Estimation of FWA MIMO Channels

Pei Xiao, Rolando Carrasco
School of EECE, University of Newcastle
Newcastle, NE1 7RU, United Kingdom
E-mail: pei.xiao@ncl.ac.uk,
r.carrasco@ncl.ac.uk

Ian Wassell
Computer Laboratory
University of Cambridge
Cambridge, CB3 0FD, United Kingdom
E-mail: i.jw24@eng.cam.ac.uk

Abstract — Channel estimation schemes for fixed wireless access (FWA) multiple-input, multiple-output (MIMO) systems are considered in this study. The use of multiple antennas in combination with advanced detection techniques, such as turbo equalization is an effective means for a FWA system to provide high quality and high data rate services. Accurate knowledge, i.e., a good estimate of the underlying channel is essential for turbo equalization to achieve good performance. In this paper, we investigate some algorithms that are suitable for estimating FWA MIMO channels. The proposed schemes are evaluated and compared using different training sequences. Based on our analysis and numerical results, some recommendations are made on how to design appropriate channel estimator and how to choose training sequence for practical FWA systems.

I. INTRODUCTION

Fixed wireless access (FWA) is quickly emerging as a significant network access alternative for the delivery of voice, data, Internet, video and multimedia type applications to business and residential customers. FWA systems offer a very cost-effective way of building an access network. Easy maintainability, low incremental costs and portability are key benefits of this wireless alternative [1]. Standardization of FWA systems is currently being undertaken by the IEEE 802.16 group [2] and the ETSI HIPERMAN group [3]. Both orthogonal frequency division multiplexing (OFDM) and single-carrier solutions have been adopted in IEEE 802.16 standard as two alternatives for FWA systems operating at 2-11 GHz bands. As stated in [1], the major challenge of designing a FWA system is to provide high data rate wireless access with wire-line quality. The high requirement for quality arises because FWA systems have to compete with cable modems and digital subscriber line (DSL) approaches which operate over fixed channels and hence are able to provide very good quality. In recent years, space-time coding has emerged as one of the most promising technologies for meeting the high data rate and high service quality requirements. Space-time codes were first introduced in [4] to provide transmit diversity in wireless fading channels using multiple antennas. There are two main types of space-time codes, namely, space-time trellis codes (STTC) [4] and space-time block codes (STBC) [5, 6].

Previous work has shown that FWA channels are dispersive, that is they introduce intersymbol interference (ISI) to the transmitted signals, which greatly degrades the system performance. An effective remedy is the use of the equalizer, or more effectively, the turbo equalizer that performs equalization and channel decoding jointly in an iterative manner.

The performance of a turbo equalizer depends largely on the accuracy of channel estimation. Perfect knowledge of the underlying channel is usually assumed in the design of such systems. However, in a practical system, the channel has to be estimated. The estimation of MIMO channels is more challenging compared to that of SISO channels since a large number of channel parameters have to be estimated. The issue of estimating frequency-selective MIMO channels has been addressed in several papers. For example, Kalman filters are used to track the channel variations in [7], and space-time decoding and channel estimation are performed jointly in spatially correlated time-varying channels. In [8], the redundancy induced by space-time block codes is utilized to blindly identify MIMO channels using the subspace method. The algorithm can be further extended to a semi-blind approach by incorporating pilot symbols. A pilot embedding scheme is proposed in [9] in order to obtain an initial estimate of the MIMO channels and improved estimates are obtained by integrating the channel estimator into the iterative decoding loop. The design of optimal training sequences for multiple-antenna systems in a dispersive environment was investigated in [10]. Some sequences with good autocorrelation and cross correlation properties were found by exhaustive search and tabulated in this paper. Although channel estimation in other systems has been studied extensively in both SISO and MIMO cases, this topic has been left relatively unaddressed for the FWA systems except in [11] where a joint frequency offset and channel impulse response estimation algorithm is proposed. It is designed under the least square criterion in an iterative manner and uses the preamble structure provided for an OFDM based FWA uplink. In this paper, we investigate algorithms suitable for estimating frequency-selective MIMO channels in the context of single-carrier FWA systems. Both classical and Bayesian approaches are studied. We show that by exploiting the quasi-static nature of the FWA channels, the performance of the channel estimators can be greatly improved by making some modifications to both the classic and the Bayesian methods.

