Optimal Training Design for MIMO OFDM Systems in Mobile Wireless Channels

Imad Barhumi, Geert Leus, *Member, IEEE*, and Marc Moonen, *Member, IEEE*

***Abstract*—This paper describes a least squares (LS) channel estimation scheme for multiple-input multiple-output (MIMO) orthogonal frequency division multiplexing (OFDM) systems based on pilot tones. We first compute the mean square error (MSE) of the LS channel estimate. We then derive optimal pilot sequences and optimal placement of the pilot tones with respect to this MSE. It is shown that the optimal pilot sequences are equipowered, equispaced, and phase shift orthogonal. To reduce the training overhead, an LS channel estimation scheme over multiple OFDM symbols is also discussed. Moreover, to enhance channel estimation, a recursive LS (RLS) algorithm is proposed, for which we derive the optimal forgetting or tracking factor. This factor is found to be a function of both the noise variance and the channel Doppler spread. Through simulations, it is shown that the optimal pilot sequences derived in this paper outperform both the orthogonal and random pilot sequences. It is also shown that a considerable gain in signal-to-noise ratio (SNR) can be obtained by using the RLS algorithm, especially in slowly time-varying channels.**

***Index Terms*—Channel estimation, MIMO, multipath fading channels, OFDM.**

I. Introduction

HIGH-DATA rate techniques in communication systems have gained considerable interest in recent years. A technique that has attracted a lot of attention is orthogonal frequency division multiplexing (OFDM), which is a multicarrier modulation technique. This is due to its simple implementation, and robustness against frequency-selective fading channels, which is obtained by converting the channel into flat fading subchannels. OFDM has been standardized for a variety of applications, such as digital audio broadcasting (DAB), digital television broadcasting, wireless local area networks (WLANs), and asymmetric digital subscriber lines (ADSLs). Combining OFDM with multiple antennas has been shown to provide a significant increase in capacity through the use of transmitter and receiver diversity [8]. However, such systems rely upon the knowledge of channel state information (CSI) at the receiver.

CSI is crucial for data detection and channel equalization. CSI can be obtained in different ways; one is based on training symbols that are *a priori* known at the receiver, whereas the other is blind, i.e., relies only on the received symbols, and acquires CSI by, e.g., exploiting statistical information and/or transmitted symbol properties (like finite alphabet, constant modulus, etc.) [3], [4]. However, compared with training, blind channel estimation generally requires a long data record. Hence, it is limited to slowly time-varying channels and entails high complexity. For these reasons, we restrict our attention to training-based channel estimation in this paper.

Typical procedures for identifying the channel based on training utilize multiple OFDM symbols that consist *completely* of pilot symbols. For single-input single-output (SISO) systems, this approach can be found in [1], [9], and [10], whereas for multiple-input multiple-output (MIMO) systems, it can be found in [5]. In such systems, the CSI is estimated prior to any transmission of data. When the CSI changes significantly, a retraining sequence is transmitted. In a fast time-varying environment, such systems must continuously retrain to re-estimate the CSI. Between retraining, these systems experience an increased BER due to their outdated channel estimates. Wiener filtering (in time and/or frequency) based on a known channel correlation function (in time and/or frequency) can be used to improve the channel estimate [2], [12].

Using *pilot tones* to obtain the CSI was first proposed in [7], where an optimal placement of the pilot tones with regard to (w.r.t.) the mean square error (MSE) of the least squares (LS) channel estimate is proposed for SISO OFDM systems. Extending this idea to MIMO OFDM systems is not straightforward, since not only the placement of the pilot tones but also the pilot sequences themselves must be optimized to obtain the minimal MSE of the LS channel estimate. Note that optimal training for SISO OFDM systems w.r.t. the MSE of the LS channel estimate, and the MSE at the output of a zero-forcing receiver based on the LS channel estimate is discussed in [14]. Optimal training for SISO OFDM systems w.r.t. the capacity based on the linear minimum mean square error (LMMSE) channel estimate is presented in [11].

In this paper, a LS channel estimation scheme for MIMO OFDM systems based on pilot tones is described. First, the MSE of the LS channel estimate is computed. Then, optimal pilot sequences and optimal placement of the pilot tones w.r.t. this MSE are derived. To reduce the training overhead, an LS channel estimation scheme over multiple OFDM symbols is also discussed.

Manuscript received January 28, 2002; revised November 27, 2002. This work was carried out at the ESAT Laboratory of the Katholieke Universiteit Leuven, in the framework of the Belgian State, Prime Minister's Office—Federal Office for Scientific, Technical and Cultural Affairs—Interuniversity Poles of Attraction Programme (2002–2007)—IUAP P5/22 ("Dynamical Systems and Control: Computation, Identification and Modeling") and P5/11 ("Mobile multimedia communication systems and networks"), and the Concerted Research Action GOA-MEFISTO-666 (Mathematical Engineering for Information and Communication Systems Technology) of the Flemish Government under Research Project FWO nr.G.0196.02 ("Design of efficient communication techniques for wireless time-dispersive multi-user MIMO systems"), and was supported in part by the Flemish Interuniversity Microelectronics Center (IMEC). The scientific responsibility is assumed by its authors. The associate editor coordinating the review of this paper and approving it for publication was Dr. Sergios Theodoridis.

The authors are with the ESAT Laboratory, Katholieke Universiteit Leuven, Leuven, Belgium (e-mail: imad.barhumi@esat.kuleuven.ac.be; geert.leus@esat.kuleuven.ac.be; marc.moonen@esat.kuleuven.ac.be).

Digital Object Identifier 10.1109/TSP.2003.811243

1053-587X/03$17.00 © 2003 IEEE

Authorized licensed use limited to: IEEE Xplore. Downloaded on October 13, 2008 at 06:55 from IEEE Xplore. Restrictions apply.

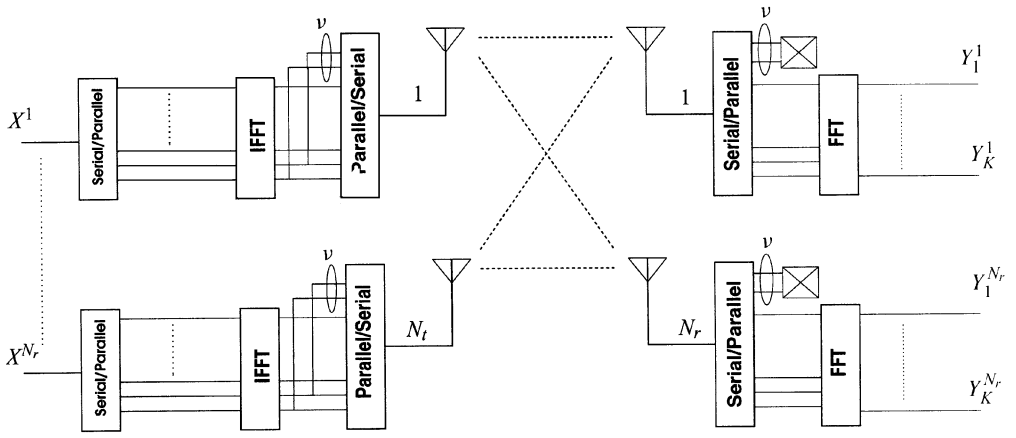


Fig. 1. System model.

Moreover, to enhance channel estimation, a recursive LS (RLS) algorithm is proposed.

This paper is organized as follows. In Section II, we briefly overview the basic system model; in Section III, we introduce LS channel estimation; the analysis of the LS channel estimator is derived in Section IV, through which we derive an optimal training strategy; Section V presents computer simulation results, and finally, conclusions are drawn in Section VI.

Notation: Upper (lower) letters will generally be used for frequency-domain (time-domain) symbols; boldface letters will be used for matrices and column vectors; $(\cdot)^H$ will denote Hermitian (conjugate transpose), $(\cdot)^\dagger$ matrix pseudo inverse, and $\lceil\cdot\rceil$ integer ceiling; $\mathcal{E}\{\cdot\}$ is used to represent expectation and tr$\{\cdot\}$ to represent trace. We will use $[\mathbf{A}]_{n,m}$ to denote the (n, m)th entry of the matrix $\mathbf{A}$; $\mathbf{I}_N$ will denote the $N \times N$ identity matrix and $\mathbf{0}_{N\times M}$ the $N \times M$ all-zero matrix. Further, diag $\{\mathbf{x}\}$ stands for the diagonal matrix with the column vector $\mathbf{x}$ on its diagonal; finally, $\jmath = \sqrt{-1}$.

