

# ADAPTIVE IDENTIFICATION AND TRACKING OF DOUBLY SELECTIVE FADING CHANNELS FOR WIRELESS MIMO-OFDM SYSTEMS

Dieter Schafhuber, Markus Rupp, Gerald Matz, and Franz Hlawatsch

Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology  
Gusshausstrasse 25/389, 1040 Vienna, Austria (Europe)  
phone: +43 1 58801 38973, fax: +43 1 58801 38999, email: Dieter.Schafhuber@ieee.org  
web: http://www.nt.tuwien.ac.at/dspgroup/time.html

## ABSTRACT

In wireless MIMO-OFDM systems, accurate channel state information at the receiver is a prerequisite for large diversity and multiplexing gains. Here, we propose a pilot symbol assisted, LMS based, adaptive channel identification/tracking scheme for time-varying, delay-spread MIMO channels. Our scheme features automatic tuning of the LMS adaptation constants and modest computational complexity. No prior knowledge of channel statistics is required. Simulation results demonstrate the good performance of our algorithm for a wide range of SNR, maximum delay, and maximum Doppler.

## 1. INTRODUCTION

Multi-input multi-output (MIMO) wireless systems using multi-element transmit and receive antennas can yield improved link reliability through spatial diversity and/or increased data rate through multiplexing techniques [1–3]. MIMO-OFDM is a promising modulation scheme for broadband MIMO communications in the case of delay-spread channels [4–6]. For MIMO-OFDM, several pilot symbol assisted channel estimation schemes have been proposed [7–10] for providing the receiver with the channel state information required for large diversity and multiplexing gains. These schemes have high computational complexity and memory requirements and presuppose (partial) knowledge of the channel statistics. The estimation of these channel statistics is difficult and needs to be repeated periodically since the channel statistics change over time.

In this paper, we present a novel adaptive channel identification and tracking scheme for doubly selective fading MIMO channels. This scheme does not require statistical prior knowledge. It is based on a variant of the least-mean-square (LMS) algorithm and uses pilot sequences with a specific orthogonality property. We also derive the optimum LMS adaptation constants minimizing the mean-square error (MSE) of the channel estimate. Since these optimum adaptation constants depend on the channel statistics, we augment the adaptive scheme by an automatic tuning of the LMS adaptation constants. We thus arrive at a fully adaptive MIMO channel estimation scheme with modest computational complexity and memory requirements. This scheme is suited to arbitrary numbers of transmit and receive antennas, including the MISO case.

The paper is organized as follows. Section 2 describes the MIMO-OFDM system model. In Section 3, we present our LMS-based channel identification/tracking scheme. In Section 4, the optimum LMS adaptation constants are derived. Section 5 presents the augmented channel identification/tracking scheme with automatic tuning of the adaptation constants. Finally, in Section 6 we assess the performance of our method through numerical simulations.

Funding by FWF grant P15156 and European IST project ANTIUM.

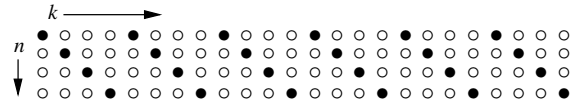


Figure 1: Location of pilot symbols (●) and data symbols (○) for one transmit antenna and  $K = 24$ ,  $P = 6$ , and  $S = 4$ .

## 2. MIMO-OFDM SYSTEM MODEL

**Modulator.** We consider a MIMO-OFDM system [4–6] with  $K$  subcarriers,  $M_T$  transmit antennas, and  $M_R$  receive antennas. The transmitted symbol vectors are  $\mathbf{a}_{n,k} = [a_{n,k}^{(1)} \dots a_{n,k}^{(M_T)}]^T$ ,  $n \in \mathbb{Z}$ ,  $k \in \{0, \dots, K-1\}$ , where  $a_{n,k}^{(i)}$  denotes the symbol transmitted at symbol time  $n$ , subcarrier  $k$ , and antenna  $i$ . The  $n$ th OFDM symbol  $\mathbf{s}_n[m]$  is obtained by applying an inverse discrete Fourier transform (IDFT) to the  $\mathbf{a}_{n,k}$  and prepending a cyclic prefix of length  $L_{cp}$ ,

$$\mathbf{s}_n[m] = \begin{cases} \frac{1}{\sqrt{KM_T}} \sum_{k=0}^{K-1} \mathbf{a}_{n,k} e^{j2\pi mk/K}, & m = -L_{cp}, \dots, K-1 \\ \mathbf{0}, & \text{else.} \end{cases}$$

Thus, the duration of each OFDM symbol is  $N = K + L_{cp}$ . The overall baseband transmit signal is  $\mathbf{s}[m] = \sum_{n=-\infty}^{\infty} \mathbf{s}_n[m - nN]$ .