II. SYSTEM MODEL

Fig. 1 defines the channels between the transmit and receive antennas. Each complex channel coefficient is denoted as h_{ij}^l where the first (second) subscript $i(j)$ is the index of the transmit (receive) antenna, the superscript l refers to the number of the channel tap. For example, h_{10}^0 denotes the channel coefficient corresponding to the first tap of the channel between $tx1$ and $rx0$. For simplicity, we assume 2 transmit antennas and 2 receive antennas in the derivation of channel estimation algorithms. However, the extension to a larger number of antennas is straightforward. The channel coefficients are assumed to remain constant during the

Figure 1: Diagram for the 2TX-2RX SUI-3 MIMO channels.

transmission of one block of data. Tailored for different terrain conditions, a set of 6 typical channel models, known as the Stanford University Interim (SUI) Channel Models have been proposed in [12] for simulation, design, development and testing of FWA systems. For the purpose of this study, we select the SUI-3 channel which has 3 taps with a tap spacing of 500ns, and maximum tap delay of 1000ns. The amplitude of the first tap $|h_{ij}^0|$ is characterized by a Ricean distribution due to the presence of line of sight propagation. The amplitudes of the other two taps $|h_{ij}^1|, |h_{ij}^2|$ are Rayleigh distributed. For simplicity, we assume the transmitted data rate is 4Mbps with QPSK modulation so that the multipath fading is modeled as a tapped-delay line with adjacent taps spaced equally at the symbol duration. The received signals at antenna $rx0$ and $rx1$ can be formed as

$$\begin{aligned} r_n &= h_{00}^2 s_{n-2}^0 + h_{00}^1 s_{n-1}^0 + h_{00}^0 s_n^0 + h_{10}^2 s_{n-2}^1 \\ &\quad + h_{10}^1 s_{n-1}^1 + h_{10}^0 s_n^1 + w_n; \\ y_n &= h_{01}^2 s_{n-2}^0 + h_{01}^1 s_{n-1}^0 + h_{01}^0 s_n^0 + h_{11}^2 s_{n-2}^1 \\ &\quad + h_{11}^1 s_{n-1}^1 + h_{11}^0 s_n^1 + v_n, \end{aligned} \quad (1)$$

where w_n, v_n are the complex additive white Gaussian noise with zero mean and variance N_0 , and s_{n-i}^j denotes the QPSK symbol transmitted from the j^{th} antenna at time instant $n-i$.

The SUI channel models also define an antenna correlation coefficient, which has to be taken into considerations for multiple antenna FWA channels. Antenna correlation is defined as the envelope correlation coefficient between signals received at two antenna elements, i.e.,

$$\rho = \left| \frac{\text{cov}(h_{ij}^l, h_{mn}^l)}{\sqrt{\text{var}(h_{ij}^l) \text{var}(h_{mn}^l)}} \right|, \quad (2)$$

where $i, j, m, n \in \{0, 1\}$, $l \in \{0, 1, 2\}$.

III. CHANNEL ESTIMATION

The task of a channel estimator is to estimate the fading coefficient h_{ij}^l given the received observation $\{r_n, y_n\}$, and some knowledge of the transmitted data. Depending on

whether some prior knowledge about the parameters to be estimated is used or not, estimation can be conducted with a classical or a Bayesian approach. Both approaches are discussed in this paper. The maximum likelihood (ML) channel estimator is based on the classical approach. The ML algorithm has been used e.g., in [13] to estimate the multipath fading channels in DS-CDMA systems. It does not require any prior knowledge of the parameters to be estimated, and the channel coefficient h_{ij}^l is assumed to be deterministic and unknown. With a Bayesian approach, h_{ij}^l is assumed to be a random variable whose particular realization needs to be estimated. The estimation accuracy can hopefully be improved by exploiting some prior knowledge of h_{ij}^l , e.g., channel and noise statistics. The Bayesian philosophy leads, in this paper, to the linear MMSE estimator for channel estimation. These channel estimation algorithms will be presented next.