II. System Model

The system under consideration is depicted in Fig. 1, which shows a MIMO OFDM system with N_t transmit antennas, N_r receive antennas, and K subcarriers. At each transmit (receive) antenna, the conventional OFDM modulator (demodulator) is used. Suppose the OFDM symbol that is transmitted from the rth antenna at time index n is denoted by the $K \times 1$ vector $\mathbf{X}^r(n)$. Before transmission, this vector is processed by an IFFT, and a cyclic prefix of length ν is added. We assume that $\nu \geq L-1$, where L is the maximum length of all channels, which is common practice in wireless communications. After removing the cyclic prefix at the qth receive antenna, we obtain the $K \times 1$ vector $\mathbf{y}^q(n)$, which can be written as [2]

$$\mathbf{y}^q(n) = \sum_{r=1}^{N_t} \mathbf{H}_{\text{cir}}^{q,r} \mathcal{F}^H \mathbf{X}^r(n) + \boldsymbol{\eta}^q(n) \tag{1}$$

where $\mathbf{H}_{\text{cir}}^{q,r}$ is a circulant matrix with first column given by $[\mathbf{h}^{q,r^T}, \mathbf{0}_{1\times(K-L)}]^T$, and $\mathbf{h}^{q,r}$ is the $L \times 1$ vector representing the length L channel impulse response from the rth transmit antenna to the qth receive antenna. Note that $\mathcal{F}$ denotes the $K \times K$ unitary DFT matrix. It is easy to show that the eigenvalue decomposition of $\mathbf{H}_{\text{cir}}^{q,r}$ leads to $\mathbf{H}_{\text{cir}}^{q,r} = \mathcal{F}^H \text{diag}\{\sqrt{K}\mathcal{F}[\mathbf{h}^{q,r^T}, \mathbf{0}_{1\times K-L}]^T\}\mathcal{F}$. Taking the FFT of $\mathbf{y}^q(n)$, we finally obtain

$$\mathbf{Y}^q(n) = \sum_{r=1}^{N_t} \text{diag}\left\{\sqrt{K}\mathcal{F}[\mathbf{h}^{q,r^T}, \mathbf{0}_{1\times(K-L)}]^T\right\} \times \mathbf{X}^r(n) + \boldsymbol{\Xi}^q(n) \tag{2}$$

where $\boldsymbol{\Xi}^q(n) = \mathcal{F}\boldsymbol{\eta}^q(n)$.

III. Least Squares Channel Estimation

In this section, a least squares (LS) channel estimation scheme is derived. Let $\mathbf{X}^r(n) = \mathbf{S}^r(n) + \mathbf{B}^r(n)$ [14], where $\mathbf{S}^r(n)$ is some arbitrary $K \times 1$ data vector, and $\mathbf{B}^r(n)$ is some arbitrary $K \times 1$ pilot sequence vector. Then, (2) can be written as

$$\begin{aligned}\mathbf{Y}^q(n) &= \sum_{r=1}^{N_t} \text{diag}\{\mathbf{X}^r(n)\}\mathbf{F}\mathbf{h}^{q,r} + \boldsymbol{\Xi}^q(n) \\ &= \sum_{r=1}^{N_t} (\text{diag}\{\mathbf{S}^r(n)\} + \text{diag}\{\mathbf{B}^r(n)\}) \\ &\quad \times \mathbf{F}\mathbf{h}^{q,r} + \boldsymbol{\Xi}^q(n)\end{aligned} \tag{3}$$

where $\mathbf{F}$ is $\sqrt{K}$ times the first L columns of $\mathcal{F}$. Defining $\mathbf{S}_{\text{diag}}^r(n) = \text{diag}\{\mathbf{S}^r(n)\}$ and $\mathbf{B}_{\text{diag}}^r(n) = \text{diag}\{\mathbf{B}^r(n)\}$, (3) can be rewritten as

$$\mathbf{Y}^q(n) = \sum_{r=1}^{N_t} \mathbf{S}_{\text{diag}}^r(n)\mathbf{F}\mathbf{h}^{q,r} + \sum_{r=1}^{N_t} \mathbf{B}_{\text{diag}}^r(n)\mathbf{F}\mathbf{h}^{q,r} + \boldsymbol{\Xi}^q(n).$$

Assuming training over g consecutive OFDM symbols, e.g., over the time indices $n \in \{0, \ldots, g-1\}$, we consider the data model

$$\mathbf{Y}^q = \mathbf{T}\mathbf{h}^q + \mathbf{A}\mathbf{h}^q + \boldsymbol{\Xi}^q \tag{4}$$

Authorized licensed use limited to: IEEE Xplore. Downloaded on October 13, 2008 at 06:55 from IEEE Xplore. Restrictions apply.

where $\mathbf{Y}^q = [\mathbf{Y}^{q^T}(0), \ldots, \mathbf{Y}^{q^T}(g-1)]^T$, $\boldsymbol{\Xi}^q = [\boldsymbol{\Xi}^{q^T}(0), \ldots, \boldsymbol{\Xi}^{q^T}(g-1)]^T$

$$\mathbf{T} = \begin{bmatrix} \mathbf{S}^1_{\text{diag}}(0)\mathbf{F} & \cdots & \mathbf{S}^{N_t}_{\text{diag}}(0)\mathbf{F} \\ \vdots & \cdots & \vdots \\ \mathbf{S}^1_{\text{diag}}(g-1)\mathbf{F} & \cdots & \mathbf{S}^{N_t}_{\text{diag}}(g-1)\mathbf{F} \end{bmatrix} \tag{5}$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{B}^1_{\text{diag}}(0)\mathbf{F} & \cdots & \mathbf{B}^{N_t}_{\text{diag}}(0)\mathbf{F} \\ \vdots & \cdots & \vdots \\ \mathbf{B}^1_{\text{diag}}(g-1)\mathbf{F} & \cdots & \mathbf{B}^{N_t}_{\text{diag}}(g-1)\mathbf{F} \end{bmatrix} \tag{6}$$

and $\mathbf{h}^q = [\mathbf{h}^{q,1^T}, \ldots, \mathbf{h}^{q,N_t^T}]^T$. The LS estimate of $\mathbf{h}^q$ can then be obtained as

$$\hat{\mathbf{h}}^q = \mathbf{A}^\dagger \mathbf{Y}^q. \tag{7}$$

We assume that the pilot sequences are designed such that the $gK \times LN_t$ matrix $\mathbf{A}$ is of full column rank LN_t, which requires $gK \geq LN_t$. The pseudo-inverse of $\mathbf{A}$ can thus be written as $\mathbf{A}^\dagger = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$ [15, p. 521]. Using (4), we then obtain

$$\hat{\mathbf{h}}^q = \mathbf{h}^q + \mathbf{A}^\dagger \mathbf{T}\mathbf{h}^q + \mathbf{A}^\dagger \boldsymbol{\Xi}^q. \tag{8}$$

To eliminate the interference term due to the data, we impose the following condition:

$$\mathbf{A}^\dagger \mathbf{T} = \mathbf{0}_{LN_t \times LN_t}. \tag{9}$$

We then obtain

$$\hat{\mathbf{h}}^q = \mathbf{h}^q + \mathbf{A}^\dagger \boldsymbol{\Xi}^q. \tag{10}$$

Note that (10) indicates that $\hat{\mathbf{h}}^q$ is a combination of the true channel vector $\mathbf{h}^q$ plus a term affected only by the noise in the system. For zero-mean noise, $\mathcal{E}\{\hat{\mathbf{h}}^q\} = \mathbf{h}^q + \mathbf{A}^\dagger \mathcal{E}\{\boldsymbol{\Xi}^q\} = \mathbf{h}^q$, i.e., $\hat{\mathbf{h}}^q$ forms an unbiased estimate of $\mathbf{h}^q$. Condition (9) holds when $\mathbf{B}^{r^H}_{\text{diag}}(n)\mathbf{S}^s_{\text{diag}}(n) = \mathbf{0}_{K\times K}$, $\forall r, s \in \{1, \ldots, N_t\}$, and $\forall n \in \{0, \ldots, g-1\}$. The only way of satisfying this is by choosing disjoint sets of tones for training and data in each OFDM symbol, i.e., zeros in $\mathbf{B}^r(n)$, where $\mathbf{S}^r(n)$ contains nonzeros, and vice versa. Note that these sets of tones are not necessarily the same for each OFDM symbol. Assuming we use P/g pilot tones per OFDM symbol (not necessarily the same set of P/g pilot tones for each OFDM symbol), we can write (7) and (10) in a simplified form:

$$\hat{\mathbf{h}}^q = \tilde{\mathbf{A}}^\dagger \tilde{\mathbf{Y}}^q = \mathbf{h}^q + \tilde{\mathbf{A}}^\dagger \tilde{\boldsymbol{\Xi}}^q \tag{11}$$

where $\tilde{\mathbf{Y}}^q = [\tilde{\mathbf{Y}}^{q^T}(0), \ldots, \tilde{\mathbf{Y}}^{q^T}(g-1)]^T$, $\tilde{\boldsymbol{\Xi}}^q = [\tilde{\boldsymbol{\Xi}}^{q^T}(0), \ldots, \tilde{\boldsymbol{\Xi}}^{q^T}(g-1)]^T$, and

$$\tilde{\mathbf{A}} = \begin{bmatrix} \tilde{\mathbf{B}}^1_{\text{diag}}(0)\tilde{\mathbf{F}}(0) & \cdots & \tilde{\mathbf{B}}^{N_t}_{\text{diag}}(0)\tilde{\mathbf{F}}(0) \\ \vdots & \cdots & \vdots \\ \tilde{\mathbf{B}}^1_{\text{diag}}(g-1)\tilde{\mathbf{F}}(g-1) & \cdots & \tilde{\mathbf{B}}^{N_t}_{\text{diag}}(g-1)\tilde{\mathbf{F}}(g-1) \end{bmatrix}.$$

In these expressions, $\tilde{\mathbf{B}}^r_{\text{diag}}(n)$ is the $P/g \times P/g$ diagonal matrix containing the nonzero entries (pilot tones) of $\mathbf{B}^r(n)$, and $\tilde{\mathbf{F}}(n)$, $\tilde{\mathbf{Y}}^q(n)$, and $\tilde{\boldsymbol{\Xi}}^q(n)$ are the corresponding P/g rows of $\mathbf{F}$, $\mathbf{Y}^q(n)$, and $\boldsymbol{\Xi}^q(n)$, respectively (the dependence of $\tilde{\mathbf{F}}(n)$ on n comes from the fact that the set of P/g pilot tones are not necessarily the same for each OFDM symbol).