**Pilot symbols.** At locations  $(n, k) \in \mathcal{P}$ , pilot symbols known to the receiver are transmitted. Here, we use the pilot location set [11]

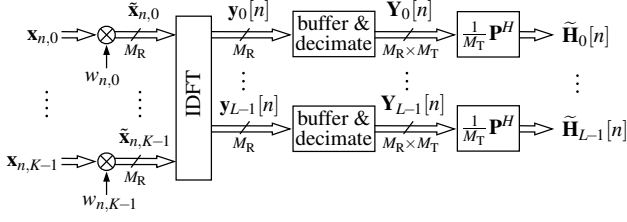
$$\mathcal{P} = \{(n, k) \mid n \in \mathbb{Z}, k = iS + (n \bmod S), i \in [0, P-1]\},$$

where  $P$  is the number of pilots per OFDM symbol and  $S = K/P$  (assumed to be an integer) is the frequency distance of the pilots. An example is shown in Fig. 1. We propose to choose the pilot symbols as  $\mathbf{a}_{n,k} = w_{n,k} \mathbf{p}_n$ , where  $w_{n,k} \in \{1, -1\}$  is a pseudo-noise sequence (for later convenience, we define  $w_{n,k} \triangleq 0$  for  $(n, k) \notin \mathcal{P}$ ) and  $\mathbf{p}_n$  is a spatial pilot spreading vector that is periodic with period  $M_T$ , i.e.,  $\mathbf{p}_{n+M_T} = \mathbf{p}_n$ . Thus, it suffices to specify  $\mathbf{P} \triangleq [\mathbf{p}_1 \dots \mathbf{p}_{M_T}]$ , which we assume to be orthogonal up to a constant:  $\frac{1}{M_T} \mathbf{P} \mathbf{P}^H = \frac{1}{M_T} \mathbf{P}^H \mathbf{P} = \mathbf{I}_{M_T}$ . This implies that the pilot sequences transmitted from different antennas are orthogonal. A convenient choice for  $\mathbf{P}$  is the  $M_T \times M_T$  DFT matrix (e.g., for  $M_T=2$  this gives  $\mathbf{p}_1 = [1 \ 1]^T$  and  $\mathbf{p}_2 = [1 \ -1]^T$ ).

**Channel.** The doubly selective fading MIMO channel is characterized by its  $M_R \times M_T$  baseband impulse response matrix  $\mathbf{H}[m, l]$ . We assume that  $\mathbf{H}[m, l] = \mathbf{0}$  for  $l \geq L$ . To avoid intersymbol interference, we further assume  $L \leq L_{cp} + 1$ . The length- $M_R$  signal vector at the receive antennas is then given by

$$\mathbf{r}[m] = \sum_{l=0}^{L-1} \mathbf{H}[m, l] \mathbf{s}[m-l] + \boldsymbol{\eta}[m],$$

where the noise  $\boldsymbol{\eta}[m]$  is assumed white and stationary with complex normal distribution  $\mathcal{CN}(\mathbf{0}, \sigma_\eta^2 \mathbf{I})$ .



**Figure 2:** Preprocessing stage of the adaptive channel identification/tracking scheme.

For a statistical characterization of the MIMO channel, we neglect spatial correlations and assume *wide-sense stationary uncorrelated scattering* (WSSUS). Let  $h_{j,i}[m,l] \triangleq [\mathbf{H}[m,l]]_{j,i}$  denote the channel between the  $i$ th transmit antenna and the  $j$ th receive antenna. The channel correlation function is then given by [12]

$$E\{h_{j,i}[m,l]h_{j',i'}^*[m-m',l']\} = r_h[m',l]\delta[l-l']\delta[j-j']\delta[i-i']. \quad (1)$$

Note that in addition to being mutually uncorrelated, the channels  $h_{j,i}[m,l]$  have identical time/delay correlation  $r_h[m,l]$ . The channel's scattering function is defined as [12]

$$S_h(l,\xi) \triangleq \sum_{m=-\infty}^{\infty} r_h[m,l]e^{-j2\pi\xi m},$$

with  $\xi$  the normalized Doppler frequency. We denote the channel's path loss by  $\sigma_h^2 \triangleq \sum_{l=0}^{L-1} \int_{-1/2}^{1/2} S_h(l,\xi) d\xi$  and the maximum Doppler frequency by  $\xi_{\max}$ .

**Demodulator.** The received signal  $\mathbf{r}[m]$  is demodulated by discarding the cyclic prefix and computing a normalized DFT,

$$\mathbf{x}_{n,k} = \frac{1}{\sqrt{K}} \sum_{m=0}^{K-1} \mathbf{r}[nN+m] e^{-j2\pi km/K}.$$

If  $\mathbf{H}[m,l]$  varies negligibly within one symbol period, the input-output relation of the overall OFDM system is obtained as

$$\mathbf{x}_{n,k} = \check{\mathbf{H}}_{n,k} \mathbf{a}_{n,k} + \check{\boldsymbol{\eta}}_{n,k},$$

with the time/frequency-domain channel coefficients

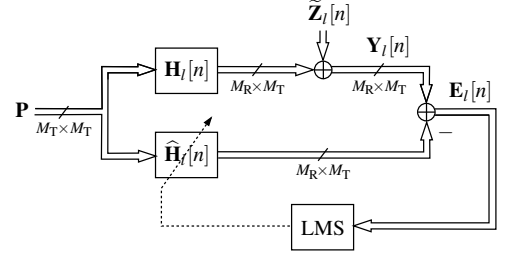
$$\check{\mathbf{H}}_{n,k} \triangleq \frac{1}{\sqrt{M_T}} \sum_{l=0}^{L-1} \mathbf{H}[nN,l] e^{-j2\pi kl/K} \quad (2)$$

and the noise  $\check{\boldsymbol{\eta}}_{n,k} \triangleq \frac{1}{\sqrt{K}} \sum_{m=0}^{K-1} \boldsymbol{\eta}[nN+m] e^{-j2\pi km/K}$ .

### 3. ADAPTIVE CHANNEL IDENTIFICATION/TRACKING

The channel identification/tracking procedure we propose consists of three steps. First, a preprocessing compensates the frequency-dependent factor of the pilot symbols and performs an IDFT to obtain a simple input-output relation in the time/delay domain. Second, a suitably modified LMS algorithm is used to identify and track the (subsampled) channel tap matrices. Finally, a DFT postprocessing yields the estimated time/frequency-domain channel coefficient matrices. In the following, we describe these steps in more detail.