A. Maximum Likelihood Estimation

Based on (1), the received signal can be expressed in matrix form as $\mathbf{R} = \mathbf{S}\mathbf{H} + \mathbf{V}$, where

$$\begin{aligned} \mathbf{R} &= \begin{bmatrix} r_n & y_n \\ r_{n+1} & y_{n+1} \\ \vdots & \vdots \\ r_{n+D-1} & y_{n+D-1} \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} w_n & v_n \\ w_{n+1} & v_{n+1} \\ \vdots & \vdots \\ w_{n+D-1} & v_{n+D-1} \end{bmatrix}; \\ \mathbf{S} &= \begin{bmatrix} s_n^0 & s_{n-1}^0 & s_{n-2}^0 & s_n^1 & s_{n-1}^1 & s_{n-2}^1 \\ s_{n+1}^0 & s_n^0 & s_{n-1}^0 & s_{n+1}^1 & s_n^1 & s_{n-1}^1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ s_{n+D-1}^0 & s_{n+D-2}^0 & s_{n+D-3}^0 & s_{n+D-1}^1 & s_{n+D-2}^1 & s_{n+D-3}^1 \end{bmatrix}; \\ \mathbf{H} &= \begin{bmatrix} h_{00}^0 & h_{00}^1 & h_{00}^2 & h_{10}^0 & h_{10}^1 & h_{10}^2 \\ h_{01}^0 & h_{01}^1 & h_{01}^2 & h_{11}^0 & h_{11}^1 & h_{11}^2 \end{bmatrix}^T. \end{aligned} \quad (3)$$

Note that the above signal model is formed by stacking D received samples from each receive antenna. Let us denote

$$\begin{aligned} \mathbf{R} &= [\mathbf{r} \quad \mathbf{y}], \quad \mathbf{V} = [\mathbf{w} \quad \mathbf{v}]; \\ \mathbf{r} &= [r_n \quad \dots \quad r_{n+D-1}]^T, \quad \mathbf{y} = [y_n \quad \dots \quad y_{n+D-1}]^T; \\ \mathbf{w} &= [w_n \quad \dots \quad w_{n+D-1}]^T, \quad \mathbf{v} = [v_n \quad \dots \quad v_{n+D-1}]^T; \\ \mathbf{h}_r &= [h_{00}^0 \quad h_{00}^1 \quad h_{00}^2 \quad h_{10}^0 \quad h_{10}^1 \quad h_{10}^2]^T; \\ \mathbf{h}_y &= [h_{01}^0 \quad h_{01}^1 \quad h_{01}^2 \quad h_{11}^0 \quad h_{11}^1 \quad h_{11}^2]^T. \end{aligned} \quad (4)$$

Then the above signal model can be split into two linear vector models $\mathbf{r} = \mathbf{S}\mathbf{h}_r + \mathbf{w}$, and $\mathbf{y} = \mathbf{S}\mathbf{h}_y + \mathbf{v}$. Given the data matrix $\mathbf{S}$, the maximum likelihood estimate [13, 14] of the vector $\mathbf{h}_r$ and $\mathbf{h}_y$ are

$$\hat{\mathbf{h}}_r = (\mathbf{S}^* \mathbf{S})^{-1} \mathbf{S}^* \mathbf{r} = \mathbf{S}^\dagger \mathbf{r}; \quad \hat{\mathbf{h}}_y = (\mathbf{S}^* \mathbf{S})^{-1} \mathbf{S}^* \mathbf{y} = \mathbf{S}^\dagger \mathbf{y},$$

where $\mathbf{S}^\dagger = (\mathbf{S}^* \mathbf{S})^{-1} \mathbf{S}^*$ denotes the left pseudoinverse of $\mathbf{S}$. The superscript operator $(\)^*$ denotes the conjugate transpose operation when applied to matrices and vectors, and simply the conjugate when applied to scalars. Define $\hat{\mathbf{H}} = [\hat{\mathbf{h}}_r \quad \hat{\mathbf{h}}_y]$, we have