As mentioned earlier, we will design the $gK \times LN_t$ matrix $\mathbf{A}$ to have full column rank LN_t. Following the above design, this is equivalent to the $P \times LN_t$ matrix $\tilde{\mathbf{A}}$ having full column rank LN_t, which requires $P \geq LN_t$. It can easily be checked that the design we will propose later satisfies this full rank condition.

IV. Channel Estimation Analysis

In this section, the MSE of the LS channel estimate is computed. Optimal pilot sequences and optimal placement of the pilot tones w.r.t. this MSE are then derived.

From (11), the MSE of the LS channel estimate is given by

$$\begin{aligned} \text{MSE} &= \frac{1}{LN_t}\mathcal{E}\left\{\left\|\hat{\mathbf{h}}^q - \mathbf{h}^q\right\|^2\right\} \\ &= \frac{1}{LN_t}\mathcal{E}\left\{\left\|\tilde{\mathbf{A}}^\dagger \tilde{\boldsymbol{\Xi}}^q\right\|^2\right\} \\ &= \frac{1}{LN_t}\,\text{tr}\left\{\tilde{\mathbf{A}}^\dagger \mathcal{E}\left\{\tilde{\boldsymbol{\Xi}}^q \tilde{\boldsymbol{\Xi}}^{q^H}\right\}\tilde{\mathbf{A}}^{\dagger^H}\right\}. \end{aligned} \tag{12}$$

For zero-mean white noise, we have $\mathcal{E}\{\tilde{\boldsymbol{\Xi}}^q \tilde{\boldsymbol{\Xi}}^{q^H}\} = \sigma_n^2 \mathbf{I}_P$. Then, the MSE can be written as

$$\text{MSE} = \frac{\sigma_n^2}{LN_t}\,\text{tr}\{(\tilde{\mathbf{A}}^H\tilde{\mathbf{A}})^{-1}\}. \tag{13}$$

Using a similar argument as in [6], we can show that in order to obtain the minimum MSE of the LS channel estimate subject to a fixed power $\mathcal{P}$ dedicated for training, we require $\tilde{\mathbf{A}}^H\tilde{\mathbf{A}} = \mathcal{P}\mathbf{I}_{LN_t}$. The minimum MSE is given by

$$\text{MSE}_{\min} = \frac{\sigma_n^2}{\mathcal{P}}. \tag{14}$$

A. Optimal Training Over One OFDM Symbol

In this subsection, we will derive the optimal pilot sequences and optimal placement of the pilot tones w.r.t. the MSE of the LS channel estimate. For simplicity, we will start with training over one OFDM symbol ($g = 1$) and then extend it to training over multiple OFDM symbols ($g > 1$).

According to Section III, when $g = 1$, training is performed over the time index $n = 0$. To simplify notation, we will omit this time index $n = 0$ in the following. First, let us rewrite $\tilde{\mathbf{A}}^H\tilde{\mathbf{A}}$ as

$$\tilde{\mathbf{A}}^H\tilde{\mathbf{A}} = \begin{bmatrix} \mathbf{C}_{1,1} & \cdots & \mathbf{C}_{1,N_t} \\ \vdots & \cdots & \vdots \\ \mathbf{C}_{N_t,1} & \cdots & \mathbf{C}_{N_t,N_t} \end{bmatrix} \tag{15}$$

where $\mathbf{C}_{r,s}$ is the (r, s)th $L \times L$ sub-matrix of $\tilde{\mathbf{A}}^H\tilde{\mathbf{A}}$, which is given by

$$\mathbf{C}_{r,s} = \tilde{\mathbf{F}}^H \tilde{\mathbf{B}}^{r^H}_{\text{diag}} \tilde{\mathbf{B}}^s_{\text{diag}} \tilde{\mathbf{F}}. \tag{16}$$

As mentioned before, to obtain the minimum MSE of the LS channel estimate subject to a fixed power $\mathcal{P}$ dedicated for training, we require $\tilde{\mathbf{A}}^H\tilde{\mathbf{A}} = \mathcal{P}\mathbf{I}_{LN_t}$, i.e.,

$$\mathbf{C}_{r,s} = \begin{cases} \mathcal{P}\mathbf{I}_L, & \text{if } r = s \\ \mathbf{0}_{L\times L}, & \text{if } r \neq s. \end{cases} \tag{17}$$

Authorized licensed use limited to: IEEE Xplore. Downloaded on October 13, 2008 at 06:55 from IEEE Xplore. Restrictions apply.

Note that with $\{k_0, k_1, \ldots, k_{P-1}\}$ being the set of P pilot tones used for training, $\tilde{\mathbf{F}}$ can be written as $\tilde{\mathbf{F}} = [\mathbf{f}_0, \ldots, \mathbf{f}_{L-1}]$, where $\mathbf{f}_l = [e^{-j2\pi lk_0/K}, e^{-j2\pi lk_1/K}, \ldots, e^{-j2\pi lk_{P-1}/K}]^T$.

First, we will consider the case $r = s$ in (17). Let the power on the pth pilot tone of the rth transmit antenna be $\mathcal{P}_p^r$, such that $\sum_{p=0}^{P-1} \mathcal{P}_p^r = \mathcal{P}$. We then obtain

$$\mathbf{C}_{r,r} = \tilde{\mathbf{F}}^H \text{diag}\{[\mathcal{P}_0^r, \ldots, \mathcal{P}_{P-1}^r]^T\}\tilde{\mathbf{F}}. \tag{18}$$

The (i, j)th entry of the sub-matrix $\mathbf{C}_{r,r}$ can then be written as

$$[\mathbf{C}_{r,r}]_{i,j} = \mathbf{f}_i^H \text{diag}\{[\mathcal{P}_0^r, \ldots, \mathcal{P}_{P-1}^r]^T\}\mathbf{f}_j \tag{19}$$

which is equivalent to

$$[\mathbf{C}_{r,r}]_{i,j} = \begin{cases} \mathcal{P}, & \text{if } i = j \\ \displaystyle\sum_{p=0}^{P-1} \mathcal{P}_p^r e^{-j2\pi k_p(j-i)/K}, & \text{if } i \neq j. \end{cases} \tag{20}$$

To satisfy the first part of (17), we thus require

$$\sum_{p=0}^{P-1} \mathcal{P}_p^r e^{-j2\pi k_p\phi/K} = \mathcal{P}\delta(\phi), \quad \forall\, \phi \in \{-L+1, \ldots, L-1\}.$$

The above condition is satisfied if and only if the following conditions are satisfied.

- C1) $\mathcal{P}_p^r = \mathcal{P}/P$, $\forall p \in \{0, \ldots, P-1\}$ and $\forall r \in \{1, \ldots, N_t\}$.
- C2) $k_p = p_0 + pV$, $\forall \phi \in \{-L+1, \ldots, L-1\} \setminus \{0\}$, where $V \in \mathbb{Z}$ such that $PV\phi/K \in \mathbb{Z}$ and $V\phi/K \notin \mathbb{Z}$, $\forall \phi \in \{-L+1, \ldots, L-1\}$, and $p_0 \in \{0, \ldots, V-1\}$ is some offset.

Note that condition C2) is obtained by using C1) and the power series expansion. C1) means that the pilot tones must be *equipowered*, whereas C2) means that the pilot tones must be *equispaced*, to achieve the first part of (17). For a minimum number of pilot tones or a maximum spacing, we have $PV = K$ or $V = K/P$. For cheap, fast, and simple implementation of the DFT, the total number of subcarriers K is chosen to be a power of 2 in practical systems. Since P should divide K, when we consider a minimum number of pilot tones or a maximum spacing, P should also be a power of 2. Hence, keeping in mind that $P \geq LN_t$, we generally select P as $P = 2^{\lceil \log_2(LN_t) \rceil}$.