**Preprocessing.** The preprocessing is depicted in Fig. 2. It is motivated by the fact that because of the WSSUS assumption (1), the time/delay-domain channel taps  $\mathbf{H}[m,l]$  are uncorrelated for different delays  $l$ . This enables a *separate* tracking of the individual channel taps without loss of optimality. Furthermore, for each symbol



**Figure 3:** Setup for adaptive identification/tracking of the  $l$ th channel tap.

period of length  $N$ , only  $L$  tap matrices  $\mathbf{H}[m,l]$  need to be estimated instead of  $K \gg L$  time/frequency-domain coefficient matrices  $\check{\mathbf{H}}_{n,k}$ .

We first compensate the  $k$ -dependent pilot symbol factor  $w_{n,k}$ . Recall that  $\mathbf{a}_{n,k} = w_{n,k} \mathbf{P} \mathbf{n}$  with  $w_{n,k} \in \{1, -1\}$  for  $(n,k) \in \mathcal{P}$ ; furthermore  $w_{n,k} = 0$  for  $(n,k) \notin \mathcal{P}$ . Hence,

$$\check{\mathbf{x}}_{n,k} \triangleq w_{n,k} \mathbf{x}_{n,k} = \begin{cases} \check{\mathbf{H}}_{n,k} \mathbf{P} \mathbf{n} + \check{\mathbf{z}}_{n,k}, & (n,k) \in \mathcal{P}, \\ \mathbf{0}, & \text{else,} \end{cases}$$

with  $\check{\mathbf{z}}_{n,k} \triangleq w_{n,k} \check{\boldsymbol{\eta}}_{n,k}$ . Assuming  $P \geq L$  in order to avoid aliasing, we can now transform  $\check{\mathbf{H}}_{n,k}$  into the time/delay domain (cf. (2)):

$$\mathbf{y}_l[n] \triangleq \frac{\sqrt{M_T}}{P} \sum_{k=0}^{K-1} \check{\mathbf{x}}_{n,k} e^{j2\pi \frac{kl}{K}} = \begin{cases} \mathbf{H}[nN,l] \mathbf{P} \mathbf{n} + \tilde{\mathbf{z}}_l[n], & l=0, \dots, L-1, \\ \tilde{\mathbf{z}}_l[n], & l=L, \dots, P-1. \end{cases}$$

Here,  $\tilde{\mathbf{z}}_l[n] \triangleq \frac{\sqrt{M_T}}{P} \sum_{k=0}^{K-1} \check{\mathbf{z}}_{n,k} e^{j2\pi kl/K}$ . Note that the IDFT involves only  $P$  nonzero  $\check{\mathbf{x}}_{n,k}$  per symbol period. Furthermore, delays  $l \geq L$  can be ignored in the subsequent channel tracking. If  $L$  is not exactly known, the worst-case choice  $L = L_{\text{cp}} + 1$  can be used.

We next assume  $M_T N \xi_{\max} \ll 1$ , i.e., the channel varies negligibly within  $M_T$  symbols periods. Thus, by idealization,  $\mathbf{H}[(n+i)N,l] = \mathbf{H}[nN,l]$  for  $i = 1, \dots, M_T - 1$ . For the  $M_R \times M_T$  matrix  $\mathbf{Y}_l[n] \triangleq [\mathbf{y}_l[nM_T] \mathbf{y}_l[nM_T+1] \dots \mathbf{y}_l[(n+1)M_T-1]]$ , we obtain

$$\mathbf{Y}_l[n] = \mathbf{H}_l[n] \mathbf{P} + \tilde{\mathbf{Z}}_l[n], \quad (3)$$

with  $\mathbf{H}_l[n] \triangleq \mathbf{H}[nM_T, l]$  and  $\tilde{\mathbf{Z}}_l[n] \triangleq [\tilde{\mathbf{z}}_l[nM_T] \tilde{\mathbf{z}}_l[nM_T+1] \dots \tilde{\mathbf{z}}_l[(n+1)M_T-1]]$ . Note that (3) involves a decimation by the factor  $M_T$ .

The last preprocessing step is a multiplication of  $\mathbf{Y}_l[n]$  by  $\mathbf{P}^H/M_T$  (correlation with the pilot sequences). Using the orthogonality of the pilot matrix  $\mathbf{P}$ , we obtain the simple ‘‘signal-plus-noise’’ relation

$$\tilde{\mathbf{H}}_l[n] \triangleq \frac{1}{M_T} \mathbf{Y}_l[n] \mathbf{P}^H = \mathbf{H}_l[n] + \mathbf{Z}_l[n], \quad (4)$$

where the elements of  $\mathbf{Z}_l[n] \triangleq \tilde{\mathbf{Z}}_l[n] \mathbf{P}^H$  are i.i.d.  $\mathcal{CN}(0, \sigma_{\check{\boldsymbol{\eta}}}^2/P)$ .

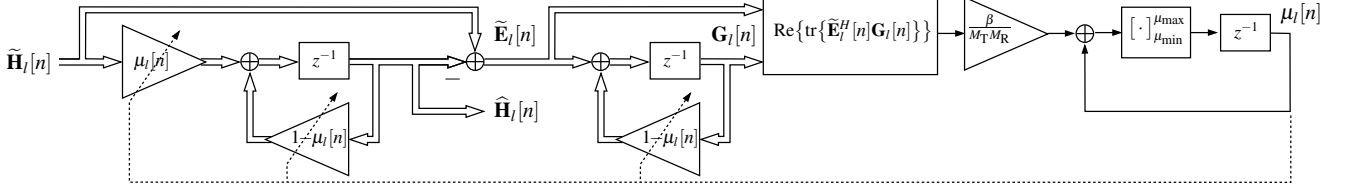
**Channel identification/tracking.** We now present an adaptive channel identification/tracking scheme that is based on (3) and (4). The setup of this scheme is depicted in Fig. 3.