$$\begin{aligned} \hat{\mathbf{H}}^{\text{ML}} &= [\mathbf{S}^\dagger \mathbf{r} \quad \mathbf{S}^\dagger \mathbf{y}] = \mathbf{S}^\dagger [\mathbf{r} \quad \mathbf{y}] = \mathbf{S}^\dagger \mathbf{R} \\ &= \mathbf{S}^\dagger (\mathbf{S}\mathbf{H} + \mathbf{V}) = \mathbf{H} + \mathbf{S}^\dagger \mathbf{V}, \end{aligned} \quad (5)$$

which is an unbiased estimate of $\mathbf{H}$ since $\mathbb{E}[\hat{\mathbf{H}}] = \mathbf{H} + \mathbf{S}^\dagger \mathbb{E}[\mathbf{V}] = \mathbf{H}$. Depending on the form of the data that can be retrieved, channel estimation can be either decision-directed or pilot-assisted. The former uses decision feedback loops and utilizes the decisions on the transmitted signals $\hat{\mathbf{S}}$ to extract the channel coefficients. The second approach makes the use of training sequence (pilot symbols), i.e., $\mathbf{S}$ is known in this case. The use of training sequence simplifies channel estimation but wastes channel resources. For quasi-static and slowly-varying fading channels, pilot-assisted channel estimation is often used in practice, and it is the approach we take here in order to simplify the estimation process. However, the proposed estimators can be extended to work in a decision-directed mode using decision feedback from an equalizer or a channel decoder.

The Cramer-Rao Lower Bound (CRLB), which is the bound on the smallest covariance matrix that can be achieved by this ML estimator is

$$\mathbf{C}_{\hat{\mathbf{h}}} = N_0 \mathbb{E}[(\mathbf{S}^* \mathbf{S})^{-1}] \geq \frac{N_0 \mathbf{I}_{2L}}{2E_b D}. \quad (6)$$

The derivation of (6) is left out here to conserve space. In order for the equality in (6) to hold, i.e., to achieve the bound of $\frac{N_0 \mathbf{I}_{2L}}{2E_b D}$, the columns of $\mathbf{S}$ must be orthogonal to each other so that $\mathbf{S}^* \mathbf{S}$ is a diagonal matrix with all the diagonal entries equal to a constant. For flat fading channels, Hadamard codes can be applied to satisfy this condition. However, for multipath channels, the training sequence in each transmit antenna not only has to be orthogonal to its shifts within V taps but also has to be orthogonal to the training sequences in other antennas and their shifts within V taps, where V is the order of channel memory [10]. Some optimal PMV codes meeting the above requirements are given in [10] for code length less than or equal to 16. Another alternative is to use the maximum-length shift-register sequences, or m-sequences for short. Each m-sequence is periodic with period $n = 2^m - 1$, where m is a positive integer. The m-sequence is a binary periodic sequence exhibiting a periodic autocorrelation $\phi(j)$ with values $\phi(j) = n$ for $j = 0$, and $\phi(j) = -1$ for $1 \leq j \leq n - 1$ [15]. This impulse-like autocorrelation implies that the power spectrum is nearly white, hence the sequence resembles white noise. Different training sequences including PVM sequences, m-sequences and random sequences are applied to the proposed channel estimators and their performance compared in Section IV of this paper.

As shown in (6), the estimation variance for the complex channel gain h_{ij}^l is $\text{var}(\hat{h}_{ij}^l) \geq \frac{N_0}{2E_b D}$. This means that the estimation error decreases linearly as the stacking factor D increases. Another way of further improving the estimation performance is to average several independent estimates. Suppose that we have N independent estimates of h_{ij}^l , denoted as $\hat{h}_{ij}^{l1}, \hat{h}_{ij}^{l2}, \dots, \hat{h}_{ij}^{lN}$. A noise resistant estimate of h_{ij}^l can be obtained as $\hat{h}_{ij}^l = \sum_{n=1}^N \hat{h}_{ij}^{ln} / N$. Assume each of the estimates is derived based on the ML algorithm describe above. For each $\hat{h}_{ij}^l$,

$$\mathbb{E}[h_{ij}^l - \hat{h}_{ij}^{ln}] = 0, \quad \mathbb{E}[(h_{ij}^l - \hat{h}_{ij}^{ln})^2] \geq \frac{N_0}{2E_b D}.$$

