tag{21}$$

This modified algorithm works for any value of $K \leq M$, including the case $K = M$ initially described and can be generalized to the subspace algorithm proposed in [16] for zero-padded (ZP) multicarrier transmissions. Moreover, when $K < M$, a channel estimation is provided blindly as soon as $R_{\overline{\mathbf{rr}}}$ reaches rank $2K$ (instead of $2M$ previously). Thus, the transmission of zeros on side carriers, which was foreseen as a potential issue, reveals itself as a means for increasing the convergence capabilities of the algorithm (intuitively increasing the amount of redundancy facilitates the channel determination). Finally, simulations suggest that it shares the same properties as the original algorithm (i.e., identifiability is guaranteed only when no channel zero is located on subcarriers).

V. SIMULATIONS

In this section, we illustrate the merits of our semi-blind channel estimator through realistic simulations conducted in the context of the existing standard HIPERLAN/2. All the results are compared with those obtained using either the classical pilot-based estimation method [19] or using a DD estimator [19], [26]. Note that an exhaustive comparison with the subspace algorithm proposed in [16] and, more generall,y between CP and ZP transmitters is reported in [27] because we only compare algorithms that are compatible with existing standards in this paper.

A. Presentation of HIPERLAN/2

We have chosen to illustrate our subspace algorithm in the specific context of the HIPERLAN/2 broadband wireless communication standard, which is similar to IEEE802.11a and MMAC. HL2 is a multicarrier system operating at 5 GHz using a 20-MHz band at typical SNR values of 0–40 dB for terminal speeds $v \leq 3$ m/s. The number of carriers is $M = 64$, and the CP is 16 samples long. As depicted in Fig. 3, 12 of the 64 carriers are null carriers. Among the $K = 52$ nonzero subcarriers, four are carrying known QPSK symbols $P_1 - P_4$, whereas the other $U = K - 4 = 48$ subcarriers convey the information-bearing sequences. With X denoting each of the

$$\widehat{\mathbf{R}}_{\overline{\mathbf{rr}}}^{(N)} := \widehat{\mathbf{R}}_{\overline{\mathbf{rr}}}^{(N-1)} + \begin{cases} \frac{1}{W}\left[\overline{\mathbf{r}}(N)\overline{\mathbf{r}}(N)^H - \widehat{\mathbf{R}}_{\overline{\mathbf{rr}}}^{(0)}\right], & \text{for } N \leq W \\ \frac{1}{W}\left[\overline{\mathbf{r}}(N)\overline{\mathbf{r}}(N)^H - \overline{\mathbf{r}}(N-W)\overline{\mathbf{r}}(N-W)^H\right], & \text{for } N > W. \end{cases} \tag{18}$$

48 information symbols drawn from four-, 16-, or, 64-quadrature amplitude modulation (QAM) constellations (depending on the target bit rate), the symbol structure is summarized in (21a), shown at the bottom of the page. The first two blocks of a frame $\tilde{\mathbf{s}}_M(0)$ and $\tilde{\mathbf{s}}_M(1)$ are known to the receiver and can be used for initial channel estimation. Relying on these training symbols and the received noisy FFT processed data $\tilde{\mathbf{r}}_M(k) = [\tilde{r}_1(k), \ldots, \tilde{r}_M(k)]$, the receiver forms an initial channel estimate as

$$\widehat{\mathbf{h}}_i(1) = \frac{1}{2}\left[\frac{\tilde{r}_i(0)}{\tilde{s}_i(0)} + \frac{\tilde{r}_i(1)}{\tilde{s}_i(1)}\right] \text{ for } 1 \le i < M. \quad (22)$$

This method is the classical pilot-based estimation algorithm usually adopted for coherent modulation [19]. Because only symbols $\tilde{s}_m(k)$ for $m = 12, 26, 40$, and 54 contain known data in subsequent blocks $\tilde{\mathbf{s}}(k)_{k\ge 2}$, one can track adaptively the channel transfer function using a running average (over, say, $B = 20$ blocks) only on these four frequencies as follows:

$$\widehat{\tilde{h}}_i(k+1) = \frac{1}{B}\sum_{l=0}^{B-1} \frac{\tilde{r}_i(k-l)}{\tilde{s}_i^{\text{pil}}(k-l)} \text{ for } i \in \{12, 26, 40, 54\}. \quad (23)$$

Actually, the standard specifies these four pilot subcarriers for synchronization and phase-tracking purposes, but they are too distant in frequency (spaced more than the channel coherence bandwidth) for estimating the channel by a simple interpolation or even for tracking channel variations. Thus, only a partial channel tracking can be achieved using (23), which may not yield accurate channel estimation in rapidly varying environments. To enhance mobility in HL2, semi-blind channel estimation is well motivated, especially with the relatively small number of carriers that enables even subspace approaches to be tried with affordable complexity.

B. Simulations Description

Simulation results have been obtained in this paper running Monte Carlo trials, where each trial corresponds to a different realization of the channel model. The channels models considered are a random FIR channel with random order upper bounded by the CP length and the channel model A provided by the HL2 standard [28]. Channel A corresponds to a typical office environment and is given in Table I, where a classical Jake's Doppler spectrum and Rayleigh fading statistics are assumed for all taps. Results are provided both for a static channel (using an averaging window of size 500) and for a time-varying channel (using an exponential window) for a terminal speed of $v = 3$ m/s, as specified in the HL2 standard.

TABLE I
CHANNEL MODEL A

Tap	Delay(ns)	Av Power(dB)	Tap	Delay(ns)	Av Power(dB)
1	0	0.0	10	90	-7.8
2	10	-0.9	11	110	-4.7
3	20	-1.7	12	140	-7.3
4	30	-2.6	13	170	-9.9
5	40	-3.5	14	200	-12.5
6	50	-4.3	15	240	-13.7
7	60	-5.2	16	290	-18.0
8	70	-6.1	17	340	-22.4
9	80	-6.9	18	390	-26.7

The figures of merit for channel estimation are 1) the time domain (TD) mean square error (MSE) defined as

$$\text{TMSE} = \frac{\sum_u |h_u - \widehat{h}_u|^2}{\sum_u |h_u|^2} \quad (24)$$

and the frequency domain (FD) MSE defined for the set $\mathcal{U}$ of indices corresponding to the U useful carriers as

$$\text{FMSE} = \frac{\sum_{u\in\mathcal{U}} |\tilde{h}_u - \widehat{\tilde{h}}_u|^2}{\sum_{u\in\mathcal{U}} |\tilde{h}_u|^2}. \quad (25)$$

The first criterion is relevant if equalization is performed in the time domain as is done in [16] or for channel shortening [14], whereas the second one is relevant if equalization is performed in the frequency domain (in that case, it is natural to consider the channel estimation error only on the carriers that are effectively conveying information). We also plot BER curves to give an insight into how channel MSE translates into BER performance.