Consider the gradient-algorithm update (cf. [13, Chapter 9])

$$\hat{\mathbf{H}}_l[n+1] = \hat{\mathbf{H}}_l[n] - \frac{\mu_l}{M_T} \nabla_{\hat{\mathbf{H}}} E\{\|\mathbf{E}_l[n]\|_F^2\}, \quad (5)$$

where  $\mu_l$  is an adaptation constant,  $\mathbf{E}_l[n]$  is an error matrix given by (cf. (3) and Fig. 3)

$$\mathbf{E}_l[n] \triangleq \mathbf{Y}_l[n] - \hat{\mathbf{H}}_l[n] \mathbf{P},$$



**Figure 4:** Adaptive identification/tracking of the  $l$ th channel tap with automatic tuning of the adaptation constant  $\mu_l[n]$ .

and  $\|\cdot\|_F^2$  denotes the Frobenius norm. Calculating the gradient  $\nabla_{\hat{\mathbf{H}}} \mathbb{E}\{\|\mathbf{E}_l[n]\|_F^2\}$  in (5), dropping the expectation, and using (4) results in the LMS-type recursion

$$\hat{\mathbf{H}}_l[n+1] = \hat{\mathbf{H}}_l[n] + \mu_l(\tilde{\mathbf{H}}_l[n] - \hat{\mathbf{H}}_l[n]), \quad (6)$$

where  $\tilde{\mathbf{H}}_l[n]$  is obtained according to (4). This recursion starts at some symbol time  $n_0$  and is initialized by  $\hat{\mathbf{H}}_l[n_0] = \tilde{\mathbf{H}}_l[n_0]$ . Because (6) is a first-order difference equation describing an IIR filter, stable operation is guaranteed if  $0 \leq \mu_l < 2$ . The choice of  $\mu_l$  is discussed in Sections 4 and 5.

**Postprocessing.** Finally, the estimated channel tap matrices  $\hat{\mathbf{H}}_l[n]$  are transformed back into the time/frequency domain as (cf. (2))

$$\hat{\mathbf{H}}_{nM_T, k} = \frac{1}{\sqrt{M_T}} \sum_{l=0}^{L-1} \hat{\mathbf{H}}_l[n] e^{-j2\pi kl/K},$$

and the missing channel coefficients are recovered through the trivial interpolation  $\hat{\mathbf{H}}_{nM_T+i, k} = \hat{\mathbf{H}}_{nM_T, k}$ ,  $i = 1, \dots, M_T-1$  (recall that the channel was assumed to vary negligibly during  $M_T$  symbol periods).

#### 4. OPTIMUM ADAPTATION CONSTANTS

We next derive the optimum adaptation constants  $\mu_l$  minimizing the channel identification MSE

$$\varepsilon \triangleq \frac{1}{\sigma_h^2 M_R} \sum_{k=0}^{K-1} \mathbb{E}\{\|\check{\mathbf{H}}_{n, k} - \hat{\mathbf{H}}_{n+1, k}\|_F^2\}.$$

(Note that  $\hat{\mathbf{H}}_{n+1, k}$  is the channel coefficient update at time  $n$ .) The MSE can be rewritten as  $\varepsilon = \frac{1}{\sigma_h^2 M_R} \sum_{l=0}^{L-1} \varepsilon_l$  with

$$\varepsilon_l \triangleq \mathbb{E}\{\|\mathbf{H}_l[n] - \hat{\mathbf{H}}_l[n+1]\|_F^2\}. \quad (7)$$

Thus, the individual adaptation constants  $\mu_l$  can be optimized separately as  $\mu_{l, \text{opt}} = \arg \min_{\mu_l} \varepsilon_l$ . To further develop the component MSEs  $\varepsilon_l$ , we note that (6) can be expressed as (cf. [14])

$$\hat{\mathbf{H}}_l[n+1] = \tilde{\mathbf{H}}_l[n] * \phi_l[n], \quad (8)$$

where  $\phi_l[n] \triangleq \mu_l(1-\mu_l)^n u[n]$  (with  $u[n]$  the unit step) is the impulse response of the tracking filter and  $*$  denotes convolution. Inserting (8) into (7), using (4), and assuming  $n_0 = -\infty$  yields

$$\varepsilon_l = \mathbb{E}\{\|\mathbf{H}_l[n] * \psi_l[n]\|_F^2\} + \mathbb{E}\{\|\mathbf{Z}_l[n] * \phi_l[n]\|_F^2\},$$

where  $\psi_l[n] \triangleq \delta[n] - \phi_l[n]$  is the impulse response of the tracking error filter. Using the fact that the scattering function of  $\mathbf{H}_l[n] = \mathbf{H}_l[nM_T N, l]$  equals  $\frac{1}{M_T N} S_h(l, \frac{\xi}{M_T N})$  (note that aliasing is avoided since  $M_T N \xi_{\text{max}} \ll 1$ ), this can be written as

$$\varepsilon_l = \frac{1}{N} \int_{-1/2}^{1/2} S_h\left(l, \frac{\xi}{M_T N}\right) |\Psi_l(\xi)|^2 d\xi + \frac{M_T \sigma_\eta^2}{P} \sum_{n=0}^{\infty} |\phi_l[n]|^2, \quad (9)$$

where  $\Psi_l(\xi)$  is the Fourier transform of  $\psi_l[n]$ . We have  $\sum_{n=0}^{\infty} |\phi_l[n]|^2 = \frac{\mu_l}{2-\mu_l}$  and

$$|\Psi_l(\xi)|^2 = \frac{2(1-\mu_l)^2 [1 - \cos(2\pi\xi)]}{1 - 2(1-\mu_l)\cos(2\pi\xi) + (1-\mu_l)^2}.$$