Using the fact that the estimates are independent from each other, it can be easily shown that

$$\begin{aligned} \mathbb{E}[h_{ij}^l - \hat{h}_{ij}^l] &= \mathbb{E}\left[h_{ij}^l - \sum_{n=1}^N \frac{\hat{h}_{ij}^{ln}}{N}\right] = 0 \\ \mathbb{E}[(h_{ij}^l - \hat{h}_{ij}^l)^2] &= \mathbb{E}\left[\left(h_{ij}^l - \sum_{n=1}^N \frac{\hat{h}_{ij}^{ln}}{N}\right)^2\right] \\ &= \frac{\sum_{n=1}^N \mathbb{E}[(h_{ij}^l - \hat{h}_{ij}^{ln})^2]}{N^2} \geq \frac{NN_0}{2E_b DN^2} = \frac{N_0}{2E_b DN}. \end{aligned} \quad (7)$$

Let us denote $\bar{\mathbf{C}}_{\hat{\mathbf{h}}}$ as the sample averaged version of the covariance matrix $\mathbf{C}_{\hat{\mathbf{h}}}$. Following the same analysis as described above, $\bar{\mathbf{C}}_{\hat{\mathbf{h}}}$ can be derived as

$$\bar{\mathbf{C}}_{\hat{\mathbf{h}}} = \frac{N_0 \mathbb{E}[(\mathbf{S}^* \mathbf{S})^{-1}]}{N} \geq \frac{N_0 \mathbf{I}_{2L}}{2E_b DN}, \quad (8)$$

where L is the number of taps per channel. For the FWA SUI-3 channel, $L = 3$. One can see from (8) that averaging several independent channel estimates also has the effect of reducing estimation errors.

B. Linear MMSE Estimation

In the following, we derive a LMMSE channel estimator and determine its mean square error (MSE) using the Bayesian Gauss-Markov Theorem. The observed data vectors $\mathbf{r}$ and $\mathbf{y}$ are modeled as $\mathbf{r} = \mathbf{S}\mathbf{h}_r + \mathbf{w}$ and $\mathbf{y} = \mathbf{S}\mathbf{h}_y + \mathbf{v}$, respectively, where $\mathbf{r}, \mathbf{y}, \mathbf{w}, \mathbf{v}$ are defined in (4), and $\mathbf{h}_r, \mathbf{h}_y \in \mathbb{C}^{2L \times 1}$ are now assumed to be random vectors whose realization are to be estimated. They have mean $\mathbb{E}[\mathbf{h}_r] = \boldsymbol{\mu}_{\mathbf{h}_r}$, $\mathbb{E}[\mathbf{h}_y] = \boldsymbol{\mu}_{\mathbf{h}_y}$ and covariance matrix $\mathbf{C}_{\mathbf{h}_r}$, $\mathbf{C}_{\mathbf{h}_y}$. The noise vectors $\mathbf{w}, \mathbf{v}$ have PDFs $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, N_0 \mathbf{I}_D)$, $\mathbf{v} \sim \mathcal{CN}(\mathbf{0}, N_0 \mathbf{I}_D)$, and are independent of $\mathbf{h}_r, \mathbf{h}_y$. The Bayesian MMSE estimate of $\mathbf{h}_r$ and $\mathbf{h}_y$ that minimizes MSE averaged over all realizations of $\mathbf{h}_r, \mathbf{h}_y$, and $\mathbf{r}, \mathbf{y}$ can be derived according to [14] as

$$\begin{aligned} \hat{\mathbf{h}}_r &= \boldsymbol{\mu}_{\mathbf{h}_r} + \mathbf{C}_{\mathbf{h}_r} \mathbf{S}^* (\mathbf{S} \mathbf{C}_{\mathbf{h}_r} \mathbf{S}^* + N_0 \mathbf{I}_D)^{-1} (\mathbf{r} - \mathbf{S} \boldsymbol{\mu}_{\mathbf{h}_r}); \\ \hat{\mathbf{h}}_y &= \boldsymbol{\mu}_{\mathbf{h}_y} + \mathbf{C}_{\mathbf{h}_y} \mathbf{S}^* (\mathbf{S} \mathbf{C}_{\mathbf{h}_y} \mathbf{S}^* + N_0 \mathbf{I}_D)^{-1} (\mathbf{y} - \mathbf{S} \boldsymbol{\mu}_{\mathbf{h}_y}). \end{aligned} \quad (9)$$