The autocorrelation matrix is refined each time a new block of symbols is received, and a new channel estimation is computed every 50 blocks of symbols. This arbitrary choice keeps the arithmetical complexity to reasonable values without sacrificing channel tracking. Note that more frequent channel estimation would be useless since the channel does not vary quickly. For reference purposes, the semi-blind algorithm is compared with the classical pilot-based estimation method previously described, as well as a DD algorithm. This algorithm does the following.

```
1) It first estimates the channel as in
the classical method.
2) It equalizes ř(k) by simply dividing
r̃_i(k) with the channel attenuation h̃_i(k - 1)
estimation obtained at step k-1 to obtain
symbol estimates as ŝ̃_i(k) = r̃_i(k)/h̃_i(k-1);
3) It updates the channel attenuations es-
timations as h̃_i(k) = 1/B ∑_{l=0}^{l=B-1} r̃_i(k-l)/Q(ŝ̃_i(k-
l)), where Q(ŝ̃_i(k - l)) denotes the hard de-
```

$$|\underbrace{0\cdots 0}_{6}\underbrace{X\cdots X}_{5}P_1\underbrace{X\cdots X}_{13}P_2\underbrace{X\cdots X}_{6}0\underbrace{X\cdots X}_{6}P_3\underbrace{X\cdots X}_{13}P_4\underbrace{X\cdots X}_{5}\underbrace{0\cdots 0}_{5}| \quad (21a)$$

Fig. 4. Time domain channel estimation MSE along the frame: $E_s/N_0 = 25$ dB.
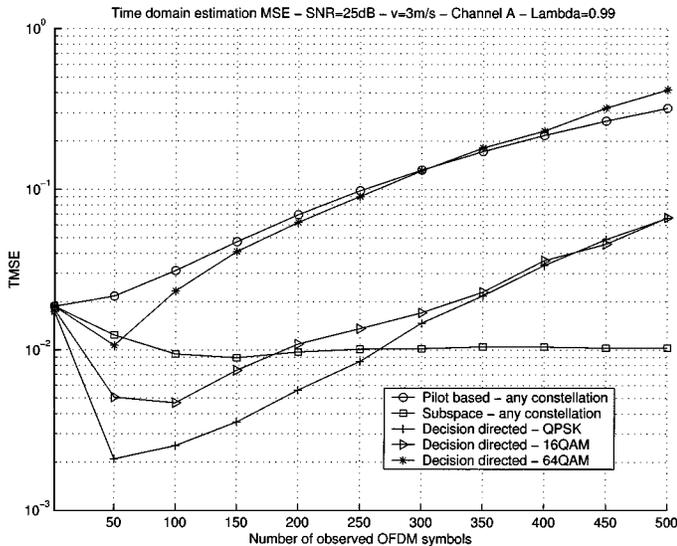

Fig. 5. FD channel estimation MSE along the frame: $E_s/N_0 = 25$ dB.

cision taken from $\hat{\hat{s}}_i(k - l)$ by the quantizer $Q(.)$ associated with the constellation.

Note that it has been chosen to average the DD estimation over $B = 20$ blocks because it has revealed experimentally to be a good tradeoff between error propagation (large values of B) and channel tracking (small values of B). Note also that a denoising of the estimation is also applied to the pilot-based and DD estimations to allow a fair comparison between algorithms. This allows us to take into account in the estimation that the channel frequency response actually corresponds in the time domain to a CIR of length upper bounded by $\approx L = M/4$ [29].

C. Simulation Results

Figs. 4 and 5 illustrate the behavior of the three channel estimation methods along the frame for transmission over the time-varying channel A at $E_s/N_0 = 25$ dB. They depict the MSE in the time and frequency domains as a function of the number of received symbols. It appears clearly that the channel estimation provided by the pilot-based method degrades quickly when the channel is varying and that tracking the channel variations is a must with long frames. The subspace algorithm tracks the channel variations and offers performance independent of the constellation. For channel estimation in the time domain, it offers good performance and clearly outperforms the DD algorithm for 16 and 64 QAM. For channel estimation in the frequency domain, the DD is preferable. Note, however, that the probability of making a wrong decision increases with the constellation size with the DD and that it can suffer error propagation, unlike the subspace algorithm (this property has often prevented the DD algorithm to be used in practice). This is especially highlighted in Fig. 5, where the effect of error propagation is clear. The channel estimation provided by the DD is degrading along the frame, whereas the subspace accuracy is approximately constant after a few observed symbols.

In order to compare more deeply the subspace and the DD algorithm, Figs. 6–8 depict the average MSE in the time
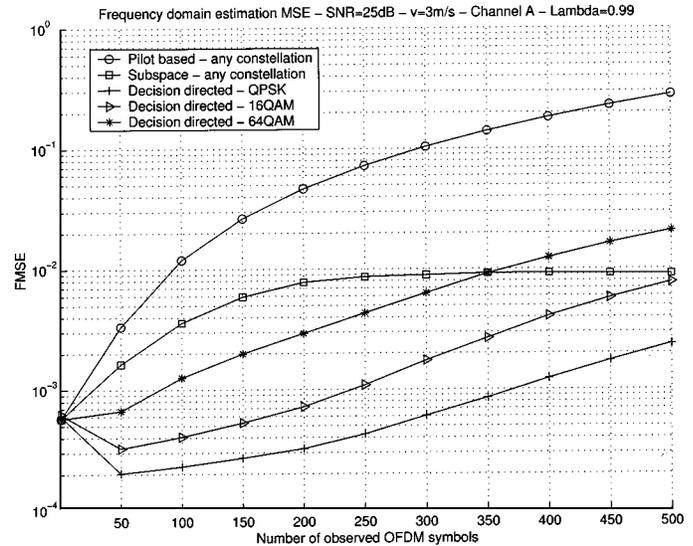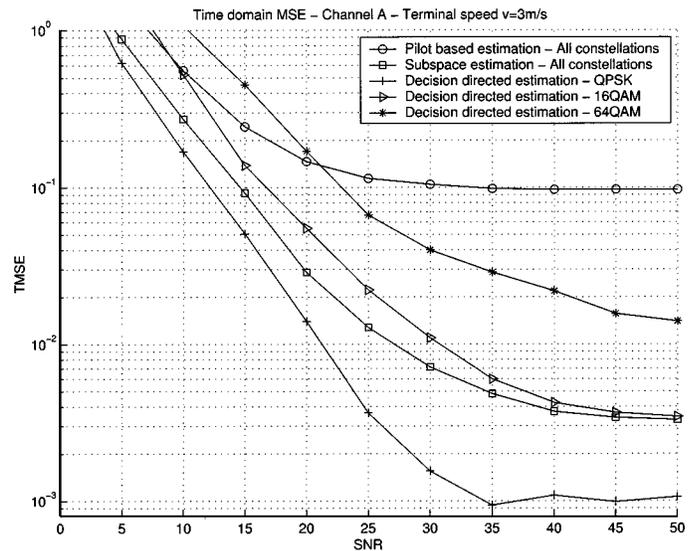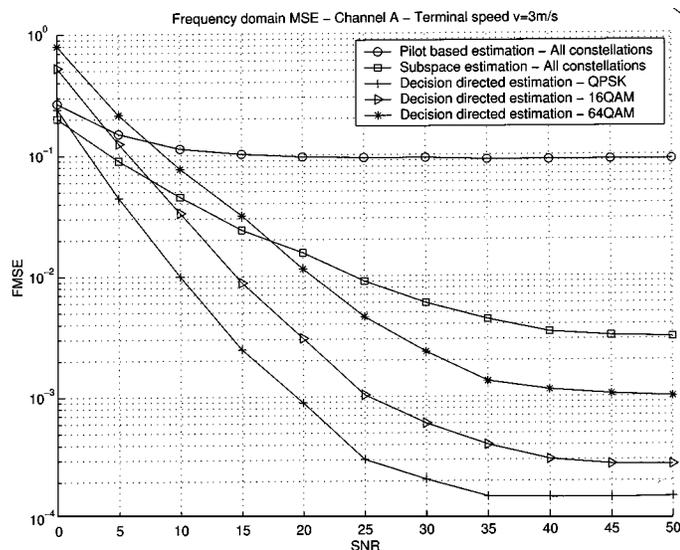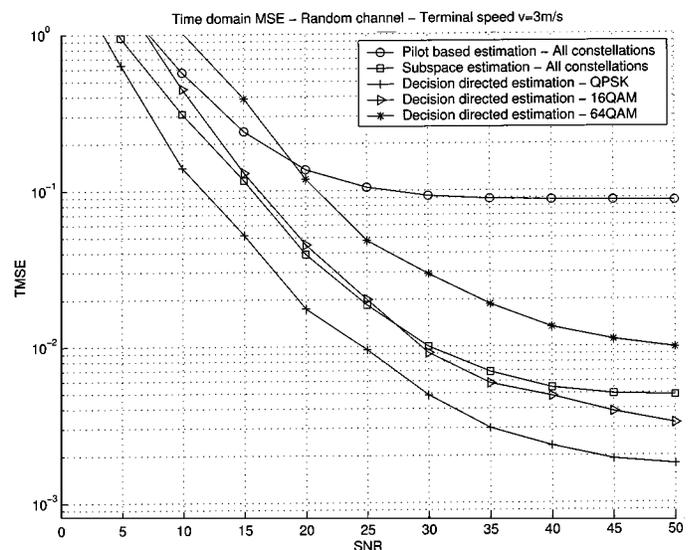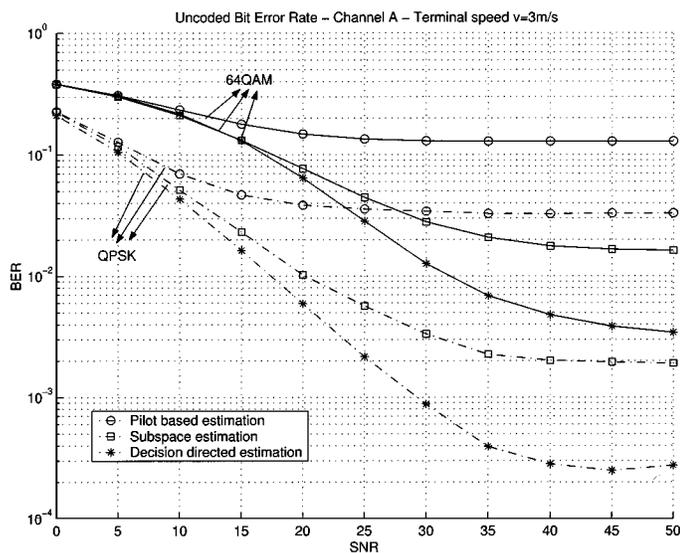