Since typically  $\xi_{\text{max}}$  is small, we use  $\cos(2\pi\xi) \approx 1 - 2\pi^2 \xi^2$  in the numerator and  $\cos(2\pi\xi) \approx 1$  in the denominator [14], whence  $|\Psi_l(\xi)|^2 \approx 4\pi^2 \left(\frac{1-\mu_l}{\mu_l}\right)^2 \xi^2$ . Thus, (9) simplifies to

$$\varepsilon_l \approx \frac{M_T \sigma_\eta^2}{P} \left[ \alpha_l \left(\frac{1-\mu_l}{\mu_l}\right)^2 + \frac{\mu_l}{2-\mu_l} \right], \quad \text{with } \alpha_l \triangleq (2\pi \bar{\xi}_l)^2 P \frac{\sigma_h^2}{\sigma_\eta^2},$$

where  $\bar{\xi}_l \triangleq \left[\frac{1}{\sigma_h^2} \int_{-1/2}^{1/2} \xi^2 S_h(l, \xi) d\xi\right]^{1/2}$  is the Doppler spread for the  $l$ th tap. The optimum adaptation constant  $\mu_{l, \text{opt}}$  minimizing the above (approximate) expression of  $\varepsilon_l$  satisfies a cubic equation whose unique positive solution can be determined using Cardano's formula. From this solution, one can obtain the following simple approximation to  $\mu_{l, \text{opt}}$ , denoted  $\tilde{\mu}_{l, \text{opt}}$ , that guarantees  $0 \leq \tilde{\mu}_{l, \text{opt}} \leq 1$  and thus stability of the LMS algorithm:

$$\tilde{\mu}_{l, \text{opt}} = \min\left\{1, \frac{3}{2} \sqrt[3]{\alpha_l}\right\}. \quad (10)$$

We note that  $\tilde{\mu}_{l, \text{opt}}$  increases for increasing Doppler spread  $\bar{\xi}_l$ , increasing number of pilot symbols  $P$ , and increasing SNR  $\sigma_h^2/\sigma_\eta^2$ . If  $\mu_l = \tilde{\mu}_{l, \text{opt}} \approx 0$ , (8) effectively acts as an integrator, yielding slow channel tracking but good noise suppression. Conversely, if  $\mu_l = \tilde{\mu}_{l, \text{opt}} \approx 1$ , (8) simplifies to  $\hat{\mathbf{H}}_l[n+1] \approx \tilde{\mathbf{H}}_l[n]$ , which results in fast channel tracking but poor noise suppression.

The Doppler spread  $\bar{\xi}_l$  and the SNR  $\sigma_h^2/\sigma_\eta^2$  required for determining  $\tilde{\mu}_{l, \text{opt}}$  change with time and are not easily available in practice. However, if the  $\mu_l$  are poorly matched to the current channel conditions, the MSE  $\varepsilon$  increases significantly. Therefore, we next propose an automatic, adaptive tuning of the adaptation constants.

#### 5. CHANNEL TRACKING WITH AUTOMATIC TUNING OF THE ADAPTATION CONSTANTS

Fig. 4 depicts the augmented identification/tracking scheme incorporating automatic tuning of the adaptation constant for the  $l$ th channel tap. This scheme connects directly to the preprocessing shown in Fig. 2. It does not require any prior knowledge.

**Automatic tuning algorithm.** For automatic tuning,  $\mu_l$  in (6) is replaced with a time-varying adaptation constant  $\mu_l[n]$  that is recursively updated according to (cf. [13, Chapter 16])

$$\mu_l[n+1] = \left[ \mu_l[n] - \frac{\beta}{2M_T M_R} \nabla_{\mu_l} \mathbb{E}\{\|\tilde{\mathbf{E}}_l[n]\|_F^2\} \right]_{\mu_{\text{min}}}^{\mu_{\text{max}}}, \quad n \geq n_0. \quad (11)$$

Here, the error matrix  $\tilde{\mathbf{E}}_l[n]$  is given by

$$\tilde{\mathbf{E}}_l[n] \triangleq \frac{1}{M_T} \mathbf{E}_l[n] \mathbf{P}^H = \tilde{\mathbf{H}}_l[n] - \hat{\mathbf{H}}_l[n],$$

$\beta$  is a (hyper)adaptation constant, and  $[\cdot]_{\mu_{\min}}^{\mu_{\max}}$  denotes clipping to the interval  $[\mu_{\min}, \mu_{\max}]$ . Note that (11) uses the same  $\beta$  and  $\mu_0$  for all  $l = 0, \dots, L-1$ . Computing the gradient in (11) and dropping the expectation results in the recursion

$$\mu_l[n+1] = \left[ \mu_l[n] + \frac{\beta}{M_T M_R} \operatorname{Re}\{\operatorname{tr}\{\tilde{\mathbf{E}}_l^H[n] \mathbf{G}_l[n]\}\} \right]_{\mu_{\min}}^{\mu_{\max}}, \quad (12)$$

where  $\operatorname{tr}\{\cdot\}$  denotes matrix trace and  $\mathbf{G}_l[n] \triangleq \nabla_{\mu} \hat{\mathbf{H}}_l[n]$ . Using (6), one can show that  $\mathbf{G}_l[n]$  can itself be computed recursively as

$$\mathbf{G}_l[n+1] = \nabla_{\mu} \hat{\mathbf{H}}_l[n+1] = (1 - \mu_l[n]) \mathbf{G}_l[n] + \tilde{\mathbf{E}}_l[n].$$

The recursion is initialized as  $\mu_l[n_0] = \mu_0$  and  $\mathbf{G}_l[n_0] = \mathbf{0}$ .