In the case where $\mathbf{C}_{\mathbf{h}_r} = \mathbf{C}_{\mathbf{h}_y} = \mathbf{C}_{\mathbf{h}}$, equations in (9) can be written in a more compact form as

$$\begin{aligned} \hat{\mathbf{H}}^{\text{MMSE}} &= [\hat{\mathbf{h}}_r \quad \hat{\mathbf{h}}_y] = [\boldsymbol{\mu}_{\mathbf{h}_r} \quad \boldsymbol{\mu}_{\mathbf{h}_y}] \\ &\quad + \mathbf{C}_{\mathbf{h}} \mathbf{S}^* (\mathbf{S} \mathbf{C}_{\mathbf{h}} \mathbf{S}^* + N_0 \mathbf{I}_D)^{-1} ([\mathbf{r} \quad \mathbf{y}] - \mathbf{S} [\boldsymbol{\mu}_{\mathbf{h}_r} \quad \boldsymbol{\mu}_{\mathbf{h}_y}]) \\ &= \boldsymbol{\mu}_{\mathbf{H}} + \mathbf{C}_{\mathbf{h}} \mathbf{S}^* (\mathbf{S} \mathbf{C}_{\mathbf{h}} \mathbf{S}^* + N_0 \mathbf{I}_D)^{-1} (\mathbf{R} - \mathbf{S} \boldsymbol{\mu}_{\mathbf{H}}), \end{aligned} \quad (10)$$

where $\boldsymbol{\mu}_{\mathbf{H}} = [\boldsymbol{\mu}_{\mathbf{h}_r} \quad \boldsymbol{\mu}_{\mathbf{h}_y}]$. We can see from (10) that the first and second order statistics of the channel are required to perform the linear MMSE channel estimation.

The performance of this estimator is measured by the error $\mathbf{e} = \mathbf{h} - \hat{\mathbf{h}}$ (where $\mathbf{h}$ represents either $\mathbf{h}_r$ or $\mathbf{h}_y$), whose mean is zero and covariance matrix (which is equivalent to the MSE matrix) is [14]

$$\begin{aligned} \text{MSE}(\mathbf{h}) &= \mathbf{C}_{\mathbf{e}} = \mathbb{E}(\mathbf{e} \mathbf{e}^T) = \mathbf{C}_{\mathbf{h}} - \mathbf{C}_{\mathbf{h}_r} \mathbf{C}_{\mathbf{r}\mathbf{r}}^{-1} \mathbf{C}_{\mathbf{r}\mathbf{h}} \\ &= \mathbf{C}_{\mathbf{h}} - \mathbf{C}_{\mathbf{h}} \mathbf{S}^* (\mathbf{S} \mathbf{C}_{\mathbf{h}} \mathbf{S}^* + \mathbf{C}_{\mathbf{w}})^{-1} \mathbf{S} \mathbf{C}_{\mathbf{h}}. \end{aligned}$$

Figure 2: Performance of the ML channel estimator for the SUI-3 MIMO channels ($D \times N = 31 \times 12$).

Using the matrix inversion lemma $(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{DA}^{-1}\mathbf{B} + \mathbf{C}^{-1})^{-1}\mathbf{DA}^{-1}$, and assigning $\mathbf{A} = \mathbf{C}_h^{-1}$, $\mathbf{B} = \mathbf{S}^*$, $\mathbf{C} = \mathbf{C}_w^{-1} = (N_0\mathbf{I})^{-1}$, $\mathbf{D} = \mathbf{S}$, yields

$$\begin{aligned} \text{MSE}(\mathbf{h}) &= (\mathbf{C}_h^{-1} + \mathbf{S}^*(N_0\mathbf{I})^{-1}\mathbf{S})^{-1} \\ &= \left(\mathbf{C}_h^{-1} + \frac{\mathbf{S}^*\mathbf{S}}{N_0} \right)^{-1} \geq \left(\mathbf{C}_h^{-1} + \frac{2E_b\mathbf{DI}}{N_0} \right)^{-1}. \end{aligned}$$

Following the same procedure as in (7), it can be shown that the MSE of this algorithm can be decreased by N fold by averaging N independent estimates, i.e.,

$$\overline{\text{MSE}}(\mathbf{h}) \geq \left(\mathbf{C}_h^{-1} + \frac{2E_b\mathbf{DI}}{N_0} \right)^{-1} / N, \quad (11)$$

where $\overline{\text{MSE}}(\mathbf{h})$ is the sample averaged version of $\text{MSE}(\mathbf{h})$.