Fig. 6. Time domain MSE as a function of E_s/N_0. Channel A: $v = 3$ m/s.

and frequency domain and the average uncoded BER as a function of the SNR (quantities are averaged over the full frame of 500 OFDM symbols for many channel realizations) for the *channel model A*. The conclusions drawn previously for $E_s/N_0 = 25$ dB can be generalized. The subspace algorithm does well for channel estimation in the TD (it outperforms the DD estimator for 16 QAM and 64 QAM) and significantly improves the estimation provided at the beginning of the frame. On the other hand, the DD is preferable for channel estimation in the FD. Since the equalization is performed in the FD in our setup, this translates into the BER curves of Fig. 8, which shows that the BER obtained using the subspace algorithm, although significantly improved compared with the case when no refinement is performed, is outperformed by the DD, even for 64 QAM.

To provide a more general insight on the relative merits of the subspace algorithm compared to the DD algorithm, Figs. 9–11 depict the same quantities for a *random FIR channel* with
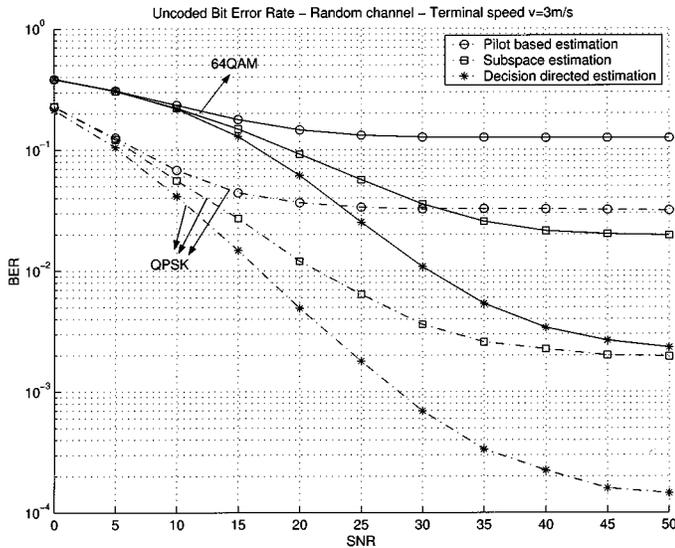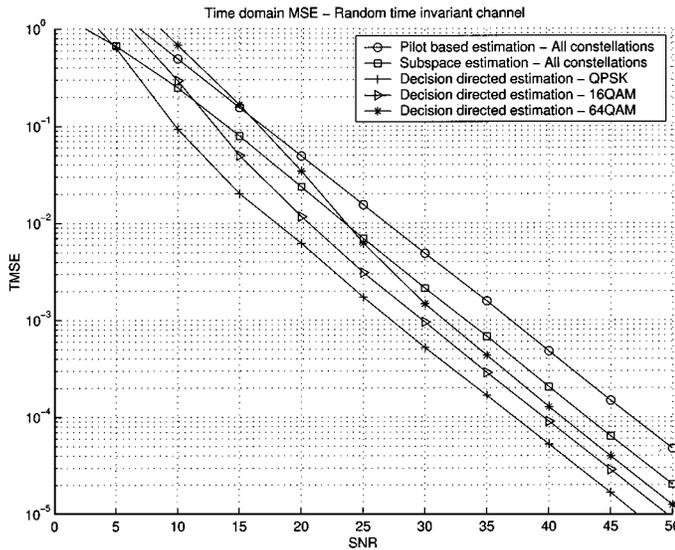
Fig. 7. FD MSE as a function of E_s/N_0. Channel A: $v = 3$ m/s.

Fig. 8. BER as a function of E_s/N_0. Channel A: $v = 3$ m/s.

Fig. 9. Time domain MSE as a function of E_s/N_0. Random channel: $v = 3$ m/s.
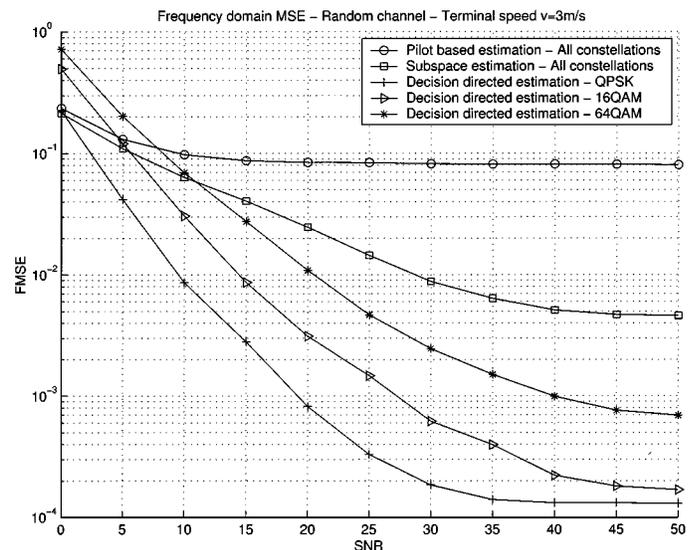
Fig. 10. FD MSE as a function of E_s/N_0. Random channel: $v = 3$ m/s.

the random channel order upper bounded by the cyclic prefix length. This scenario is more general but does not take into account the fact that the channel usually has a decreasing statistical power profile, as channel model A does. With this new scenario, DD is slightly favored compared with the subspace algorithm, but the conclusions drawn for channel A extend to this new simulation scenario: The subspace algorithm offers good performance for time domain estimation and a moderate one for frequency domain estimation. Moreover, it can be inferred from the comparison between results with channel A and the random channel that the subspace algorithm performance is enhanced when the channel has a decreasing power profile. This is probably due to better matrix conditioning.

