Reduced-Complexity Training Schemes for Multiple-Antenna Broadband Transmissions

C. Fragouli, N. Al-Dhahir, W. Turin
AT& T Shannon Laboratories
180 Park Ave,
Florham Park, NJ 07932

Abstract– **This paper addresses the problem of training sequence design for multiple-antenna transmissions over quasi-static frequency-selective channels. As performance metric for channel estimation, mean square error is adopted. To achieve the minimum mean square error, the training sequences transmitted from the multiple antennas must have impulse-like auto-correlation and zero cross-correlation. We reduce the problem of designing multiple training sequences to the much easier and well-understood problem of designing a single training sequence with impulse-like auto-correlation. To this end, we propose to encode the training sequences with a space-time code, that may be the same or different from the space-time code that encodes the information symbols. Designing one instead of multiple training sequences reduces the search space significantly and simplifies the construction of optimal or suboptimal training sequences.**

I. Introduction

This paper addresses the problem of constructing training sequences to estimate the impulse response coefficients of frequency-selective channels, in a system with multiple transmit/receive antennas that employs a space-time code. Consider the multiple–transmit single–receive[1] transmission scenario. The receiver observes the superposition of training sequences transmitted through different channels. The training sequences that achieve the channel estimation minimum mean square error[2] have an impulse-like auto-correlation sequence and additionally zero cross-correlation. This last property makes the channel estimation problem different for multiple-antenna systems from single-antenna systems, and has motivated research in this area.

Training-based estimation for a single-input single-output (SISO) frequency-selective channel has been widely investigated in the literature (see for example [1] and the references therein). Training for multiple-transmit-antenna systems, can be classified in two main approaches. The first approach employs optimal training sequences[3] such as the perfect roots-of-unity sequences (PRUS) (see [2] and the references therein). A PRUS generally does not belong to a standard constant amplitude constellation of size 2^m such as BPSK, QPSK etc.

The second approach employs exhaustive searches to identify suboptimal training sequences from a standard constellation to achieve a simpler transmitter/receiver implementation. Tables of such sequences from a BPSK alphabet are provided, for example, in [3]. Such exhaustive searches are in many cases computationally prohibitive. As an example, in the third generation TDMA proposal EDGE [4] (Enhanced Data for GSM Evolution), 8-PSK constellation symbols are transmitted in blocks of $N_i = 116$ information symbols and $N_t = 26$ training symbols. No optimal training sequence exists for this N_t and constellation size. For two-transmit antennas, an exhaustive search would involve $8^{2\times 26}$ sequences. Restricting the training sequence alphabet size to BPSK would reduce the search space to $2^{2\times 26}$ sequences, which is still large and would increase the minimum achievable mean square error.

This paper proposes to reduce the problem of designing multiple training sequences with impulse-like auto-correlation and zero cross-correlation to designing a single training sequence with impulse-like auto-correlation. This makes exhaustive searches more practical and thus facilitates the identification of good training sequences from standard constellation alphabets. In some cases no search is necessary since optimal sequences are available from published results in the literature.

The paper is organized as follows. Section II presents the channel model and formally defines the optimal training sequences. Section III proposes methods to generate multiple training sequences starting from a single one. Section IV presents simulation results and the paper is concluded in Section V.

Contact author: N. Al-Dhahir, tel: (973)-360-8407, fax: (973)-360-8178

[1] The number of receiver antennas does not affect the training setup as long as the antennas are spaced wide enough to ensure uncorrelated received signals.

[2] Throughout this paper we use the mean square error as performance metric for channel estimation.

[3] By optimal sequences, we mean sequences that achieve the channel estimation minimum mean square error.