The choice of  $\mu_{\min}$ ,  $\mu_{\max}$ ,  $\beta$ , and (during startup)  $\mu_0$  influences the tracking performance but is not very critical. Stability of the adaptive filter is ensured if  $0 \leq \mu_{\min} \leq \mu_{\max} < 2$ . We obtained good results with  $\mu_{\min} = 0.01$ ,  $\mu_{\max} = 1$ ,  $\beta = 0.5$ , and  $\mu_0 = 0.1$ .

If the channel statistics do not change too fast,  $\mu_l[n]$  will be close to the current value of  $\mu_{l,\text{opt}}$ . Thus,  $\mu_l[n]$  provides information about the current rate of channel variation. This is useful, e.g., for adaptive modulation techniques.

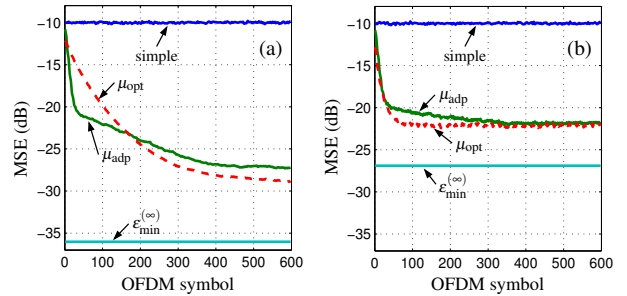
**Complexity.** The various stages of the adaptive channel estimation scheme have the following computational complexity (number of operations per OFDM symbol). IDFT preprocessing (see Fig. 2):  $\mathcal{O}(M_R P \log_2 P)$ ; pilot symbol correlation:  $\mathcal{O}(L M_T M_R)$ ; augmented channel tracking (shown in Fig. 4 for a single tap):  $\mathcal{O}(L M_R)$ ; DFT postprocessing:  $\mathcal{O}(M_R K \log_2 K)$ . For typical system parameters, the complexity is dominated by the preprocessing. The number of memory locations required is roughly  $2 M_T M_R (L + K)$ . Thus, our channel tracking scheme is very efficient in terms of computation and memory requirements.

## 6. SIMULATION RESULTS

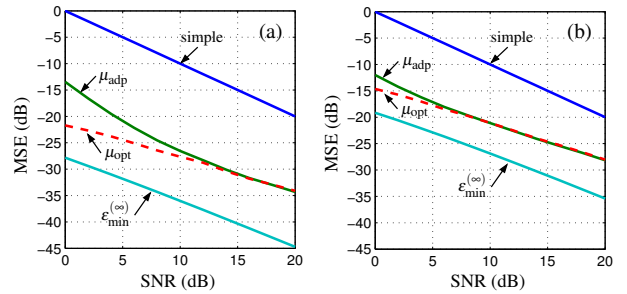
We simulated a MIMO-OFDM system with parameters  $M_T = M_R = 2$ ,  $K = 64$ ,  $P = 16$ , and  $L_{\text{cp}} = 15$ . Assuming a bandwidth/sampling rate of  $B = 6.4$  MHz, the subcarrier separation is 100 kHz and the cyclic prefix duration is  $2.5 \mu\text{s}$ . For a QPSK symbol alphabet, the spectral efficiency of this MIMO-OFDM system is 2.4 bit/s/Hz.

The fading MIMO channel has uncorrelated components  $h_{j,i}[m,l]$  satisfying the WSSUS assumption. The component channels were simulated according to [15] with exponential delay profile and Jakes Doppler profile, i.e.,  $S_h(l, \xi) = \exp(-l/\kappa) / \sqrt{\xi_{\max}^2 - \xi^2}$  for  $l = 0, \dots, L-1$ ,  $|\xi| < \xi_{\max}$  and  $S_h(l, \xi) = 0$  else (here,  $\kappa = L / \log_e(2L)$ ). Unless stated otherwise, we used  $\tau_{\max} \triangleq L/B = 1.875 \mu\text{s}$  (i.e.,  $L = 12$ ) and either  $v_{\max} \triangleq \xi_{\max} B = 10$  Hz (“slow channel”) or  $v_{\max} = 100$  Hz (“fast channel”).

Because  $L$  is assumed unknown to the receiver, we used  $L_{\text{cp}} + 1 = 16$  individual adaptive LMS channel estimators. In the figures, “ $\mu_{\text{opt}}$ ” and “ $\mu_{\text{adp}}$ ” will label the results obtained with the optimum adaptation constant  $\mu_{l,\text{opt}}$  and with the automatically tuned adaptation constant (using  $\mu_{\min} = 0.01$ ,  $\mu_{\max} = 1$ ,  $\mu_0 = 0.1$ , and  $\beta = 0.5$ ), respectively. For comparison, we also show the results obtained with a “simple” channel estimator that consists only of the pre- and postprocessing, i.e.,  $\hat{\mathbf{H}}_l[n] = \tilde{\mathbf{H}}_l[n]$ ,  $l = 0, \dots, L_{\text{cp}}$ . As a benchmark, the figures also include the MMSE  $\epsilon_{\min}^{(\infty)}$  of the causal, infinite-length MMSE channel estimator operating on the  $\tilde{\mathbf{H}}_l[n]$  [16].