We now investigate the conditions imposed by the second part of (17), i.e., $r \neq s$. Let us assume equispaced pilot tones with maximum spacing, that is $k_p = p_0 + pK/P$. The (i, j)th entry of $\mathbf{C}_{r,s}$ can then be written as

$$\begin{aligned}[\mathbf{C}_{r,s}]_{i,j} &= e^{-j2\pi p_0(j-i)/K} \sum_{p=0}^{P-1} \left[\tilde{\mathbf{B}}_{\text{diag}}^{r^H} \tilde{\mathbf{B}}_{\text{diag}}^{s}\right]_{p,p} e^{-j2\pi p(j-i)/P} \\ &= e^{-j2\pi p_0(j-i)/K} \sum_{p=0}^{P-1} \left[\tilde{\mathbf{B}}_{\text{diag}}^{r^H} \mathbf{D}_{j-i} \tilde{\mathbf{B}}_{\text{diag}}^{s}\right]_{p,p} \end{aligned} \tag{21}$$

where $\mathbf{D}_\phi$ represents the $P \times P$ phase shift matrix with phase shift ϕ

$$\mathbf{D}_\phi = \text{diag}\{[1, e^{-j2\pi\phi/P}, \ldots, e^{-j2\pi\phi(P-1)/P}]^T\}.$$

TABLE I
CONSTRAINTS ON OPTIMAL PILOT SEQUENCES FOR VARIOUS SCENARIOS

Configuration	PS Requirement
Single TX	Equipowered + Equispaced [7]
Multiple TX	Equipowered + Equispaced
Flat Fading: $L = 1$	+ Orthogonal
Multiple TX	Equipowered + Equispaced
Frequency-Selective	+ Phase Shift Orthogonal
Fading: $L > 1$	$\phi \in \{-L+1, \ldots, L-1\}$

It is clear from (21) that the second part of (17) is satisfied when

$$\sum_{p=0}^{P-1} \left[\tilde{\mathbf{B}}_{\text{diag}}^{r^H} \mathbf{D}_\phi \tilde{\mathbf{B}}_{\text{diag}}^{s}\right]_{p,p} = 0 \tag{22}$$
$$\forall\, \phi \in \{-L+1, \ldots, L-1\}$$
$$\forall\, r, s, \in \{1, \ldots, N_t\}, \text{ with } r \neq s.$$

When $L = 1$ (flat fading), the pilot sequences on different transmit antennas must be orthogonal. However, when $L > 1$ (frequency-selective fading), the pilot sequences on different transmit antennas must be not only orthogonal but *phase shift orthogonal* for phase shifts in the range $\phi \in \{-L+1, \ldots, L-1\}$.

Note that phase shift orthogonality in the frequency domain corresponds to circular shift orthogonality in the time domain. In other words, the pilot sequence of one antenna must not only be orthogonal to the pilot sequences of other antennas but to circularly shifted copies of these sequences as well.

For the purpose of comparison, we list various scenarios and the constraints they impose on the optimal pilot sequences in Table I.

Optimal pilot sequences can now be designed as

$$\begin{aligned}[\tilde{\mathbf{B}}_{\text{diag}}^{r}]_{p,p} &= \sqrt{\mathcal{P}/P} e^{-j2\pi n_r p/P} \\ &\forall\, p \in \{0, \ldots, P-1\} \\ &\forall\, r \in \{1, \ldots, N_t\} \end{aligned}$$

where the set $\{n_r\}_{r=1}^{N_t}$ has to be selected in a special way. Since

$$[\tilde{\mathbf{B}}_{\text{diag}}^{r^H} \tilde{\mathbf{B}}_{\text{diag}}^{s}]_{p,p} = \mathcal{P}/P e^{-j2\pi(n_s - n_r)p/P}$$

it is clear that in order to satisfy (22), we need $(n_s - n_r + j - i)/P \notin \mathbb{Z}, \forall i, j \in \{0, \ldots, L-1\}$, and $\forall r, s \in \{1, \ldots, N_t\}$ with $r \neq s$. One possible choice is $n_r = (r-1)L$, $\forall r \in \{1, \ldots, N_t\}$. For an arbitrary unit modulus sequence $c(p)$ of length P ($|c(p)| = 1$, $\forall p \in \{0, \ldots, P-1\}$), it is also worth noticing that when $[\tilde{\mathbf{B}}_{\text{diag}}^{r}]_{p,p} = \sqrt{\mathcal{P}/P} e^{-j2\pi n_r p/P}$ is optimal, then $[\tilde{\mathbf{B}}_{\text{diag}}^{r}]_{p,p} = \sqrt{\mathcal{P}/P} e^{-j2\pi n_r p/P} c(p)$ is also optimal.

B. Optimal Training Over Multiple OFDM Symbols

We will now consider training over multiple OFDM symbols ($g > 1$). According to Section III, when $g > 1$, training is

Authorized licensed use limited to: IEEE Xplore. Downloaded on October 13, 2008 at 06:55 from IEEE Xplore. Restrictions apply.

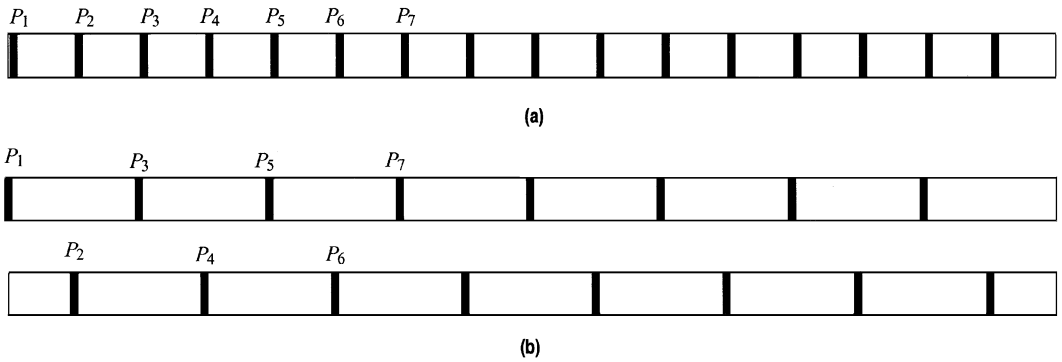


Fig. 2. (a) Training over one OFDM symbol. (b) Training over two OFDM symbols.

performed over the time indices $n \in \{0, \ldots, g-1\}$. First, let us rewrite $\tilde{\mathbf{A}}^H \tilde{\mathbf{A}}$ as

$$\tilde{\mathbf{A}}^H \tilde{\mathbf{A}} = \begin{bmatrix} \mathbf{C}_{1,1} & \cdots & \mathbf{C}_{1,N_t} \\ \vdots & \cdots & \vdots \\ \mathbf{C}_{N_t,1} & \cdots & \mathbf{C}_{N_t,N_t} \end{bmatrix} \tag{23}$$

where

$$\mathbf{C}_{r,s} = \sum_{n=0}^{g-1} \tilde{\mathbf{F}}^H(n) \tilde{\mathbf{B}}^{r^H}_{\text{diag}}(n) \tilde{\mathbf{B}}^{s}_{\text{diag}}(n) \tilde{\mathbf{F}}(n). \tag{24}$$

To obtain the minimum MSE of the LS channel estimate subject to a fixed power $\mathcal{P}$ dedicated for training, we again require (17) to be satisfied. Note that with $\{k_0(n), k_1(n), \ldots, k_{P/g-1}(n)\}$ being the set of P/g pilot tones used for training at time index n, $\tilde{\mathbf{F}}(n)$ can be written as $\tilde{\mathbf{F}}(n) = [\mathbf{f}_0(n), \ldots, \mathbf{f}_{L-1}(n)]$, where $\mathbf{f}_l(n) = [e^{-\jmath 2\pi l k_0(n)/K}, \ldots, e^{-\jmath 2\pi l k_{P/g-1}(n)/K}]^T$.

First, we will consider the case $r = s$ in (17). Let the power on the pth pilot tone of the rth transmit antenna at time index n be $\mathcal{P}_p^r(n)$ such that $\sum_{n=0}^{g-1} \sum_{p=0}^{P/g-1} \mathcal{P}_p^r(n) = \mathcal{P}$. In a similar fashion as before, to satisfy the first part of (17), we require

$$\sum_{n=0}^{g-1} \sum_{p=0}^{P/g-1} \mathcal{P}_p^r(n) e^{-\jmath 2\pi k_p(n)\phi/K} = \mathcal{P}\delta(\phi) \qquad \forall \phi \in \{-L+1, \ldots, L-1\}.$$

Up to an order ambiguity of the pilot tones, i.e., which set of P/g pilot tones is used during which OFDM symbol, the above condition is satisfied, if and only if the following conditions are satisfied.

- C1) $\mathcal{P}_p^r(n) = \mathcal{P}/P$, $\forall n \in \{0, \ldots, g-1\}$, $\forall p \in \{0, \ldots, P/g-1\}$, and $\forall r \in \{1, \ldots, N_t\}$.
- C2) $k_p(n) = p_0 + (n+pg)V$, $\forall n \in \{0, \ldots, g-1\}$, and $\forall p \in \{0, \ldots, P/g-1\}$, where $V \in \mathbb{Z}$, such that $PV\phi/K \in \mathbb{Z}$ and $V\phi/K \notin \mathbb{Z}, \forall \phi \in \{-L+1, \ldots, L-1\} \setminus \{0\}$, and $p_0 \in \{0, \ldots, V-1\}$ is some offset.

Notice the similarity with the conditions stated in Section IV-A. For a minimum number of pilot tones or a maximum spacing, we again have $PV = K$ or $V = K/P$.