II. Channel model and optimal training sequences

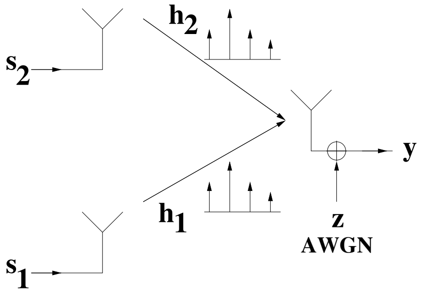

Fig. 1. System with two-transmit and one-receive antennas.

To simplify the presentation, consider the system of Fig. 1 that employs two-transmit and one-receive antennas. The analysis can be generalized to multiple transmit/receive antennas.

Two signals $\mathbf{s}_1$ and $\mathbf{s}_2$ are simultaneously transmitted over two frequency-selective channels $\mathbf{h}_1 = [h_1(0) \ \dots \ h_1(L-1)]^T$ and $\mathbf{h}_2 = [h_2(0) \ \dots \ h_2(L-1)]^T$, where $(.)^T$ denotes the transpose operation. Each channel is modeled as an FIR filter with L taps. We assume that the channels $\mathbf{h}_1$ and $\mathbf{h}_2$ remain constant over the transmission of a block and vary independently from block to block. The input sequences $\mathbf{s}_1$ and $\mathbf{s}_2$ belong to a finite signal constellation and are transmitted in data blocks where each block consists of N_i information symbols and N_t training symbols. For two-transmit antennas the receiver uses the $2N_t$ known training symbols to estimate the $2L$ unknown channel coefficients.

The observed training sequence output, that does not have interference from information or preamble symbols, can be expressed as

$$\mathbf{y} = \mathbf{S}\mathbf{h} + \mathbf{z} = [\mathbf{S}_1 \ \ \mathbf{S}_2] \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \end{bmatrix} + \mathbf{z},$$

where $\mathbf{y}$ and $\mathbf{z}$ are of dimension $(N_t - L + 1) \times 1$ and $\mathbf{S}_1$ and $\mathbf{S}_2$ are Toeplitz matrices of dimension $(N_t - L + 1) \times L$

$$\mathbf{S}_i = \begin{bmatrix} s_i(L-1) & \dots & s_i(0) \\ s_i(L) & \dots & s_i(1) \\ \vdots & & \vdots \\ s_i(N_t-1) & \dots & s_i(N_t-L) \end{bmatrix}, \ \text{for } i = 1, 2.$$

We assume that $\mathbf{z}$ is Additive White Gaussian Noise (AWGN) with auto-correlation matrix $R_z = E\mathbf{z}\mathbf{z}^H = 2\sigma^2\mathbf{I}$ where $\mathbf{I}$ is the identity matrix. The linear least square channel estimates, assuming that $\mathbf{S}$ has full column rank, can be calculated as [5]

$$\hat{\mathbf{h}} = \begin{bmatrix} \hat{\mathbf{h}}_1 \\ \hat{\mathbf{h}}_2 \end{bmatrix} = (\mathbf{S}^H\mathbf{S})^{-1}\mathbf{S}^H\mathbf{y}, \qquad (3)$$

where $(.)^H$ and $(.)^{-1}$ denote the complex-conjugate (Hermitian) transpose and the inverse, respectively. The channel estimation mean square error (MSE) is defined as

$$MSE = E[(\mathbf{h} - \hat{\mathbf{h}})^H(\mathbf{h} - \hat{\mathbf{h}})] = 2\sigma^2 tr((\mathbf{S}^H\mathbf{S})^{-1}),$$

where $tr(.)$ denotes the trace of a matrix. The minimum MSE (MMSE) is equal to

$$MMSE = \frac{2\sigma^2 L}{(N_t - L + 1)}, \qquad (5)$$

which is achieved if and only if [6]

$$\mathbf{S}^H\mathbf{S} = \begin{bmatrix} \mathbf{S}_1^H\mathbf{S}_1 & \mathbf{S}_2^H\mathbf{S}_1 \\ \mathbf{S}_1^H\mathbf{S}_2 & \mathbf{S}_2^H\mathbf{S}_2 \end{bmatrix} = (N_t - L + 1)\mathbf{I}. \qquad (6)$$

The sequences $\mathbf{s}_1$ and $\mathbf{s}_2$ that satisfy (6) are henceforth referred to as *optimal* sequences. Equation (6) implies that the optimal sequences have an impulse-like auto-correlation sequence and zero cross-correlation.