Note that both the subspace and the DD suffers from error floor phenomenon with time-varying channels. Actually, this can be explained by the number of observations required to obtain a channel estimation in the noiseless case (about $2K$). Thus, at high SNR, the error floor occurs due to the channel variations,

which cannot be accurately tracked since the algorithm requires an averaging window of length greater than $2K$. The DD also exhibits an error floor, but it is easily possible to lower it by reducing the number of symbols used for the averaging as the SNR increases.
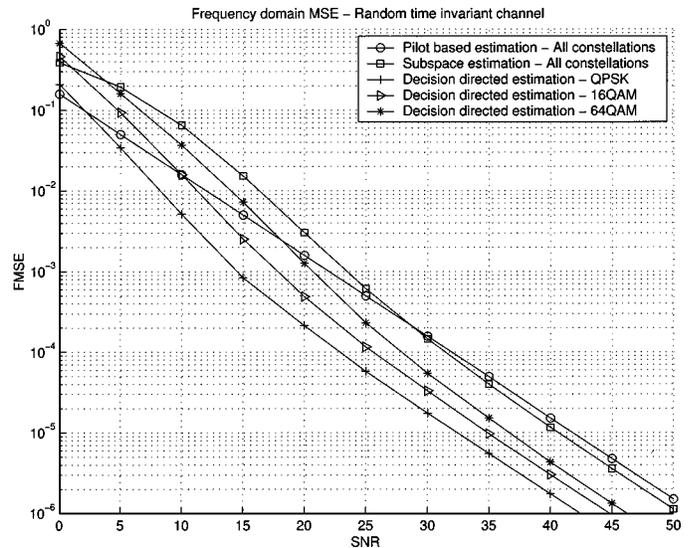
Finally, we provide in Figs. 12–14 some curves obtained for a *static random channel*. As with time-varying channels, the subspace algorithm is better for channel estimation in the time domain rather than in the frequency domain but is also outperformed by the DD algorithm. Note, however, that BER performances are almost the same with or without channel tracking and, hence, that both the DD and the subspace algorithm are useless in a static channel context. Indeed, with static channels, the only errors to occur are due to the thermal noise because the channel estimation obtained at the beginning of the frame using pilot symbols is accurate enough to avoid errors due to an inaccurate channel estimation.

Fig. 11.   BER as a function of E_s/N_0 . Random channel: $v = 3$ m/s.

Fig. 13.   FD MSE as a function of E_s/N_0. Random static channel.

Fig. 12.   Time domain MSE as a function of E_s/N_0. Random static channel.

Fig. 14.   BER as a function of E_s/N_0. Random static channel.

D. Discussion

The proposed method has been shown to offer good performance and to compare favorably with the DD algorithm for channel estimation in the time domain. For channel estimation in the frequency domain, even though it works and does improve the channel estimation accuracy, it is outperformed by the DD algorithm. This loss in performance is mainly due to the fact that the proposed subspace algorithm, as does any subspace method, suffers of important error floors phenomenon due to the averaging window involved by the autocorrelation matrix estimation. This limitates the interest of subspace methods compared with DD approaches to practical applications with slow-varying channels, provided that one accepts to be subject to error propagation phenomena, which may happen with DD algorithms.

Besides, this conclusion concerning the frequency domain equalization must be made mild because this is only a first attempt to apply that class of methods to a real context and because only a basic implementation has been considered herein.
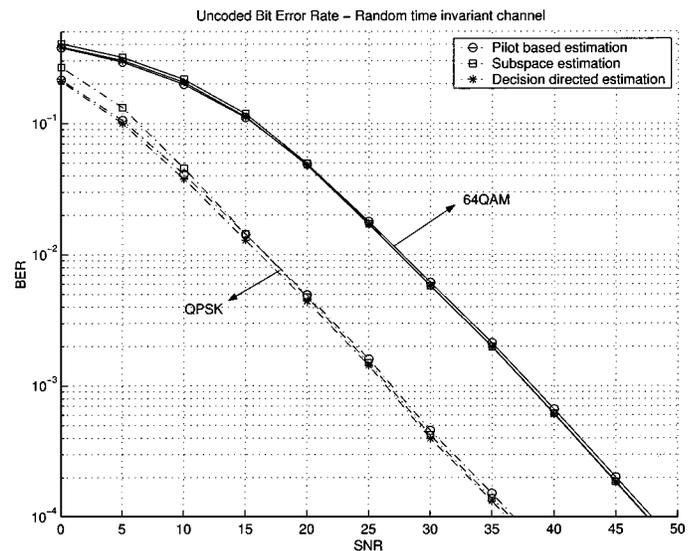
Performance could be improved following the steps proposed in [23] and [30]. In any case, even in this frequency domain equalization framework, the proposed method does work in practice, even if the performance is not always as good as that of the DD one. Moreover, since the performance of that class of method does not depend on the constellation size, they could be useful for applications using varying or very large size constellations (e.g., 256 QAM) [3] and/or constellations unknown to the receiver. Further, the subspace method also has the interesting characteristic to be able to work in a fully blind context.

More important, the subspace method has the very attractive feature of estimating directly the channel impulse response, whereas the DD first estimates the frequency domain channel attenuations. This explains why the DD algorithms often perform worse for channel estimation in the time domain. Indeed, the channel attenuations on the guard carriers cannot be estimated with the DD yielding an inaccurate estimation in the time domain.

Furthermore, unlike DD, this feature allows us to extend the method to identify channels with length greater than the CP, which is a problem that occurs, for example, in digital subscriber line (DSL) contexts [3]. This can be useful because a shortening of the time domain channel impulse response [14] is usually performed at the receiver in that case. Thus, this requires us to estimate the taps located after the CP, which is difficult with methods operating in the frequency domain. In contrast, the proposed subspace algorithm [31] has been extended toward this aim, as reported in [32] and [33].

Finally, it is important to note that the subspace approach relies on the use of a redundancy (the cyclic prefix) whose structure is imposed by the transmitter and is designed for equalization purposes but not for channel estimation. This may be the most limiting factor of that kind of method since it imposes both the minimal size of the averaging window as well as the conditioning of the matrix to be decomposed by SVD. In [34], a subspace method is proposed for linearly block-precoded OFDM systems [35] in which the redundancy can be arbitrarily chosen since it is introduced by the block precoder. This allows us to obtain improved performance and to use the method in realistic contexts, provided that the compatibility constraint is relaxed.

VI. CONCLUSIONS

This paper has presented a new blind channel estimation method for OFDM systems. Making use of the redundancy introduced by the cyclic prefix to identify the channel, it preserves the classical OFDM transmitter structure and, thus, applies to most existing systems. The method can operate in a fully blind context and does not require initialization. It can also be used to improve the estimation obtained from pilot symbols using semi-blind procedures, as proposed in this paper. The most important feature of the method is that it estimates directly the channel impulse response rather than the channel attenuations on subcarriers. Thus, unlike decision-directed algorithms, the proposed algorithm can be extended to estimate channels longer than the cyclic prefix, which is important for channel impulse response shortening algorithms. Simulations have shown that the proposed method offers good performance in practice, especially for channel estimation in the time domain. For channel estimation in the frequency domain, some limitations may reduce the practical impact of subspace approaches compared with decision-directed algorithms estimating directly the channel attenuations from symbol decisions.

APPENDIX A
IMPLEMENTATION OF THE SUBSPACE ALGORITHM FOR ANY VALUE OF L, K, OR M

In this Appendix, we detail how the subspace algorithm can be implemented for any value of M and L verifying $M \geq 2L$.