IV. ANALYTICAL AND SIMULATION RESULTS

Analytical and simulation results are presented in this section to assess and compare the performance of the proposed channel estimation schemes. The simulation curves are obtained by averaging the simulation results over 1000 channel realizations. The FWA channel coefficients vary from one block to another, however, they are assumed to remain constant during the transmission of one block of data. It is therefore a quasi-static channel. The antenna correlation coefficient defined in (2) is set to $\rho = 0.4$. The noise variance N_0 and path delays as well as the channel mean value and covariance are assumed to be known to the receiver.

The performance of the ML and linear MMSE channel estimator using random sequence and m-sequence is shown in Fig. 2 and 3, respectively. The employed m-sequence has a degree of $m = 5$, and is obtained with the generator polynomial $g(p) = p^5 + p^3 + 1$. It is a periodic binary sequence with period $n = 2^5 - 1 = 31$. The binary sequence is then converted to a symbol sequence in such a way that 0 is mapped to symbol $\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}$, and 1 is mapped to symbol $\frac{-1}{\sqrt{2}} - j\frac{1}{\sqrt{2}}$. The symbol sequence and its cyclic shifted

Figure 3: Performance of the MMSE channel estimator for the SUI-3 MIMO channels ($D \times N = 31 \times 12$).

version are transmitted from antenna $tx0$ and $tx1$, respectively. Note that there has to be enough shift to guarantee low cross correlation between the two antennas. We use the mean square error (MSE) as the performance measurement (MSE is equivalent to the variance of the estimation error in the ML case since the ML estimator is unbiased). As shown in Fig. 2, the simulation results for the ML estimator is in close agreement with the CRLB expressed by (8) for the m-sequence. In case random QPSK symbols are transmitted as training sequence, the theoretical curve is derived semi-analytically by averaging the matrix $(\mathbf{S}^*\mathbf{S})^{-1}$ in the simulations, and scaling it with factor $\frac{N_0}{N}$. It exhibits a small discrepancy compared with the CRLB because the matrix $\mathbf{S}$ in (8) is not strictly orthogonal for random sequences. Fig. 3 shows that the simulated MSE for the linear MMSE estimator slightly deviates from the theoretical analysis expressed in (11) at low SNR. However, the simulated curve for the m-sequence gradually converges to the performance lower bound as SNR increases. By comparing Fig. 2 with 3, it can be observed that the ML estimator yields comparable performance to the MMSE estimator. Considering the fact that the ML estimator does not require any prior knowledge about the channel to be estimated, it is the preferred choice for the SUI-3 FWA channel under investigation.

We know from the analysis and simulations that the estimation error decreases as the stacking factor D or averaging factor N increases, which implies two ways of improving channel estimation performance. For example, at low E_b/N_0 , we can achieve the same performance obtained with a higher E_b/N_0 by increasing the values of D and/or N . However, increasing the value of D also increases the complexity of the channel estimation due to the matrix inverse operation in (5) and (10). Note that the performance improvement comes at the penalty of wasting system resources since increasing the values of D and/or N means transmitting more pilot symbols. It also should be noted that the sample averaging technique is only useful for static and slowly time-varying (e.g., quasi-static) ISI chan-

Figure 4: Comparison of different training sequences ($D \times N = 31 \times 12$).

nels. For fast time-varying channels, some other techniques, e.g., sample smoothing [13], or the Kalman filter [16] can be used.

Different training sequences are compared in Fig. 4 for the ML estimator. The chosen PVM sequences are $X_1 = \{111010001110\}$ and $X_2 = \{111101101000\}$, which are taken from [10]. In this case, the parameter settings are $P = 12, V = 2, M = 2$. The binary sequence to symbol sequence conversion process is the same as stated above. It can be seen from Fig. 4 that the m-sequence yields almost identical performance to the optimal PVM sequence. They both achieve CRLB. The performance of random sequence is slightly worse. However, the gap is very small, the performance loss by using random sequence is only about 0.2 dB.