III. Complexity Reduction methods

This section provides two methods to reduce the complexity of designing training sequences for multiple-antenna systems to that of designing a single training sequence.

A. A block code for training symbols

This section proposes to encode the training symbols with a block code that takes an input sequence $\mathbf{s}$ with impulse-like auto-correlation and produces sequences $\mathbf{s}_1$ and $\mathbf{s}_2$ with zero cross-correlation. Such a code for two transmit antennas is depicted in Fig. 2. The received out-

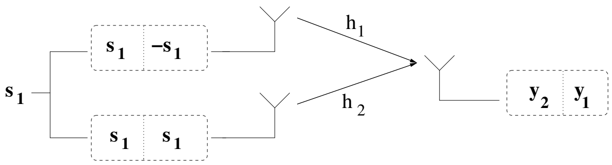

Fig. 2. A simple two-antenna training scheme for frequency-selective channels.

put can be expressed as

$$\begin{aligned}\left[\begin{array}{c}\mathbf{y}_1\\ \mathbf{y}_2\end{array}\right] &= \underbrace{\left[\begin{array}{cc}-\mathbf{S}_1 & \mathbf{S}_1\\ \mathbf{S}_1 & \mathbf{S}_1\end{array}\right]}_{\mathbf{S}}\left[\begin{array}{c}\mathbf{h}_1\\ \mathbf{h}_2\end{array}\right]+\left[\begin{array}{c}\mathbf{z}_1\\ \mathbf{z}_2\end{array}\right]\\ &= \left[\begin{array}{cc}\mathbf{S}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{S}_1\end{array}\right]\underbrace{\left[\begin{array}{cc}-\mathbf{I} & \mathbf{I}\\ \mathbf{I} & \mathbf{I}\end{array}\right]}_{\mathbf{U}}\left[\begin{array}{c}\mathbf{h}_1\\ \mathbf{h}_2\end{array}\right]+\left[\begin{array}{c}\mathbf{z}_1\\ \mathbf{z}_2\end{array}\right],\end{aligned} \quad (7)$$

where $\mathbf{y}_i$ and $\mathbf{z}_i$, $i = 1, 2$, are of dimension $(N_t - L + 1) \times 1$, and $\mathbf{S}_1$ is of dimension $(N_t - L + 1) \times L$. Multiplying the received output with the transpose-conjugate matrix $\mathbf{S}^H$ we get

$$\left[\begin{array}{c}\mathbf{r}_1\\ \mathbf{r}_2\end{array}\right] = \left[\begin{array}{cc}2\mathbf{S}_1^H\mathbf{S}_1 & \mathbf{0}\\ \mathbf{0} & 2\mathbf{S}_1^H\mathbf{S}_1\end{array}\right]\left[\begin{array}{c}\mathbf{h}_1\\ \mathbf{h}_2\end{array}\right]+\left[\begin{array}{c}\overline{\mathbf{z}}_1\\ \overline{\mathbf{z}}_2\end{array}\right], \quad (8)$$

where the noise now becomes

$$\left[\begin{array}{c}\overline{\mathbf{z}}_1\\ \overline{\mathbf{z}}_2\end{array}\right] = \left[\begin{array}{cc}-\mathbf{S}_1^H & \mathbf{S}_1^H\\ \mathbf{S}_1^H & \mathbf{S}_1^H\end{array}\right]\left[\begin{array}{c}\mathbf{z}_1\\ \mathbf{z}_2\end{array}\right]. \quad (9)$$