We now investigate the conditions imposed by the second part of (17), i.e., $r \neq s$. Let us assume equispaced pilot tones with maximum spacing, that is, $k_p(n) = p_0 + (n+pg)K/P$. The (i, j)th entry of $\mathbf{C}_{r,s}$ can then be written as

$$\begin{aligned} &[\mathbf{C}_{r,s}]_{i,j} \\ &\quad = e^{-\jmath 2\pi p_0 (j-i)/K} \sum_{n=0}^{g-1} \sum_{p=0}^{P/g-1} \left[\tilde{\mathbf{B}}^{r^H}_{\text{diag}}(n) \tilde{\mathbf{B}}^{s}_{\text{diag}}(n)\right]_{p,p} \\ &\qquad \times e^{-\jmath 2\pi (n+pg)(j-i)/P} \\ &\quad = e^{-\jmath 2\pi p_0 (j-i)/K} \\ &\qquad \times \sum_{n=0}^{g-1} \sum_{p=0}^{P/g-1} \left[\tilde{\mathbf{B}}^{(r)H}_{\text{diag}}(n) \mathbf{D}_{j-i}(n) \tilde{\mathbf{B}}^{(s)}_{\text{diag}}(n)\right]_{p,p} \end{aligned} \tag{25}$$

where $\mathbf{D}_\phi(n)$ represents the $P/g \times P/g$ phase shift matrix with phase shift ϕ and offset determined by n: $\mathbf{D}_\phi(n) = \text{diag}\{[e^{-\jmath 2\pi \phi n/P}, e^{-\jmath 2\pi \phi (n+g)/P}, \ldots, e^{-\jmath 2\pi \phi (n+(P/g-1)g)/P}]^T\}$. It is clear from (25) that the second part of (17) is satisfied when

$$\sum_{n=0}^{g-1} \sum_{p=0}^{P/g-1} \left[\tilde{\mathbf{B}}^{r^H}_{\text{diag}}(n) \mathbf{D}_\phi(n) \tilde{\mathbf{B}}^{s}_{\text{diag}}(n)\right]_{p,p} = 0 \tag{26}$$

$$\forall \phi \in \{-L+1, \ldots, L-1\}$$

$$\forall r, s, \in \{1, \ldots, N_t\}, \text{ with } r \neq s.$$

As before, optimal pilot sequences can now be designed as

$$\begin{aligned} [\tilde{\mathbf{B}}^{r}_{\text{diag}}(n)]_{p,p} &= \sqrt{\mathcal{P}/P} e^{-\jmath 2\pi n_r (n+pg)/P} \\ &\forall n \in \{0, \ldots, g-1\}, \forall p \in \{0, \ldots, P/g-1\} \\ &\forall r \in \{1, \ldots, N_t\} \end{aligned}$$

where the set $\{n_r\}_{r=1}^{N_t}$ has to be selected in a special way. Since

$$[\tilde{\mathbf{B}}^{r^H}_{\text{diag}}(n) \tilde{\mathbf{B}}^{s}_{\text{diag}}(n)]_{p,p} = \mathcal{P}/P e^{-\jmath 2\pi (n_s - n_r)(n+pg)/P}$$

it is clear that in order to satisfy (22), we need $(n_s - n_r + j - i)/P \notin \mathbb{Z}, \forall i, j \in \{0, \ldots, L-1\}$, and $\forall r, s \in \{1, \ldots, N_t\}$ with $r \neq s$. As before, one possible choice is $n_r = (r-1)L$, $\forall r \in \{1, \ldots, N_t\}$.

Hence, we can design optimal pilot sequences as in Section IV-A, arbitrarily split each sequence of length P into g subsequences of length P/g, and arbitrarily assign each subsequence to a different OFDM symbol (see, for example, Fig. 2 for training over two consecutive OFDM symbols).

Authorized licensed use limited to: IEEE Xplore. Downloaded on October 13, 2008 at 06:55 from IEEE Xplore. Restrictions apply.

C. Channel Estimation Enhancement

In this subsection, we consider a slowly time-varying channel and describe an RLS algorithm for channel estimation enhancement, where previously received frames of g OFDM symbols can be used to estimate the channel in the current frame. For simplicity, we only consider $g = 1$ in this section. However, the obtained results can easily be generalized to $g > 1$. For convenience, the receive antenna index q is omitted. The channel vector $\hat{\mathbf{h}}$ and the matrix $\tilde{\mathbf{A}}$ will now depend on the time index n. The channel vector $\hat{\mathbf{h}}(n)$ is estimated as

$$\hat{\mathbf{h}}(n) = \underbrace{\begin{bmatrix} \vdots \\ \lambda^k \tilde{\mathbf{A}}(n-k) \\ \vdots \\ \lambda \tilde{\mathbf{A}}(n-1) \\ \tilde{\mathbf{A}}(n) \end{bmatrix}^{\dagger}}_{\tilde{\mathbf{A}}_{\mathrm{RLS}}^{(n)}} \underbrace{\begin{bmatrix} \vdots \\ \lambda^k \tilde{\mathbf{Y}}(n-k) \\ \vdots \\ \lambda \tilde{\mathbf{Y}}(n-1) \\ \tilde{\mathbf{Y}}(n) \end{bmatrix}}_{\tilde{\mathbf{Y}}_{\mathrm{RLS}}^{(n)}} \tag{27}$$

where $\lambda < 1$ is called the forgetting or tracking factor. Using the fact that $\tilde{\mathbf{A}}$ is an orthogonal matrix (optimal pilot sequences derived in the previous sections are used), it can be easily shown that

$$\hat{\mathbf{h}}(n) = \frac{1-\lambda^2}{1-\lambda^{2(n+1)}} \tilde{\mathbf{A}}_{\mathrm{RLS}}^{(n)H} \tilde{\mathbf{Y}}_{\mathrm{RLS}}^{(n)} / \mathcal{P}. \tag{28}$$

At time index $n+1$, we can then write

$$\hat{\mathbf{h}}(n+1) = \frac{1-\lambda^2}{1-\lambda^{2(n+2)}} \times \left(\lambda^2 \tilde{\mathbf{A}}_{\mathrm{RLS}}^{(n)H} \tilde{\mathbf{Y}}_{\mathrm{RLS}}^{(n)} + \tilde{\mathbf{A}}^H(n+1) \tilde{\mathbf{Y}}(n+1) \right) \Big/ \mathcal{P}. \tag{29}$$

Substituting (28) in (29) yields

$$\begin{aligned} \hat{\mathbf{h}}(n+1) &= \lambda^2 \frac{1-\lambda^{2(n+1)}}{1-\lambda^{2(n+2)}} \hat{\mathbf{h}}(n) + \frac{1-\lambda^2}{1-\lambda^{2(n+2)}} \\ &\quad \times \tilde{\mathbf{A}}^H(n+1) \tilde{\mathbf{Y}}(n+1)/\mathcal{P} \qquad (30a) \\ &= \lambda^2 \frac{1-\lambda^{2(n+1)}}{1-\lambda^{2(n+2)}} \hat{\mathbf{h}}(n) + \frac{1-\lambda^2}{1-\lambda^{2(n+2)}} \mathbf{h}(n+1) \\ &\quad + \frac{(1-\lambda^2)}{1-\lambda^{2(n+2)}} \tilde{\mathbf{A}}^H(n+1) \tilde{\boldsymbol{\Xi}}(n+1)/\mathcal{P}. \qquad (30b) \end{aligned}$$

From (30a), it is clear that a low-complexity algorithm for channel estimation can be used, where rather than storing and tracking a large matrix, we can simply update our new channel estimate by only LN_tP^2 multiplications. We define the error of the new channel estimate as

$$\begin{aligned} \boldsymbol{\epsilon}_{n+1} &= \hat{\mathbf{h}}(n+1) - \mathbf{h}(n+1) \\ &= \lambda^2 \frac{1-\lambda^{2(n+1)}}{1-\lambda^{2(n+2)}} \hat{\mathbf{h}}(n) - \lambda^2 \frac{1-\lambda^{2(n+1)}}{1-\lambda^{2(n+2)}} \mathbf{h}(n+1) \\ &\quad + \frac{1-\lambda^2}{1-\lambda^{2(n+2)}} \tilde{\mathbf{A}}^H(n+1) \tilde{\boldsymbol{\Xi}}(n+1)/\mathcal{P}. \end{aligned} \tag{31}$$

For $n \to \infty$, we may assume $(1-\lambda^{2(n+1)})/(1-\lambda^{2(n+2)}) \approx 1$ and $1-\lambda^{2(n+2)} \approx 1$. Defining $\beta = \lambda^2$, (31) then becomes

$$\boldsymbol{\epsilon}_{n+1} = \beta(\hat{\mathbf{h}}(n) - \mathbf{h}(n+1)) + (1-\beta)\tilde{\mathbf{A}}^H(n+1)\tilde{\boldsymbol{\Xi}}(n+1)/\mathcal{P}.$$

Substituting $\boldsymbol{\epsilon}_n = \hat{\mathbf{h}}(n) - \mathbf{h}(n)$, the error of the new channel estimate can be written as

$$\begin{aligned} \boldsymbol{\epsilon}_{n+1} &= \beta \boldsymbol{\epsilon}_n + \beta(\mathbf{h}(n) - \mathbf{h}(n+1)) \\ &\quad + (1-\beta)\tilde{\mathbf{A}}^H(n+1)\tilde{\boldsymbol{\Xi}}(n+1)/\mathcal{P}. \end{aligned} \tag{32}$$

Assuming $\boldsymbol{\epsilon}_n$ is uncorrelated with $\mathbf{h}(n) - \mathbf{h}(n+1)$, the MSE of the new channel estimate can be written as

$$\begin{aligned} \mathcal{E}\{\boldsymbol{\epsilon}_{n+1}^H \boldsymbol{\epsilon}_{n+1}\} &= \mathrm{tr}\{\mathcal{E}\{\boldsymbol{\epsilon}_{n+1}\boldsymbol{\epsilon}_{n+1}^H\}\} \\ &= \beta^2 \mathrm{tr}\{\mathcal{E}\{\boldsymbol{\epsilon}_n \boldsymbol{\epsilon}_n^H\} + 2\mathbf{R}_{hh}(0) - 2\mathbf{R}_{hh}(1)\} \\ &\quad + (1-\beta)^2 \sigma_n^2 L N_t/\mathcal{P} \end{aligned} \tag{33}$$

where $\mathbf{R}_{hh}(k) = \mathcal{E}\left\{\mathbf{h}(n)\mathbf{h}^H(n-k)\right\}$.