Consider the I/O relationship (1): $\mathbf{r}_{cp}(k) = \mathbf{H}_0 \mathbf{s}_{cp}(k) + \mathbf{H}_1 \mathbf{s}_{cp}(k-1) + \mathbf{n}_P(k)$. Due to the CP, the vector $\mathbf{s}_{cp}(k)$ can be split into three subvectors of size L, $M-L$, and L, respectively: $\mathbf{s}_{cp}(k) = (\mathbf{s}_0(k)^T, \mathbf{s}_1(k)^T, \mathbf{s}_2(k)^T)^T$ with $\mathbf{s}_0(k) = \mathbf{s}_2(k)$. Similarly, both $\mathbf{r}_{cp}(k)$ and $\mathbf{n}_P(k)$ can be split into three subvectors of size L, $M-L$, and L as $\mathbf{r}_{cp}(k) = (\mathbf{r}_0(k)^T, \mathbf{r}_1(k)^T, \mathbf{r}_2(k)^T)^T$

and $\mathbf{n}_P(k) = (\mathbf{n}_0(k)^T, \mathbf{n}_1(k)^T, \mathbf{n}_2(k)^T)^T$. From these decompositions, each of matrices $\mathbf{H}_0$ and $\mathbf{H}_1$ can be split into nine submatrices, leading to the following relationship:

$$
\begin{aligned}
\mathbf{r}_{cp}(k) &= [\mathbf{r}_0(k), \mathbf{r}_1(k), \mathbf{r}_2(k)]^T \\
&= \begin{bmatrix} \mathbf{C}_0 & 0 & 0 \\ \mathbf{C}'_1 & \mathbf{C}'_0 & 0 \\ 0 & \mathbf{C}''_1 & C_0 \end{bmatrix} \begin{bmatrix} \mathbf{s}_2(k) \\ \mathbf{s}_1(k) \\ \mathbf{s}_2(k) \end{bmatrix} \\
&\quad + \begin{bmatrix} 0 & 0 & \mathbf{C}_1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{s}_2(k-1) \\ \mathbf{s}_1(k-1) \\ \mathbf{s}_2(k-1) \end{bmatrix} \\
&\quad + [\mathbf{n}_0(k), \mathbf{n}_1(k), \mathbf{n}_2(k)]^T
\end{aligned}
\tag{26}
$$

where matrices $\mathbf{C}_0$ and $\mathbf{C}_1$ are defined as previously, where $\mathbf{C}'_0$ is the $(M-L) \times (M-L)$ Toeplitz matrix with first column $[h_0, \ldots, h_L, 0, \ldots, 0]^T$ and first row $[h_0, 0, \ldots 0]^T$ and where $\mathbf{C}'_1$ and $\mathbf{C}''_1$ are defined as $\mathbf{C}'_1 = [\mathbf{C}_1^T, \mathbf{0}_{L \times (M-2L)}]^T$ and $\mathbf{C}''_1 = [\mathbf{0}_{L \times (M-2L)}, \mathbf{C}_1]$. Following the lines of Section III, we can define $\bar{\mathbf{r}}(k)$ as

$$
\bar{\mathbf{r}}(k) = [\mathbf{r}_1(k-1)^T, \mathbf{r}_2(k-1)^T, \mathbf{r}_0(k)^T, \mathbf{r}_1(k)^T, \mathbf{r}_2(k)^T]^T
\tag{27}
$$

to get the relationship that follows, from which it is possible to derive the subspace algorithm:

$$
\begin{aligned}
\bar{\mathbf{r}}(k) &= [\mathbf{r}_1(k-1), \mathbf{r}_2(k-1), \mathbf{r}_0(k), \mathbf{r}_1(k), \mathbf{r}_2(k)]^T \\
&= \begin{bmatrix} \mathbf{C}'_0 & \mathbf{C}'_1 & 0 & 0 \\ \mathbf{C}''_1 & \mathbf{C}_0 & 0 & 0 \\ 0 & \mathbf{C}_1 & 0 & \mathbf{C}_0 \\ 0 & 0 & \mathbf{C}'_0 & \mathbf{C}'_1 \\ 0 & 0 & \mathbf{C}''_1 & \mathbf{C}_0 \end{bmatrix} \begin{bmatrix} \mathbf{s}_1(k-1) \\ \mathbf{s}_2(k-1) \\ \mathbf{s}_1(k) \\ \mathbf{s}_2(k) \end{bmatrix} \\
&\quad + [\mathbf{n}_1(k-1), \mathbf{n}_2(k-1), \mathbf{n}_0(k), \mathbf{n}_1(k), \mathbf{n}_2(k)]^T.
\end{aligned}
\tag{28}
$$

APPENDIX B
IDENTIFIABILITY ISSUES

This appendix focuses on the noise subspace of matrix $\mathbf{H}(\mathbf{h})$ and provides some results concerning identifiability.

A. Structure and Dimension of the Noise Subspace

Since permuting the columns of a matrix changes neither its signal subspace nor its noise subspace, the demonstration is conducted in the following with the matrix $\mathbf{H}(\mathbf{h})$ defined in (29) instead of the matrix defined in (5) since it simplifies the later developments.

$$
\mathbf{H}(\mathbf{h}) = \begin{bmatrix}
\mathbf{C}_0 & 0 & 0 & \mathbf{C}_1 & 0 & 0 & 0 & 0 \\
\mathbf{C}_1 & \mathbf{C}_0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \mathbf{C}_1 & \mathbf{C}_0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \mathbf{C}_1 & \mathbf{C}_0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \mathbf{C}_1 & \mathbf{C}_0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \mathbf{C}_1 & \mathbf{C}_0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \mathbf{C}_1 & \mathbf{C}_0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \mathbf{C}_1 & \mathbf{C}_0 \\
0 & 0 & 0 & 0 & \mathbf{C}_0 & 0 & 0 & \mathbf{C}_1
\end{bmatrix}.
\tag{29}
$$

Consider a vector $\mathbf{g}_i = [\mathbf{g}_i^{1T}, \ldots, \mathbf{g}_i^{9T}]^T$ in the noise subspace of matrix $\mathbf{H}(\mathbf{h})$. Let $\mathbf{C}$ be the $2L \times L$ matrix defined as $\mathbf{C} = [\mathbf{C}_0^T, \mathbf{C}_1^T]^T$; the orthogonality relationship $\mathbf{g}_i^H \mathbf{H}(\mathbf{h}) = \mathbf{0}$ is equivalent to

$$\left[\mathbf{g}_i^{1H}, \mathbf{g}_i^{2H}\right] \mathbf{C} = \mathbf{0} \tag{30}$$

$$\left[\mathbf{g}_i^{2H}, \mathbf{g}_i^{3H}\right] \mathbf{C} = \mathbf{0} \tag{31}$$

$$\left[\mathbf{g}_i^{3H}, \mathbf{g}_i^{4H}\right] \mathbf{C} = \mathbf{0} \tag{32}$$

$$\left[\mathbf{g}_i^{4H}, \mathbf{g}_i^{5H} + \mathbf{g}_i^{1H}\right] \mathbf{C} = \mathbf{0} \tag{33}$$

$$\left[\mathbf{g}_i^{5H} + \mathbf{g}_i^{9H}, \mathbf{g}_i^{6H}\right] \mathbf{C} = \mathbf{0} \tag{34}$$

$$\left[\mathbf{g}_i^{6H}, \mathbf{g}_i^{7H}\right] \mathbf{C} = \mathbf{0} \tag{35}$$

$$\left[\mathbf{g}_i^{7H}, \mathbf{g}_i^{8H}\right] \mathbf{C} = \mathbf{0} \tag{36}$$

$$\left[\mathbf{g}_i^{8H}, \mathbf{g}_i^{9H}\right] \mathbf{C} = \mathbf{0}. \tag{37}$$