If we choose $\mathbf{S}_1^H\mathbf{S}_1 = (N_t - L + 1)\mathbf{I}$, then

$$\left[\begin{array}{c}\mathbf{y}_1\\ \mathbf{y}_2\end{array}\right] = 2(N_t - L + 1)\left[\begin{array}{c}\mathbf{h}_1\\ \mathbf{h}_2\end{array}\right]+\left[\begin{array}{c}\overline{\mathbf{z}}_1\\ \overline{\mathbf{z}}_2\end{array}\right]. \quad (10)$$

Moreover, if the noise is white, i.e.

$$E\left(\left[\begin{array}{c}\mathbf{z}_1\\ \mathbf{z}_2\end{array}\right]\left[\mathbf{z}_1^H\mathbf{z}_2^H\right]\right) = 2\sigma^2\mathbf{I}, \quad (11)$$

then the linear processing at the receiver does not color it

$$\left[\begin{array}{c}\overline{\mathbf{z}}_1\\ \overline{\mathbf{z}}_2\end{array}\right]\left[\overline{\mathbf{z}}_1^H\ \overline{\mathbf{z}}_2^H\right] = 2(N_t - L + 1)\sigma^2\mathbf{I}. \quad (12)$$

The MSE of this scheme achieves the lower bound

$$MSE = \frac{2\sigma^2}{2(N_t - L + 1)}.$$

Note that we could use any orthogonal matrix $\mathbf{U}$ in (7) such that $\mathbf{U}^H\mathbf{U} = 2\mathbf{I}$ to encode the training sequence for two transmit antennas. Generally, for n transmit antennas, we could use a matrix $\mathbf{U}$ such that $\mathbf{U}^H\mathbf{U} = n\mathbf{I}$.

If an optimal sequence $\mathbf{s}_1$ does not exist, we can instead use a sequence with good auto-correlation properties. The block code structure would still guarantee the orthogonality between $\mathbf{s}_1$ and $\mathbf{s}_2$. However, in this case the noise will not be white.

This method assumes no interference between the two consecutive transmitted training symbol blocks in Fig. 2. Thus, either the interference can be taken into account in the MSE calculation, or two training blocks may be used assuming that the channel remains constant over two consecutive blocks.

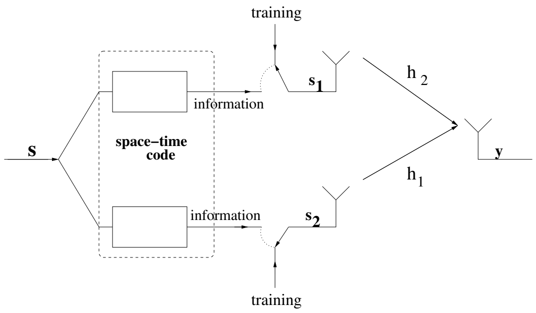

Fig. 3. A training scheme with independent sequences $\mathbf{s}_1$ and $\mathbf{s}_2$ to estimate the channels $\mathbf{h}_1$ and $\mathbf{h}_2$.

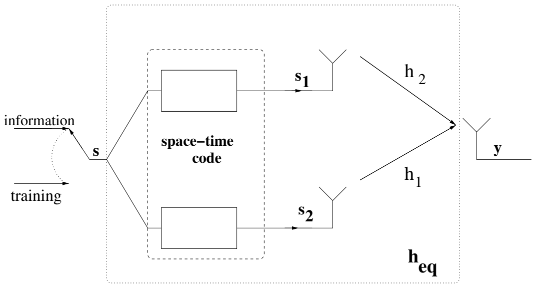

Fig. 4. A training scheme that directly estimates $\mathbf{h}_{eq}$ with sequences $\mathbf{s}_1$ and $\mathbf{s}_2$ generated from a space-time encoded sequence $\mathbf{s}$.