For $n \to \infty$, we may assume $\mathcal{E}\{\boldsymbol{\epsilon}_n^H \boldsymbol{\epsilon}_n\} = \mathcal{E}\{\boldsymbol{\epsilon}_{n+1}^H \boldsymbol{\epsilon}_{n+1}\}$, which allows us to rewrite (33) as

$$\begin{aligned} (1-\beta^2)\mathcal{E}\{\boldsymbol{\epsilon}_n^H \boldsymbol{\epsilon}_n\} &= \beta^2 \mathrm{tr}\{2\mathbf{R}_{hh}(0) - 2\mathbf{R}_{hh}(1)\} \\ &\quad + (1-\beta)^2 \sigma_n^2 L N_t/\mathcal{P}. \end{aligned} \tag{34}$$

For i.i.d. channel taps that are correlated in time, we obtain

$$\mathbf{R}_{hh}(k) = r_{hh}(k)\mathbf{I}_{LN_t}.$$

Substituting $\mathbf{R}_{hh}(0)$ and $\mathbf{R}_{hh}(1)$ in (34), the MSE of the channel estimate can be written as

$$\begin{aligned} \mathrm{MSE}(\beta) = \frac{1}{LN_t}\mathcal{E}\{\boldsymbol{\epsilon}_n^H \boldsymbol{\epsilon}_n\} &= \frac{\beta^2}{1-\beta^2}(2r_{hh}(0) - 2r_{hh}(1)) \\ &\quad + \frac{(1-\beta)^2}{1-\beta^2}\sigma_n^2/\mathcal{P}. \end{aligned} \tag{35}$$

From (35), an optimal β can be derived as

$$\begin{aligned} \beta_{opt} &= 1 + \frac{2r_{hh}(0) - 2r_{hh}(1)}{2\sigma_n^2/\mathcal{P}} \\ &\quad - \sqrt{\frac{1}{4}\left(\frac{2r_{hh}(0) - 2r_{hh}(1)}{\sigma_n^2/\mathcal{P}}\right)^2 + \left(\frac{2r_{hh}(0) - 2r_{hh}(1)}{\sigma_n^2/\mathcal{P}}\right)}. \end{aligned} \tag{36}$$

Defining the degree of nonstationarity as in [7]

$$\eta_s = \frac{2r_{hh}(0) - 2r_{hh}(1)}{\sigma_n^2/\mathcal{P}}$$

we can write (36) as

$$\beta_{opt} = 1 + \frac{\eta_s}{2} - \sqrt{\frac{\eta_s^2}{4} + \eta_s}. \tag{37}$$

Authorized licensed use limited to: IEEE Xplore. Downloaded on October 13, 2008 at 06:55 from IEEE Xplore. Restrictions apply.

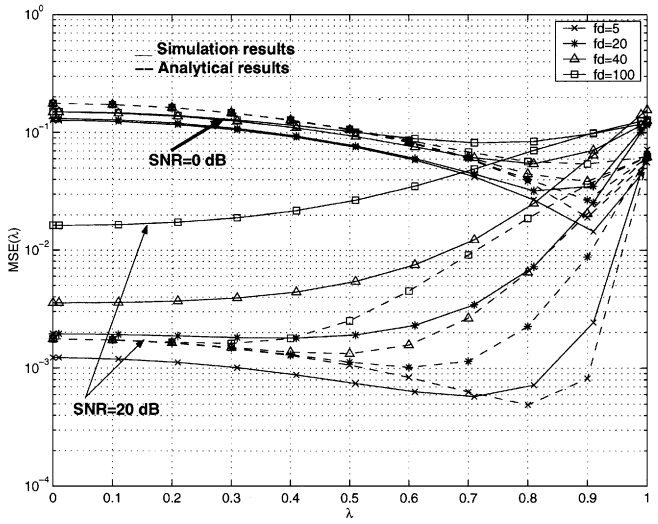


Fig. 3. MSE versus λ; simulation and analytical results for SNR $= 0$ dB and SNR $= 20$ dB.

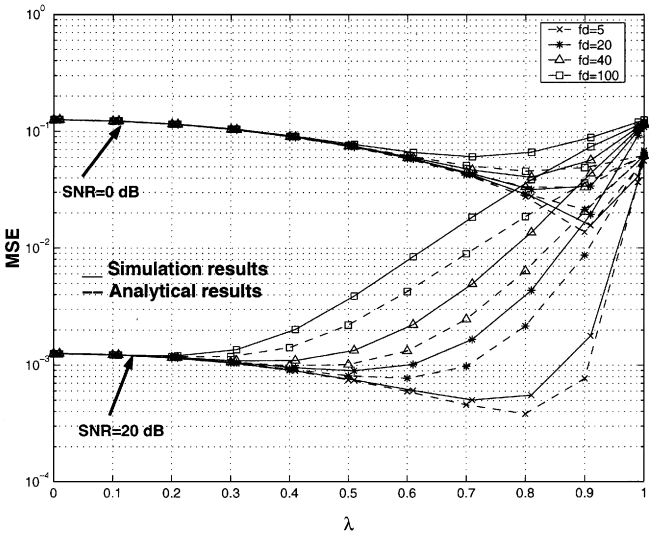


Fig. 4. MSE versus λ; simulation and analytical results for SNR $= 0$ dB and SNR $= 20$ dB (channel is fixed over an entire OFDM symbol).

A similar analysis can be performed in the frequency domain. From (32), the error of the channel estimate can be written as (assume again $n \to \infty$)

$$\Upsilon(z) = \frac{\beta\left(z^{-1}-1\right)}{1-\beta z^{-1}} H(z) + \frac{1-\beta}{1-\beta z^{-1}} \tilde{\mathbf{A}}^H \Xi(z)/\mathcal{P}. \quad (38)$$

Substituting $z = e^{\jmath\omega}$, the MSE of the channel estimate becomes

$$\begin{aligned} \mathrm{MSE}(\beta) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \Upsilon(e^{\jmath\omega})\Upsilon^*(e^{\jmath\omega}) d\omega \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \left(F(e^{\jmath\omega})F^*(e^{\jmath\omega})S_{hh}(f) \right. \\ &\qquad \left. + G(e^{\jmath\omega})G^*(e^{\jmath\omega})\frac{\sigma_n^2}{P} \right) d\omega \end{aligned} \quad (39)$$

where $F(e^{\jmath\omega}) = (\beta(e^{-\jmath\omega}-1))/(1-\beta e^{-\jmath\omega})$, $G(e^{\jmath\omega}) = (1-\beta)/(1-\beta e^{-\jmath\omega})$, and $S_{hh}(f) = \sum_{k=-\infty}^{\infty} r_{hh}(k) e^{-\jmath\omega k}$. From Jakes' model, the correlation function $r_{hh}(k)$ can be written as [13]

$$r_{hh}(k) = \sigma_h^2 J_0(2\pi f_d T_f k) \quad (40)$$

where σ_h^2 is the channel power, $J_0(\cdot)$ is the zeroth-order Bessel function, T_f is the OFDM symbol duration, and f_d is the Doppler spread. In Fig. 3, we compare the analytical results obtained from (39) with simulation results (details about the setup follow in the next section). Notice that there is a difference between the analytical and simulation results. This can be explained as follows. In our analysis so far, we assume that the channel is fixed over an entire OFDM symbol. In reality (and in our simulations), however, the channel varies continuously. To check whether our analytical results are accurate, we therefore compare, in Fig. 4, the analytical results obtained from (39) with simulation results, assuming the channel is fixed over an entire OFDM symbol.

Now, the analytical results are clearly more accurate. The small difference is due to the fact that the analytical results assume $n \to \infty$, whereas the simulation results are obtained by averaging the square error of the channel estimate over the first 100 OFDM symbols. Therefore, the analytical MSE is consistently below the simulated MSE. However, in Fig. 3 as well as in Fig. 4, the optimal β (or λ) is more or less the same for both the analytical and simulation results.