Let $H(z) = \sum_{i=0}^{L} h_i z^{-i}$ be the channel transfer function and $p_1, \ldots, p_L$ be its L roots, which are assumed to be distinct [if $H(z)$ has multiple roots, it is easy to extend the demonstration using the generalized Vandermonde vectors [36]], and let $\mathbf{p}_i$ be the Vandermonde vector of size $2L$ associated with p_i: $\mathbf{p}_i = \left[1, p_i^{-1} \ldots, p_i^{-(2L-1)}\right]^H$. The left null space of matrix $\mathbf{C}$ is spanned by the L Vandermonde vectors $\mathbf{p}_1, \ldots, \mathbf{p}_L$ associated with the L roots of $H(z)$. Let $\mathbf{V}$ and $\mathbf{D}$ be the $L \times L$ matrices defined as

$$\mathbf{V} = \begin{bmatrix} p_1^{-L} & \cdots & p_L^{-L} \\ \vdots & & \vdots \\ p_1^{-(2L-2)} & \cdots & p_L^{-(2L-2)} \\ p_1^{-(2L-1)} & \cdots & p_L^{-(2L-1)} \end{bmatrix} \tag{38}$$

and

$$\mathbf{D} = \begin{bmatrix} p_1^L & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & p_L^L \end{bmatrix}. \tag{39}$$

A basis for the left null space of matrix $\mathbf{C}$ is given by the columns of matrix $\left[(\mathbf{VD})^T \ \mathbf{V}^T\right]^T = [\mathbf{p}_1, \ldots, \mathbf{p}_L]$. Hence, for any vector $\mathbf{g}_i$ in the noise subspace, there exists a size L vector $\mathbf{a}_i^{12}$ such that $\left[\mathbf{g}_i^{1T} \ \mathbf{g}_i^{2T}\right]^T = \left[(\mathbf{VD})^T \ \mathbf{V}^T\right]^T \mathbf{a}_i^{12}$. Proceeding similarly with (31) to (36), it can be shown that any vector in the noise subspace is only defined by two size L vectors $\mathbf{a}_i = [a_i(1), \ldots, a_i(L)]^T$ and $\mathbf{b}_i = [b_i(1), \ldots, b_i(L)]^T$ since it has the following structure:

$$\mathbf{g} = \begin{bmatrix} \mathbf{g}_i^1 \\ \mathbf{g}_i^2 \\ \mathbf{g}_i^3 \\ \mathbf{g}_i^4 \\ \mathbf{g}_i^5 \\ \mathbf{g}_i^6 \\ \mathbf{g}_i^7 \\ \mathbf{g}_i^8 \\ \mathbf{g}_i^9 \end{bmatrix} = \begin{bmatrix} \mathbf{VD}^4 \mathbf{a}_i \\ \mathbf{VD}^3 \mathbf{a}_i \\ \mathbf{VD}^2 \mathbf{a}_i \\ \mathbf{VD} \mathbf{a}_i \\ \mathbf{V}(\mathbf{I}_L - \mathbf{D}^4)\mathbf{a}_i = \mathbf{V}(\mathbf{D}^4 - \mathbf{I}_L)\mathbf{b}_i \\ \mathbf{VD}^3 \mathbf{b}_i \\ \mathbf{VD}^2 \mathbf{b}_i \\ \mathbf{VD} \mathbf{b}_i \\ \mathbf{V} \mathbf{b}_i \end{bmatrix}. \tag{40}$$

If no channel zero is located on subcarriers (i.e., $p_i^N \neq 1, \forall i$), then matrix $(\mathbf{I}_L - \mathbf{D}^4)$ is invertible, and $\forall i, \mathbf{a}_i = -\mathbf{b}_i$. Thus, any vector $\mathbf{g}_i$ in the noise subspace is uniquely defined by $\mathbf{b}_i$. Hence, the noise subspace of $\mathbf{H}(\mathbf{h})$ has dimension L, and $\mathbf{H}(\mathbf{h})$ is full column rank. If there are $Z \leq L$ zeros located among the M subcarriers, let $Z = \{k_1, \ldots, k_Z\}$ be the set of index corresponding to these zeros (that is, such that $p_{k_i}^M = 1$). In this case, $(\mathbf{I}_L - \mathbf{D}^4)$ is no longer invertible. However, $\mathbf{V}(\mathbf{I}_L - \mathbf{D}^4)\mathbf{a}_i = \mathbf{V}(\mathbf{D}^4 - \mathbf{I}_L)\mathbf{b}_i$ involves that $a_i(k) = -b_i(k)$ for each $k \notin \mathcal{Z}$. This proves that the noise subspace of $\mathbf{H}(\mathbf{h})$ has dimension $L + Z$ since any vector $\mathbf{g}_i$ in the noise subspace is uniquely defined by the size $L + Z$ vector $\mathbf{c}_i = [a_i(k_1), \ldots, a_i(k_Z), \mathbf{b}_i^T]^T$.

B. Uniqueness of the Solution Using the Entire Noise Subspace

Let $\mathbf{h}' = [h_0' \ldots h_L']$ be a vector such that $\mathbf{g}_i^H \mathbf{H}(\mathbf{h}') = \mathbf{0}$ for $1 \leq i \leq L + Z$, where $\mathbf{G} = [\mathbf{g}_1, \ldots, \mathbf{g}_{L+Z}]$ is a basis for the noise subspace. $\mathbf{G}^H \mathbf{H}(\mathbf{h}') = \mathbf{0}$ is equivalent to $\mathbf{G}_c^H \mathbf{H}(\mathbf{h}') = \mathbf{0}$, where $\mathbf{G}_c$ stands for the canonical basis corresponding to the vectors $\mathbf{c}_i = \mathbf{1}_i^{L+Z}$ for $1 \leq i \leq L + Z$ (note that if $Z = 0$, $\mathbf{c}_i$ simply reduces to $\mathbf{b}_i$), where $\mathbf{1}_i^S$ denotes the vector of size S with 1 at position i and 0 elsewhere. Considering (37), $\mathbf{G}_c^H \mathbf{H}(\mathbf{h}') = \mathbf{0}$ involves

$$\mathbf{1}_i^{LH} \left[\mathbf{D}^H \mathbf{V}^H, \mathbf{V}^H\right] \begin{bmatrix} \mathbf{C}_0(\mathbf{h}') \\ \mathbf{C}_1(\mathbf{h}') \end{bmatrix} = \mathbf{0} \text{ for } 1 \leq i \leq L$$

$$\Longleftrightarrow$$

$$\left[\mathbf{D}^H \mathbf{V}^H, \mathbf{V}^H\right] \begin{bmatrix} \mathbf{C}_0(\mathbf{h}') \\ \mathbf{C}_1(\mathbf{h}') \end{bmatrix} = \mathbf{0} \tag{41}$$

and hence, $H'(z) := \sum_{i=0}^{L} h_i' z^{-i}$ shares the same roots as $H(z)$, which proves that $\mathbf{h}' = \alpha \mathbf{h}$.

C. Channel Estimation Using Only a Part of the Noise Subspace

It has been proven that the channel can be uniquely identified under the condition that the noise subspace dimension $(L+Z)$ is known. However, the number Z of zeros located on subcarriers can only be upper bounded by L since the channel is unknown, and the noise subspace dimension can only be assumed to be greater or equal to L. Thus, it is of interest to know if uniqueness is guaranteed if only L independent vectors of the noise subspace are considered.