B. Training symbols encoded with a space-time code

In this section, we propose to use a single training sequence encoded by the same space-time encoder as the information symbols. This approach differs from previous work in the literature in that the two transmitted training sequences $\mathbf{s}_1$ and $\mathbf{s}_2$ are not independent sequences (Fig. 3); instead they are generated by a space-time encoded sequence $\mathbf{s}$ (Fig. 4).

The code structure imposes a constraint on the possible generated $\mathbf{s}_1$ and $\mathbf{s}_2$ sequences to achieve diversity and coding gain as shown in [7]. This constraint amounts to a reduction of the search space from C^{2N_t} to C^{N_t} (assuming equal input and output alphabet size C and two transmit antennas), making exhaustive searches more practical and thus facilitating the identification of good training sequences. The search space can be further reduced by exploiting special characteristics of the employed space-time code. As an example, in the following we examine a specific space-time trellis and a specific space-time block code.

B.1 Space-time Trellis Code

A space-time code with m binary memory elements and inputs/outputs from a constellation of size C has always an implementation with at most m memory elements that store constellation symbols. The receiver can incorporate the space-time trellis code structure in the channel model to create an equivalent single-input single-output channel $\mathbf{h}_{eq}$, of length $m + L$. The trellis will involve C^{m+L-1} states, which does not depend on the number of transmit antennas. Our proposed approach uses one training sequence at the input of the space-time trellis encoder to directly estimate $\mathbf{h}_{eq}$ used by the joint space-time equalizer/decoder.

Some special space-time trellis codes allow a nonlinear implementation with less than m (non binary) memory elements leading to an equivalent channel $\mathbf{h}_{eq}$ of length smaller than $m + L$ which makes joint equalization/decoding possible without a prohibitive increase in complexity. As an example consider the 8-state 8-PSK space-time trellis code [7] for two transmit and one receive antennas. Let the complex 8-PSK symbols be described as $s(k) = exp(2i\pi p_k/8)$, where the integer $p_k = u_{1,k} + 2u_{2,k} + 4u_{3,k}$ takes values in $\{0, 1, \ldots 7\}$ and $u_{i,k}$, $i = 1 \ldots 3$ denote the binary inputs at time k. The transmitted signals at time k can be expressed as

$$s_1(k) = s(k), \quad s_2(k) = (-1)^{p_{k-1}} s(k-1). \tag{14}$$

The received signal at time k can be expressed as

$$y(k) = \sum_{i=0}^{L} h_{eq}(i,k) s(k-i) + z(k),$$

where $h_{eq}(i,k)$ are the taps of the equivalent input-dependent channel given by

$$h_{eq}(i,k) = \begin{cases} h_1(0) & i = 0 \\ (-1)^{p_{k-L}} h_2(L-1) & i = L \\ h_1(i) + (-1)^{p_{k-i}} h_2(i-1) & 0 < i < L. \end{cases} \tag{16}$$

This nonlinear channel representation allows us to easily incorporate the space-time code structure in the equivalent channel $\mathbf{h_{eq}}$ by increasing the number of channel taps by one. Thus, the special structure of this code allows us to reduce the number of joint equalizer/decoder states from 8^{2L-1} to 8^L.

For the examined code the number of unknown parameters is $2L$ whether we estimate $\mathbf{h}_{eq}$ or $\mathbf{h}_1$ and $\mathbf{h}_2$.

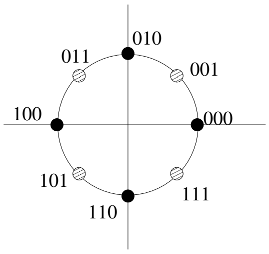

Fig. 5. 8-PSK constellation labeling in binary notation.

To further reduce the training sequence search space for the 8-state 8-PSK space-time trellis code [7] we can take advantage of the $\mathbf{h_{eq}}$ channel special structure.