V. Simulations

We assume channels with $L = 8$ taps. These taps are simulated as i.i.d. and correlated in time with a correlation function according to Jakes' model $r_{hh}(\tau) = \sigma_h^2 J_0(2\pi f_d \tau)$. We consider $K = 128$ subcarriers and a cyclic prefix of length $\nu = 8$. The number of pilot tones dedicated for training is $P = 16$, which satisfies the minimum number of training and maximum spacing. Hence, when training is performed over g consecutive OFDM symbols, $P/g = 16/g$ pilot tones are used for training in each OFDM symbol. The OFDM symbol duration is $T_f = 1.13$ ms. QPSK signaling is applied. Finally, 2 transmit and 4 receive antennas are assumed. The performance of the system is measured in terms of the MSE of the channel estimate, and the bit error rate (BER) versus SNR for a zero-forcing equalizer based on the channel estimate. The SNR is defined as SNR $= LN_t\sigma_h^2 E_s/\sigma_n^2$, where E_s is the QPSK symbol power (the power dedicated for training is $\mathcal{P} = (P/(K+L))\mathcal{P}_{tot}$, where $\mathcal{P}_{tot}$ is the total power used to transmit a single OFDM symbol). We run the simulations for different Doppler spreads $f_d = 5, 20, 40$, and 100 Hz.

In our simulations, we evaluate a variety of choices for the pilot sequences:

i) equipowered, equispaced random pilot tones;
ii) equipowered, equispaced, orthogonal pilot tones;
iii) equipowered, equispaced, phase shift orthogonal pilot tones.

As shown in Figs. 5 and 6, using phase shift orthogonal pilot sequences outperforms the use of random or orthogonal pilot sequences in terms of MSE of the channel estimate and BER. We can see a 2-dB gain in SNR for phase shift orthogonal over or-

Authorized licensed use limited to: IEEE Xplore. Downloaded on October 13, 2008 at 06:55 from IEEE Xplore. Restrictions apply.

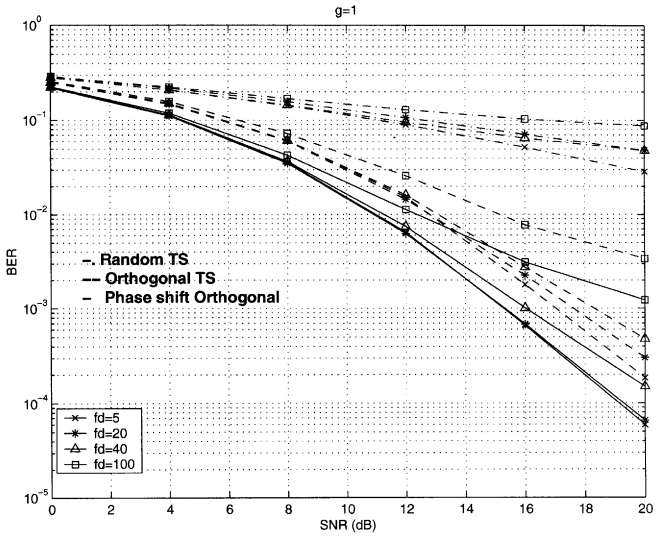


Fig. 5. BER versus SNR for training over one OFDM symbol.

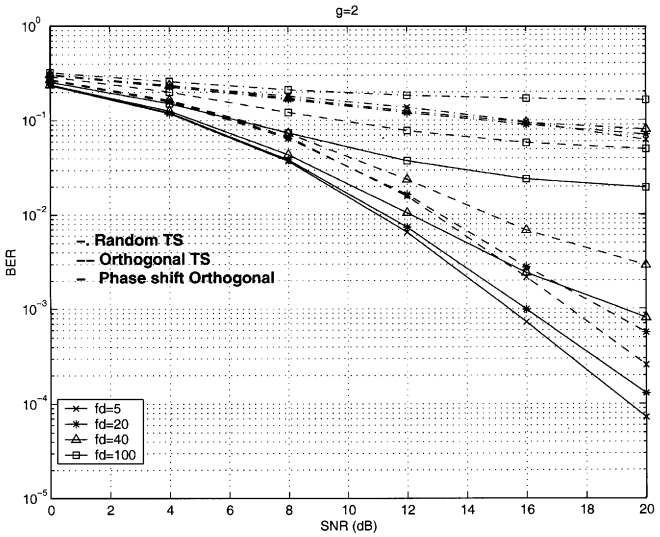


Fig. 7. BER versus SNR for training over two consecutive OFDM symbols.

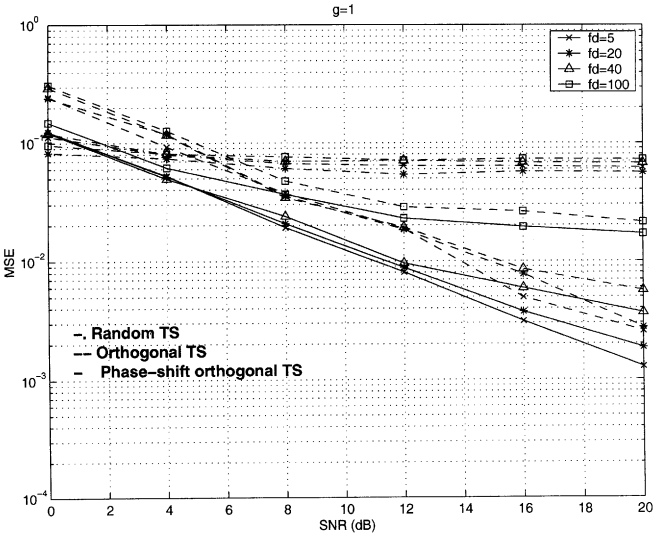


Fig. 6. MSE versus SNR for training over one OFDM symbol.

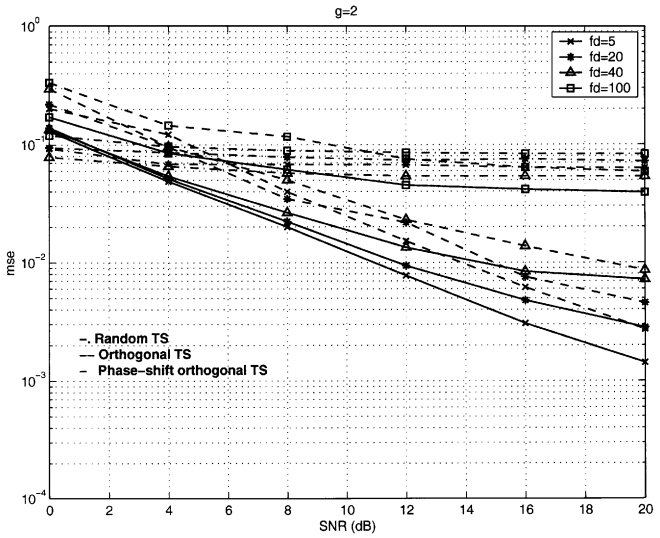


Fig. 8. MSE versus SNR for training over two consecutive OFDM symbol.

thogonal pilot sequences at a BER of 10^{-2} and Doppler spread $f_d = 5$ Hz and a 3.5-dB gain in SNR at a BER of 10^{-2} and Doppler spread $f_d = 100$ Hz. Random pilot sequences are clearly useless. Similar results hold when training over two and four consecutive OFDM symbols is considered (see Figs. 7–10). It is found that training over multiple OFDM symbols pays off especially for slowly time-varying channels. For example, for channels with a Doppler spread $f_d = 5$ Hz, training can be performed over two or four consecutive OFDM symbols without any performance loss, whereas for fast time-varying channels, this scheme will experience an increased BER and becomes even prohibitive for very fast time-varying channels, as shown in Figs. 7–10.

Using the RLS method will enhance the channel estimation especially for channels with a small Doppler spread. As can be seen from Figs. 11 and 12, we can achieve a 2-dB gain in SNR for the scheme with RLS over the scheme without RLS at a BER of 10^{-2} and Doppler spread $f_d = 5$ Hz, whereas no gain is obtained at Doppler spread $f_d = 100$ Hz.

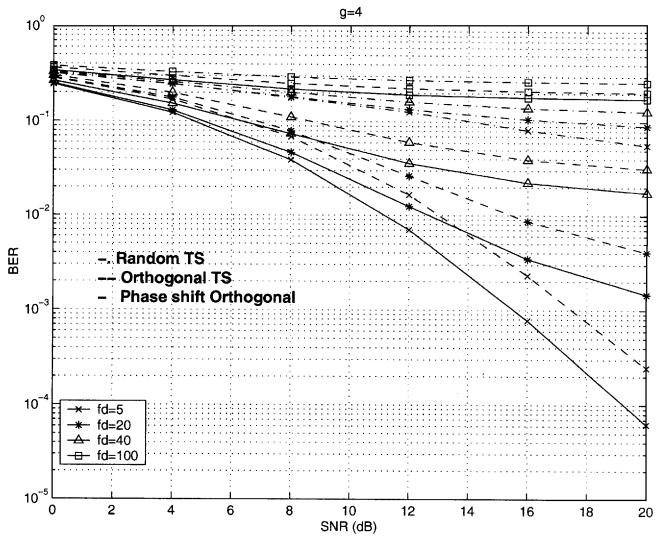


Fig. 9. BER versus SNR for training over four consecutive OFDM symbols.

Authorized licensed use limited to: IEEE Xplore. Downloaded on October 13, 2008 at 06:55 from IEEE Xplore. Restrictions apply.