Uniqueness is ensured if no zero is located on subcarriers because the entire noise subspace is considered in that case. Let us assume now that one zero is located on a subcarrier (the following developments can easily be extended to the case where there are several zeros located on the subcarrier) and assume without any loss of generality that $p_1^M = 1$. In this case, any vector in the noise subspace $\mathbf{g}_i$ is uniquely specified by a size $L + 1$ vector of size $\mathbf{c}_i = [a_i(1), \mathbf{b}_i^T]^T$. If L independent vectors $\mathbf{g}_i$ for $1 \leq i \leq L$ of the noise subspace are considered, the corresponding vectors $\mathbf{c}_i$ for $1 \leq i \leq L$ are independent. Hence, (41) holds, which proves that $\mathbf{h}' = \alpha \mathbf{h}$. If the vectors $\mathbf{b}_i$ for $1 \leq i \leq L$ are not independent, a vector $\mathbf{h}'$ satisfying $\mathbf{g}_i^H \mathbf{H}(\mathbf{h}') = \mathbf{0}$ for $1 \leq i \leq L$ such that $\mathbf{h}' \neq \alpha \mathbf{h}$ can be found. For example, let us assume that the L independent vectors $\mathbf{g}_i$ used to identify the channel are the vectors $\mathbf{g}_i^c$ of the canonical basis $\mathbf{G}_c$ for $1 \leq i \leq L + 1$ and $i \neq 3$. Let $\mathbf{h}'$ be the size-L vectors of norm 1 defined by the data of the $L - 1$ roots p_i'

of $H'(z)$ with $p'_i = p_i$ for $i \neq 2$ and $p'_2 \neq p_2$. In that case, $(\mathbf{g}_1^c)^H \mathbf{H}(\mathbf{h}') = \mathbf{0}$ and $(\mathbf{g}_2^c)^H \mathbf{H}(\mathbf{h}') = \mathbf{0}$ since p_1 is a root of $H'(z)$. In addition, $(\mathbf{g}_i^c)^H \mathbf{H}(\mathbf{h}') = \mathbf{0}$ for $4 \leq i \leq L + 1$ since $p_3, \ldots, p_L$ are roots of $H'(z)$. Hence, $\mathbf{h}' \neq \alpha \mathbf{h}$ but satisfies the orthogonality relations. Thus, identifiability is not guaranteed when some channel zeros are located on subcarriers.

Besides, the number R of different roots between $H(z)$ and $H'(z)$ is smaller than Z, which is proved in the following when $Z = 1$ (the demonstration can be extended to $Z > 1$). In order to do this, let us compute the spectral decomposition of matrix $\mathbf{B} := [\mathbf{b}_1, \ldots, \mathbf{b}_L]$, which is given by $\mathbf{B} = \mathbf{U}_B^H \Delta \mathbf{V}_B$, where $\mathbf{U}_B$ and $\mathbf{V}_B$ are square orthogonal matrices of size $L \times L$ and where Δ is a diagonal matrix with main diagonal entries $d_1, \ldots, d_L$ with $d_1 \geq \cdots \geq d_L$. From (37), $\Delta \mathbf{U}_B [(\mathbf{VD})^H, \mathbf{V}^H] \begin{bmatrix} \mathbf{C}_0(\mathbf{h}') \\ \mathbf{C}_1(\mathbf{h}') \end{bmatrix} = \mathbf{0}$. If the L eigenvectors $\mathbf{b}_1, \ldots, \mathbf{b}_L$ are not independent, then matrix Δ is noninvertible, and $d_1 \geq \cdots \geq d_{L-1} > d_L = 0$. In that case, it can be shown that for each $i \in \{1, \ldots, L - 1\}$, $\sum_{j=1}^{L-1} [\mathbf{U}_B]_{i,j} \mathbf{p}_j$ is a linear combination of the Vandermonde vectors $\mathbf{p}'_i$ associated with the roots p'_i of $H'(z)$. Let us assume that $R > 1$, and consider, without any loss of generality, that $p'_i \neq p_i$ for $1 \leq i \leq R$. Using the fact that the columns of a tall Vandermonde matrix built from distinct roots are linearly independent, it can be shown that for each $i \in \{1, \ldots, D - 1\}$, $[\mathbf{U}_B]_{i,1} = [\mathbf{U}_B]_{i,2} = 0$, and therefore, that the two first columns of $\mathbf{U}_B$ are proportional, which is impossible because $\mathbf{U}_B$ is orthogonal. Hence, $R \leq Z = 1$.

D. Identifiability With the Semi-Blind Algorithm

Consider the semi-blind algorithm defined by (15), and assume that the number Z of channel zeros located on subcarriers is smaller than the number of pilot subcarriers F. Let $\mathbf{h}'$ be a vector satisfying (15). Equation (15) involves (6), and hence, at most Z roots of $H'(z)$ are different from those of $H(z)$. Denote these $R < F$ roots as $p_1, \ldots, p_R$ and $p'_1, \ldots, p'_R$ for $H(z)$ and $H'(z)$, respectively. Since both $\mathbf{h}$ and $\mathbf{h}'$ must satisfy $\mathbf{F}^{\mathrm{pil}} \mathbf{h} = \tilde{\mathbf{h}}^{\mathrm{pil}}$ for the pilot subcarriers frequencies $p_1^{\mathrm{pil}}, \ldots, p_F^{\mathrm{pil}}$, $\alpha_1 \prod_{i=1}^{R} (p_j^{\mathrm{pil}} - p_i) = \alpha_2 \prod_{i=1}^{R} (p_j^{\mathrm{pil}} - p'_i)$ must hold for every $j \in \{1, \ldots, F\}$, where α_1 and α_2 are two nonzero normalization constants. Hence, every p_j^{pil} is a root of the degree-R polynomial $P(x) = \alpha_1 \prod_{i=1}^{R} (x - p_i) - \alpha_2 \prod_{i=1}^{R} (x - p'_i)$. Therefore, $P(x)$ is equal to zero for any value of x since it is a degree-$R < F$ polynomial with F distinct roots. Thus, α_1 must be equal to α_2, and $p_1, \ldots, p_R$ must be equal to $p'_1, \ldots, p'_R$ up to a permutation, and hence, $\mathbf{h}' = \mathbf{h}$. Therefore, the uniqueness of the solution provided by (15) is ensured if the number of channel zeros located on the unit circle is smaller than the number of pilot subcarriers.

REFERENCES

[1] "Radio broadcasting system, digital audio broadcasting (DAB) to mobile, portable, and fixed receivers," Eur. Telecommun. Stand. Inst., Sophia-Antipolis, Valbonne, France, ETS 300 401, 1995–1997.

[2] "Digital broadcasting system television, sound, and data services; Framing structure, channel coding, and modulation digital terrestrial television," Eur. Telecommun. Stand. Inst., Sophia-Antipolis, Valbonne, France, ETS 300 744, 1996.

[3] "The DWMT: A multicarrier transceiver for ADSL using M-band wavelets," ANSI Stand. T1E1.4 Comm. Contrib., 1993.

[4] "Broadband radio access networks (BRAN); High performance radio local area networks (HIPERLAN) Type 2; System overview," Eur. Telecomm. Stand. Inst., Sophia-Antipolis, Valbonne, France, ETR101 683 114, 1999.

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

[6] H. Liu, G. Xu, L. Tong, and T. Kailath, "Recent developments in blind channel equalization : From cyclostationnarity to subspaces," *Signal Process.*, vol. 50, no. 1-2, pp. 83–99, Apr 1996.

[7] M. de Courville, P. Duhamel, P. Madec, and J. Palicot, "Blind equalization of OFDM systems based on the minimization of a quadratic criterion," in *Proc. Int. Conf. Commun.*, vol. 3, Dallas, TX, June 1996, pp. 1318–1321.

[8] G. B. Giannakis, "Filterbanks for blind channel identification and equalization," *IEEE Signal Processing Lett.*, vol. 4, pp. 184–187, June 1997.

[9] R. Heath and G. B. Giannakis, "Exploiting input cyclostationnarity for blind channel identification in OFDM systems," *IEEE Trans. Signal Processing*, vol. 47, pp. 848–856, Mar 1999.

[10] B. Muquet and M. de Courville, "Blind and semi-blind channel identification methods using second order statistics for OFDM systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 5, Phoenix, AZ, Mar 1999, pp. 2745–2748.

[11] M. Tsatsanis and G. B. Giannakis, "Transmitter induced cyclostationarity for blind channel equalization," *IEEE Trans. Signal Processing*, vol. 45, pp. 1785–1794, July 1997.