**Figure 5:** Convergence behavior (MSE vs. symbol time) for (a) the slow channel and (b) the fast channel.



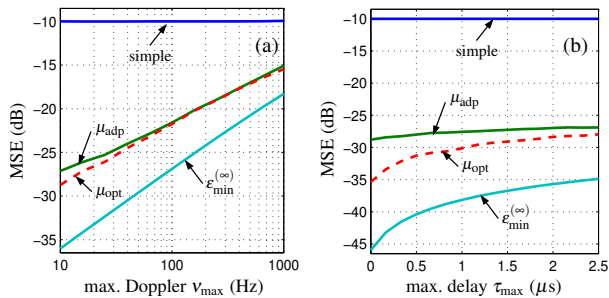
**Figure 6:** MSE vs. SNR for (a) the slow channel and (b) the fast channel.

**Convergence.** In Fig. 5, we show the convergence behavior of our adaptive algorithms at an SNR of 10 dB. The MSE results were obtained by averaging over 100 simulation runs. Initially, all methods achieve the same channel estimation MSE of about  $-10$  dB. Within some 30 OFDM symbols, our fully adaptive method decreases the MSE by another 10 dB for both the slow and the fast channel. After final convergence, the excess MSE over  $\epsilon_{\min}^{(\infty)}$  is only about 7 dB for the slow channel and 5 dB for the fast channel. It is also seen that the  $\mu_{\text{adp}}$  and  $\mu_{\text{opt}}$  results are quite similar.

**MSE vs. SNR.** In Fig. 6, we show the MSE (averaged over 50,000 OFDM symbols after convergence of the adaptive filters) vs. the SNR. It is seen that  $\mu_{\text{adp}}$  and  $\mu_{\text{opt}}$  perform identically except for low SNR. However, even at an SNR of 0 dB, the MSE achieved with  $\mu_{\text{adp}}$  is less than about  $-12$  dB for both channels. Furthermore,  $\mu_{\text{adp}}$  gains 8–15 dB over the simple estimator while its excess MSE over  $\epsilon_{\min}^{(\infty)}$  is about 12 dB for the slow channel and 7 dB for the fast channel (only weakly dependent on the SNR).

**MSE vs. maximum delay and Doppler.** Fig. 7(a) shows the MSE after convergence vs. the maximum Doppler frequency  $v_{\max}$ , at fixed  $\tau_{\max} = 1.875 \mu\text{s}$  and an SNR of 10 dB. It is seen that the MSE of all methods except for the simple estimator increases with  $v_{\max}$ ; this is because larger Doppler frequencies result in smaller channel coherence times. However, the excess MSE of  $\mu_{\text{adp}}$  and  $\mu_{\text{opt}}$  over  $\epsilon_{\min}^{(\infty)}$  decreases for increasing  $v_{\max}$ . It is furthermore seen that our method yields reliable channel estimates even at the extremely large Doppler frequency of 1.000 Hz.

Fig. 7(b) shows the MSE vs. the maximum delay  $\tau_{\max}$ , at fixed  $v_{\max} = 10$  Hz and an SNR of 10 dB. It can be seen that the MSE of  $\mu_{\text{adp}}$  increases only very slowly with  $\tau_{\max}$ . This is because the clipping in (12) bounds  $\mu_l[n]$  away from zero, which results in residual noise for taps that would ideally be zero. The MSE of  $\mu_{\text{adp}}$  is lower than  $-27$  dB for all  $\tau_{\max}$  up to the cyclic prefix duration of  $2.5 \mu\text{s}$ .



**Figure 7:** MSE vs. (a) maximum Doppler shift and (b) maximum delay.

The largest loss of  $\mu_{\text{adp}}$  relative to  $\mu_{\text{opt}}$  and  $\epsilon_{\text{min}}^{(\infty)}$  occurs at  $\tau_{\text{max}} = 0$  and equals 7 dB and 17 dB, respectively.

**BER vs. SNR.** To evaluate the bit error rates (BER) achieved with our  $2 \times 2$  MIMO-OFDM system, we performed a simple space-frequency orthogonal block precoding of the  $(K-P)M_T = 96$  data symbols associated to each OFDM symbol. An  $8 \times 8$  (normalized) DFT matrix was applied to 12 groups of 8 symbols each. To exploit frequency (delay) diversity, each group of symbols was then mapped to the two transmit antennas using four maximally separated subcarriers. At the receiver, ML decoding of each symbol group was performed using a sphere decoder [17] that was provided with the respective channel estimates. An ideal ML receiver using the true channel coefficients (labeled “ideal”) serves as a benchmark.

For comparison, we also simulated a SISO OFDM system ( $M_T = M_R = 1$ ) with 16-QAM symbols but otherwise identical system parameters. This system has the same spectral efficiency of 2.4 bit/s/Hz as the MIMO system. The transmitter here distributes groups of 4 precoded data symbols across the subcarriers. Thus, the frequency diversity is the same as for the MIMO system but there is no spatial diversity. For the SISO system, we only implemented the ideal ML receiver using the true channel coefficients.