V. CONCLUSIONS

Turbo equalization is widely proposed to combat the detrimental effects of ISI, which is one of the major impairments in wireless communication systems. Most turbo equalizers rely on accurate channel information, which in practice needs to be estimated. A modified version of the ML and the linear MMSE estimator with sample stacking and averaging has been proposed in this paper and shown to be very efficient for estimating static or quasi-static frequency-selective channels. By comparison, the ML estimator is a preferred solution for the FWA MIMO systems. Different training sequences are compared in this paper, and the results show that the choice of training sequences is not very significant in the estimation performance. Although originally designed for the FWA MIMO channels, the proposed channel estimation algorithms can be applied to other static or quasi-static channels and other STBC coded systems.

REFERENCES

[1] H. Bolcskei, A. Paulraj, K. Hari, R. Nabar, W. Lu. "Fixed Broadband Wireless Access: State of the Art, Challenges, and Future Directions". *IEEE Communication Magazine*, pp. 100–108, Jan. 2001.

[2] IEEE 802.16 Working Group on Broadband Wireless Access Standards. available at <http://grouper.ieee.org/groups/802/16/>. 2002.

[3] Draft ETSI TS 101, Broadband Radio Access Networks (BRAN), HIPERMAN. available at <http://portal.etsi.org/bran/Summary.asp>, 2002.

[4] V. Tarokh, N. Seshadri, A. Calderbank. "Space-time codes for high data rates wireless communications: performance criterion and code construction". *IEEE Transactions on Information Theory*, vol. 44, pp. 744–765, March 1998.

[5] A. Alamouti. "A simple transmit diversity technique for wireless communications". *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[6] V. Tarokh, N. Seshadri, A. Calderbank. "Space-time block coding from orthogonal designs". *IEEE Transactions on Information Theory*, vol. 45, pp. 1456–1467, July 1999.

[7] V. Nguyen, L. White. "Joint space-time trellis decoding and channel estimation in correlated fading channels". *IEEE Signal Processing Letters*, vol. 11, no. 7, pp. 633–636, July 2004.

[8] J. Choi. "Equalization and semi-blind channel estimation for space-time block coded signals over a frequency-selective fading channel". *IEEE Transactions on Signal Processing*, vol. 52, no. 3, pp. 774–785, March 2004.

[9] H. Zhu, B. Farhang-Boroujeny, C. Schlegel. "Pilot embedding for joint channel estimation and data detection in MIMO communication systems". *IEEE Communications Letter*, vol. 7, no. 1, Jan. 2003.

[10] S. Yang, J. Wu. "Optimal binary training sequence design for multiple-antenna systems over dispersive fading channels". *IEEE Transactions on Vehicular Technology*, vol. 51, no. 5, pp. 1271–1276, Sept. 2002.

[11] N. Balamurali, D. Jalihal. "An efficient algorithm for joint carrier frequency offset and channel estimation in IEEE 802.16 OFDM systems". *Proceedings 1st International Symposium on Wireless Communication Systems*, pp. 428–432, Sept. 2004.

[12] V. Erceg. "An empirically based path loss model for wireless channels in suburban environments". *IEEE JSAC*, vol. 17, no. 7, pp. 1205–1211, July 1999.

[13] P. Xiao, E. Ström. "Iterative demodulation of M-ary orthogonal signaling formats in coded DS-CDMA systems with soft interference cancellation and channel estimation". *IEEE Transactions on Communications*, vol. E89-B, no. 1, pp. 150–161, January 2006.

[14] S. Kay. *Fundamentals of statistical signal processing*, Prentice Hall, NJ, 1998.

[15] J. Proakis. *Digital Communications*, 3rd edition, McGraw-Hill, 1995.

[16] R. Iltis. "Joint estimation of PN code delay and multipath using the extended Kalman filter". *IEEE Transactions on Communications*, vol. 38, no. 10, pp. 1677–1685, Oct. 1990.