[12] U. Tureli and H. Liu, "Blind carrier synchronization and channel identification for OFDM communications," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 6, Seattle, WA, May 1998, pp. 3509–3512.

[13] X. Wang and K. J. Ray Liu, "Adaptive channel estimation using cyclic prefix in multicarrier modulation system," *IEEE Commun. Lett.*, vol. 3, pp. 291–293, Oct 1999.

[14] P. Melsa, R. C. Younce, and C. E. Rohrs, "Impulse response shortening for discrete multitone transceivers," *IEEE Trans. Commun.*, vol. 44, pp. 1662–1672, Dec 1996.

[15] S. Halford and G. B. Giannakis, "Direct blind equalization for transmitter induced cyclostationnarity," in *Proc. IEEE Workshop Signal Process. Adv. Wireless Commun.*, Paris, France, Apr 1997, pp. 117–120.

[16] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers — Part II : blind channel estimation, synchronization and direct equalization," *IEEE Trans. Signal Processing*, vol. 47, pp. 2007–2022, July 1999.

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

[18] E. de Carvalho and D. Slock, "Cramer-Rao bounds for semi-blind, blind and training sequence based channel estimation," in *Proc. IEEE Workshop Signal Process. Adv. Wireless Commun.*, Paris, France, Apr 1997.

[19] V. Mignone and A. Morello, "CD3-OFDM: A novel demodulation scheme for fixed and mobile receivers," *IEEE Trans. Commun.*, vol. 44, pp. 1144–1151, Sept 1996.

[20] F. Tufvesson and T. Maseng, "Pilot assisted channel Estimation for OFDM in mobile cellular systems," in *Proc. IEEE Veh. Technol. Conf.*, Phoenix, AZ, May 1997, pp. 1639–1643.

[21] M. Alard and R. Lassale, "Principles of modulation and channel coding for digital broadcasting for mobile receivers," *Eur. Broadcasting Union Rev. Tech.*, vol. 224, pp. 168–190, Aug 1987.

[22] G. Xu, H. Lui, L. Tong, and T. Kailath, "A least-squares approach to blind channel identification," *IEEE Trans. Signal Processing*, vol. 43, pp. 2982–2993, Dec 1995.

[23] V. Buchoux, O. Cappé, E. Moulines, and A. Gorokhov, "On the performance of semi-blind subspace-based channel estimation," *IEEE Trans. Signal Processing*, vol. 48, pp. 1750–1759, June 2000.

[24] A. Gorokhov and P. Loubaton, "Semi-blind second order identification of convolutive channels," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 5, Munich, Germany, Apr 1997, pp. 3905–3908.

[25] O. Edfors, M. Sandell, J. van de Beek, S. Wilson, and P. Borjesson, "Analysis of DFT-based channel estimators for OFDM," Lulea Univ. Technol., Lulea, Sweden, Res. Rep. TULEA, vol. 17, 1996.

[26] P. Frenger and A. Svensson, "Decision-directed coherent detection in multicarrier systems on rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 48, pp. 490–498, Mar 1999.

[27] B. Muquet, Z. Wang, G. B. Giannakis, M. de Courville, and P. Duhamel, "Cyclic-prefix or zero-padding for multicarrier transmissions ?," *IEEE Trans. Commun.*, 2002, to be published.

[28] "Channel Models for HIPERLAN/2 in different indoor scenarios," Eur. Telecommun. Stand. Inst., Sophia-Antipolis, Valbonne, France, 3ERI085B, 1998.

[29] J. van de Beek, O. Edfors, M. Sandell, S. Wilson, and P. Borjesson, "On channel estimation in OFDM systems," in *Proc. IEEE Veh. Technol. Conf.*, vol. 2, Chicago, IL, July 1995, pp. 815–819.

[30] K. Abed-Meraim, J. F. Cardoso, A. Gorokhov, P. Loubaton, and E. Moulines, "On subspace methods for blind identification of single-input multiple-output FIR systems," *IEEE Trans. Signal Processing*, vol. 45, pp. 42–55, Jan 1997.

[31] B. Muquet, M. de Courville, P. Duhamel, and V. Buzenac, "A subspace based blind and semi-blind channel identification method for OFDM systems," in *Proc. SPAWC*, May 1999, pp. 170–173.

[32] X. Cai and A. Akansu, "A subspace method for blind channel identification in OFDM systems," in *Proc. Int. Conf. Commun.*, vol. 2, New Orleans, LA, June 2000, pp. 929–933.

[33] X. Zhuang, Z. Ding, and A. L. Swindlehurst, "A statistical subspace method for blind channel identification in OFDM communications," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Istanbul, Turkey, June 2000.

[34] G. B. Giannakis, P. Anghel, and Z. Wang, "All-digital unification and equalization of generalized multi-carrier transmissions through frequency-selective uplink channels," *IEEE Trans. Commun.*, Mar 2000, submitted for publication.

[35] Z. Wang and G. B. Giannakis, "Block-precoding for MUI/ISI-resilient generalized multi-carrier CDMA with multirate capabilities," *IEEE Trans. Commun.*, 2001, submitted for publication.

[36] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: John Hopkins Univ. Press, 1983.

Bertrand Muquet (M'01) was born in France in 1973. He received the Engineering degree in electrical engineering from the Ecole Supérieure d'Electricité (Supélec), Gif-sur-Yvette, France, in 1996 and the "Agrégation" degree in applied physics from the Ecole Normale Supérieure de Cachan, Cachan, France, in 1997. He received the Ph.D. degree from Ecole Nationale Supérieure des Télécommunications, Paris, France, in 2001.

From 1998 to 2001, he was a Research Engineer working on OFDM systems at Motorola Labs, Paris. He is now with Stepmind, a company based in France doing wireless communications (GSM, GPRS, EDGE, HIPERLAN/2, IEEE802.11a). His area of expertise lies in signal processing and digital communications with emphasis on multicarrier and OFDM systems, blind channel estimation and equalization, and iterative and turbo algorithms.

Marc de Courville (M'97) was born in Paris, France, on April 21, 1969. He graduated from the Ecole National Supérieure des Télécommunications, Paris (Engineering University) in 1993 and received the Ph.D. degree, from the Ecole Nationale Supérieure des Télécommunications (ENST), Paris, also in 1996.

His research interests include digital communications and digital signal processing. Since 1996, he has been with Motorola Labs, Paris, and is now a Research Team Manager involed in projects dealing with multicarrier systems for current and future generations of wireless local area networks.

Pierre Duhamel (F'98) was born in France in 1953. He received the Ing. degree in electrical engineering from the National Institute for Applied Sciences (INSA), Rennes, France, in 1975 and the Dr.Ing. degree in 1978 and Doctoratès sciences degree in 1986, both from Orsay University, Orsay, France.

From 1975 to 1980, he was with Thomson-CSF, Paris, France, where his research interests were in circuit theory and signal processing, including digital filtering and analog fault diagnosis. In 1980, he joined the National Research Center in Telecommunications (CNET), Issy les Moulineaux, Paris, France, where his research activities were first concerned with the design of recursive CCD filters. Later, he worked on fast algorithms for computing Fourier transforms and convolutions and applied similar techniques to adaptive filtering, spectral analysis, and wavelet transforms. From 1993 to September 2000, he was a Professor with the Ecole Nationale Supérieure des Télécommunications (ENST), Paris, with research activities focused on signal processing for communications. He was Head of the Signal and Image Processing Department from 1997 to 2000. He is now with the Laboratoire de Signaux et Systemes (CNRS/LSS), Gif sur Yvette, France, where he is developing studies in signal processing for communications (including equalization, iterative decoding, and multicarrier systems) and signal/image processing for multimedia applications, including source coding, joint source/channel coding, watermarking, and audio processing.