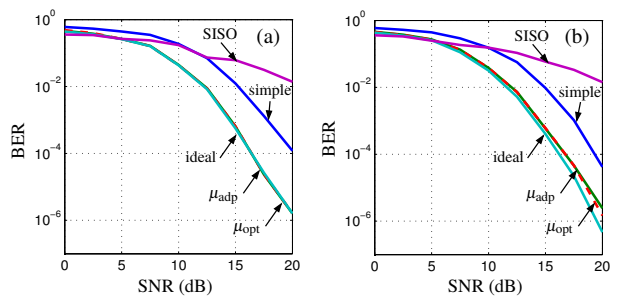
Fig. 8 shows that the performance of the ML receivers with adaptive channel tracking is virtually identical to that of the ideal ML receiver. Moreover, the receivers with adaptive channel tracking clearly outperform the receiver incorporating the simple channel estimation scheme. At a BER of  $10^{-2}$ , for example, the proposed scheme achieves an SNR gain of about 3.5 dB over the simple channel estimator. Finally, the diversity advantage of the MIMO systems relative to the SISO system is also clearly visible.

## 7. CONCLUSIONS

We proposed a pilot symbol assisted, LMS based adaptive channel identification and tracking scheme for MIMO-OFDM systems operating over doubly selective fading channels. Our algorithm features automatic tuning of the LMS adaptation constant, does not require prior knowledge of channel statistics, and has modest computational complexity and memory requirements. Simulation results for a  $2 \times 2$  MIMO-OFDM system indicate that reliable channel estimates can be obtained for all practically relevant ranges of SNR, delay spread, and Doppler spread. The BER achieved with an ML receiver using our channel identification/tracking scheme is almost identical to the BER of an ideal ML receiver using the true channel coefficients.

## ACKNOWLEDGMENT

The authors would like to thank H. Artés and D. Seethaler for helpful discussions and for providing parts of the simulation software.



**Figure 8:** BER vs. SNR for (a) the slow channel and (b) the fast channel.

## REFERENCES

- [1] G. J. Foschini and M. J. Gans, “On limits of wireless communications in a fading environment when using multiple antennas,” *Wireless Personal Communications*, vol. 6, pp. 311–335, 1998.
- [2] I. E. Telatar, “Capacity of multi-antenna gaussian channels,” Tech. Rep. BL0112170-950615-07TM, AT&T Bell Laboratories, June 1995.
- [3] V. Tarokh, N. Seshadri, and A. R. Calderbank, “Space-time codes for high data rate wireless communications: Performance criterion and code construction,” *IEEE Trans. Inf. Theory*, vol. 44, pp. 744–765, March 1998.
- [4] D. Agrawal, V. Tarokh, A. Naguib, and N. Seshadri, “Space-time coded OFDM for high data rate wireless communications over wide-band channels,” in *Proc. IEEE VTC-98*, (Ottawa, Canada), pp. 2232–2236, May 1998.
- [5] H. Bölcskei, D. Gesbert, and A. J. Paulraj, “On the capacity of OFDM-based spatial multiplexing systems,” *IEEE Trans. Comm.*, vol. 50, pp. 225–234, Feb. 2002.
- [6] H. Sampath, S. Talwar, J. Tellado, V. Erceg, and A. Paulraj, “A fourth-generation MIMO-OFDM broadband wireless system: Design, performance, and field trial results,” *IEEE Comm. Mag.*, vol. 40, pp. 143–149, Sept. 2002.
- [7] Y. Li, N. Seshadri, and S. Ariyavisitakul, “Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels,” *IEEE J. Sel. Areas Comm.*, vol. 17, pp. 461–471, March 1999.
- [8] Y. Li, J. H. Winters, and N. R. Sollenberger, “MIMO-OFDM for wireless communications: Signal detection with enhanced channel estimation,” *IEEE Trans. Comm.*, vol. 50, pp. 1471–1477, Sept. 2002.
- [9] Y. Li, “Simplified channel estimation for OFDM systems with multiple transmit antennas,” *IEEE Trans. Wireless Comm.*, vol. 1, pp. 67–75, Jan. 2002.
- [10] I. Barhumi, G. Leus, and M. Moonen, “Optimal training sequences for channel estimation in MIMO OFDM systems in mobile wireless channels,” in *Int. Zurich Seminar on Broadband Communications*, (Zurich, Switzerland), pp. 441–446, Feb. 2002.
- [11] R. Negi and J. Cioffi, “Pilot tone selection for channel estimation in a mobile OFDM system,” *IEEE Trans. Consumer Electron.*, vol. 44, pp. 1122–1128, Aug. 1998.
- [12] P. A. Bello, “Characterization of randomly time-variant linear channels,” *IEEE Trans. Comm. Syst.*, vol. 11, pp. 360–393, 1963.
- [13] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs (NJ): Prentice Hall, 3rd ed., 1996.
- [14] J. Lin, J. G. Proakis, F. Ling, and H. Lev-Ari, “Optimal tracking of time-varying channels: A frequency domain approach for known and new algorithms,” *IEEE J. Sel. Areas Comm.*, vol. 13, pp. 141–154, Jan. 1995.
- [15] D. Schafhuber, G. Matz, and F. Hlawatsch, “Simulation of wideband mobile radio channels using subsampled ARMA models and multi-stage interpolation,” in *Proc. 11th IEEE Workshop on Statistical Signal Processing*, (Singapore), pp. 571–574, Aug. 2001.
- [16] D. Schafhuber, G. Matz, and F. Hlawatsch, “Adaptive Wiener filters for time-varying channel estimation in wireless OFDM systems,” in *Proc. IEEE ICASSP-2003*, (Hong Kong), April 2003.
- [17] U. Fincke and M. Post, “Improved methods for calculating vectors of short length in a lattice, including a complexity analysis,” *Math. of Comp.*, vol. 44, pp. 463–471, April 1985.