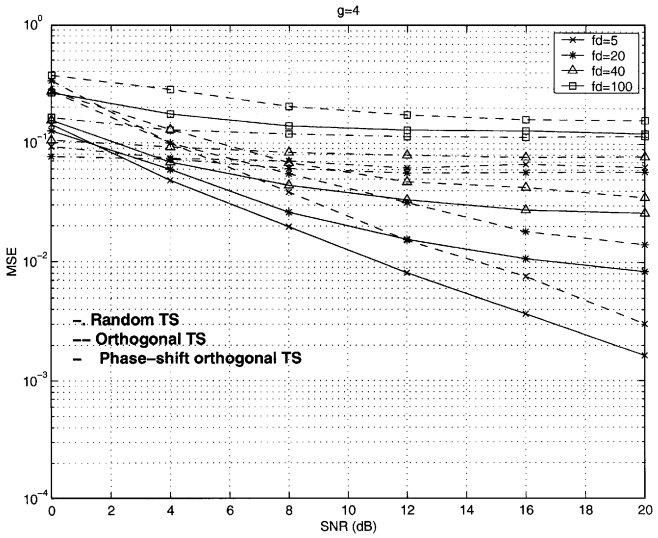


Fig. 10. MSE versus SNR for training over four consecutive OFDM symbol.

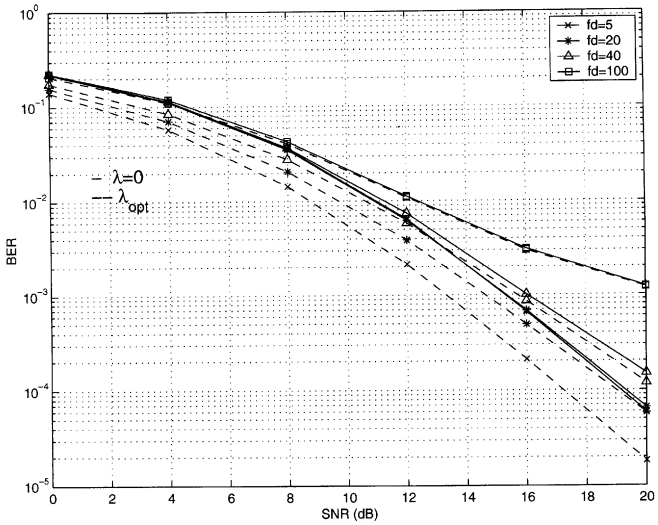


Fig. 11. BER versus SNR with optimal tracking factor λ.

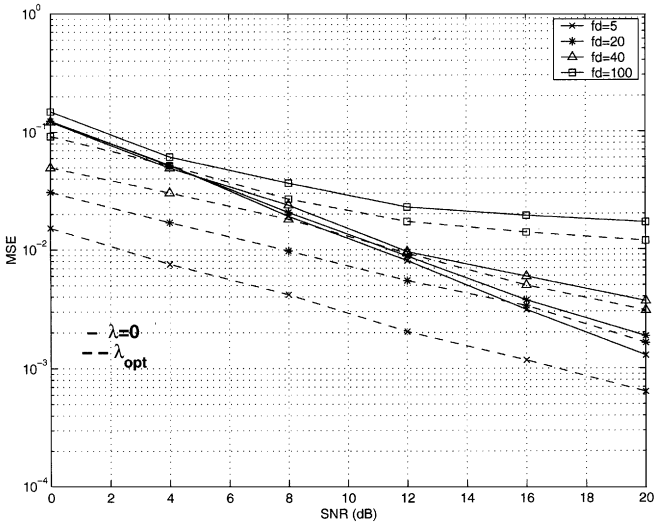


Fig. 12. MSE versus SNR with optimal tracking factor λ.

VI. Conclusions

In this paper, an LS channel estimation scheme for MIMO OFDM systems based on pilot tones has been proposed. To obtain the minimum MSE of the LS channel estimate, the pilot sequences must be equipowered, equispaced, and phase shift orthogonal. Increasing the number of transmit antennas requires more pilot tones for training and, hence, decreases the efficiency. This effect can be mitigated by estimating the channel parameters over multiple OFDM symbols when the channel is slowly time-varying.

References

[1] L. Deneire, P. Vandenameele, L. Van der Perre, B. Gyselinckx, and M. Engels, "A low complexity ML channel estimator for OFDM," in *Proc. IEEE Int. Conf. Commun.*, Helsinki, Finland, June 11–14, 2001.

[2] Y. Li, N. Seshadri, and S. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 461–471, Mar. 1999.

[3] S. Zhou and G. B. Giannakis, "Finite-alphabet based channel estimation for OFDM and related multicarrier systems," *IEEE Trans. Commun.*, vol. 49, pp. 1402–1414, Aug. 2001.

[4] H. Bölcskei, R. W. Heath, Jr., and A. J. Paulraj, "Blind channel identification and equalization in OFDM-based multi-antenna systems," *IEEE Trans. Signal Processing*, vol. 50, pp. 96–109, Jan. 2002.

[5] W. G. Jeon, K. H. Paik, and Y. S. Cho, "An efficient channel estimation technique for OFDM systems with transmitter diversity," in *Proc. IEEE Int. Symp. Pers., Indoor Mobile*, vol. 2, 2000, pp. 1246–1250.

[6] T. L. Tung, K. Yao, and R. E. Hudson, "Channel estimation and adaptive power allocation for performance and capacity improvement of multiple-antenna OFDM systems," in *Proc. Third IEEE Signal Process. Workshop Signal Process. Adv. Wireless Commun.*, Mar. 2001.

[7] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM system," *IEEE Trans. Consum. Electron.*, vol. 44, pp. 1112–1128, Aug. 1998.

[8] H. Bölcskei, D. Gesbert, and A. J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Commun.*, vol. 50, pp. 225–234, Feb. 2002.

[9] O. Edfords, M. Sandell, J.-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.

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

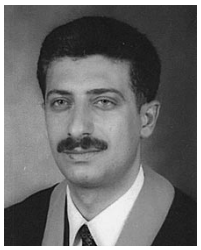
[11] S. Ohno and G. B. Giannakis, "Capacity maximizing pilots for wireless OFDM over rapidly fading channels," in *Proc. Int. Symp. Signals, Syst., Electron.*, Tokyo, Japan, July 24–27, 2001, pp. 246–249.

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

[13] W. Jakes, *Microwave Mobile Communications*. Piscataway, NJ: IEEE, 1974, Classic Reissue.

[14] S. Ohno and G. B. Giannakis, "Optimal training and redundant precoding for block transmissions with applications to wireless OFDM," in *Proc. IEEE ICASSP*, Salt Lake City, UT, May 2001, pp. 2389–2392.

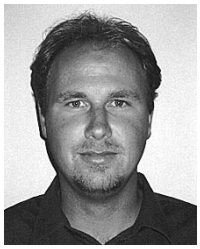
[15] C. D. Cantrell, *Modern Mathematical Methods for Physicists and Engineers*. Cambridge, U.K.: Cambridge Univ. Press, 2000.



Imad Barhumi was born in Palestine in 1972. He received the B.Sc. degree in electrical engineering from Birzeit University, Birzeit, Palestine, in 1996 and the M.Sc. in telecommunications from University of Jordan, Amman, Jordan, in 1999. Currently, he is pursuing the Ph.D. degree in electrical engineering (signal processing) at the Electrical Engineering Department (ESAT), Katholieke Universiteit Leuven, Leuven, Belgium.

From 1999 to 2000, he was with the Electrical Engineering8 Department, Birzeit University, as a lecturer. His research interests are in the area of signal processing for communications.

Authorized licensed use limited to: IEEE Xplore. Downloaded on October 13, 2008 at 06:55 from IEEE Xplore. Restrictions apply.



Geert Leus (M'01) was born in Leuven, Belgium, in 1973. He received the electrical engineering degree and the Ph.D. degree in applied sciences from the Katholieke Universiteit Leuven (KU Leuven) in 1996 and 2000, respectively.

Currently, he is a Postdoctoral Fellow of the Fund for Scientific Research–Flanders (FWO–Vlaanderen) at the Electrical Engineering Department, KU Leuven. During the summer of 1998, he visited Stanford University, Stanford, CA, and from March 2001 until May 2002, he was a Visiting Researcher and Lecturer at the University of Minnesota, Minneapolis. His research interests are in the area of signal processing for communications.

Dr. Leus is a member of the IEEE Signal Processing for Communications Technical Committee. He is also an Associate Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and the IEEE SIGNAL PROCESSING LETTERS.



Marc Moonen (M'94) received the electrical engineering and Ph.D. degrees in applied sciences from the Katholieke Universiteit Leuven (KU Leuven), Leuven, Belgium, in 1986 and 1990 respectively.

He has been a research associate with the Belgian National Fund for Scientific Research since 1994 and an Associate Professor since 2000, both at the Electrical Engineering Department, KU Leuven. His research activities are in digital signal processing, digital communications, and audio signal processing. He is a member of the editorial boards of *Integration, the VLSI Journal* and *Applied Signal Processing*.

Dr. Moonen received the 1994 K.U. Leuven Research Council Award, the 1997 Alcatel Bell (Belgium) Award (with P. Vandaele), and was a 1997 "Laureate of the Belgium Royal Academy of Science." He was chairman of the IEEE Benelux Signal Processing Chapter from 1998 to 2002 and a EURASIP AdCom Member (European Association for Signal, Speech and Image Processing) since 2000. He is an Associate Editor of IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II.

Authorized licensed use limited to: IEEE Xplore. Downloaded on October 13, 2008 at 06:55 from IEEE Xplore. Restrictions apply.