For a given block and constant $\mathbf{h}_1$ and $\mathbf{h}_2$, the input sequence determines the sequence of $\mathbf{h}_{eq}$ values. From (16) by transmitting only "even" training symbols from the set $C_e = \{0, 2, 4, 6\}$ we observe the $L + 1$ channel taps

$$\mathbf{h}_e = [h_1(0) \quad h_1(i) + h_2(i-1) \quad h_2(L-1)],$$
$$1 \le i \le L-1,$$

while by transmitting only "odd" training symbols from the set $C_o = \{1, 3, 5, 7\}$ we observe the $L + 1$ channel taps

$$\mathbf{h}_o = [h_1(0) \quad h_1(i) - h_2(i-1) \quad -h_2(L-1)],$$
$$1 \le i \le L-1.$$

After estimating $\mathbf{h}_e$ and $\mathbf{h}_o$, we can obtain $\mathbf{h}_1 = \frac{\mathbf{h}_e + \mathbf{h}_o}{2}$ and $\mathbf{h}_2 = \frac{\mathbf{h}_e - \mathbf{h}_o}{2}$.

The middle taps of $\mathbf{h}_e$ and $\mathbf{h}_o$ correspond to distinct unknowns. If all taps of $\mathbf{h}_e$ and $\mathbf{h}_o$ were distinct unknowns, we could use half the training sequence length to estimate each channel separately without increase in channel estimation MSE. To estimate $\mathbf{h}_e$, we would use training symbols in the C_e sub-constellation (Fig. 5), while to estimate $\mathbf{h}_o$ we would use training symbols in the C_o sub-constellation. To estimate both, we propose to use half of the input training symbols from each sub-constellation. That is, for even training length N_t, we propose to use a training sequence of the form $\mathbf{s} = [\mathbf{s}_e \quad \mathbf{s}_o]$ where $\mathbf{s}_e$ has length $N_t/2$ and takes values in the C_e sub-constellation and $\mathbf{s}_o$ has length $N_t/2$ and takes values in the C_o sub-constellation.

Note that if $\mathbf{s}_e$ is a good sequence (in terms of minimizing MSE) for the estimation of $\mathbf{h}_e$, the sequence $\mathbf{s}_o$ created as

$$\mathbf{s}_o = a\, \mathbf{s}_e : \quad \text{for} \;\; a = exp(\frac{i\pi k}{4}) \tag{19}$$

and any $k = 1,3,5,7$, achieves the same MSE for the estimation of $\mathbf{h}_o$. Thus, instead of searching over all possible $N^{N_t} = 8^{N_t}$ sequences $\mathbf{s}$, we can further restrict the search space to the $4^{\frac{N_t}{2}}$ sequences $\mathbf{s}_e$.

Because the first and last taps of $\mathbf{h}_e$ and $\mathbf{h}_o$ are common, there is an advantage in jointly estimating them. An exhaustive search can identify sequences $\mathbf{s}_e$ and $\mathbf{s}_o = a\ \mathbf{s}_e$ such that the overall matrix $\mathbf{S}$ corresponding to the whole training sequence $[\mathbf{s}_e \quad \mathbf{s}_o]$ achieves the minimum MSE. In this case the value of a may make a difference. Section IV provides a table of such sequences and simulation results.

B.2 Space-time block code

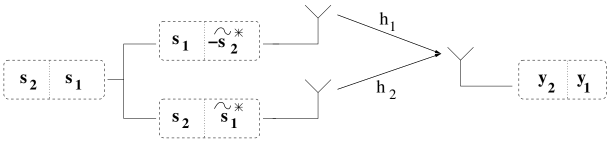

Fig. 6. Space-time block code for two transmit antennas over frequency-selective channels.

This section presents training schemes suitable for the space-time block code in [8] which is an extension of the code in [9] for frequency-selective channels. The encoder maps two consecutive input blocks $\mathbf{s}_1$ and $\mathbf{s}_2$ to the blocks $[\mathbf{s}_1 \quad -\tilde{\mathbf{s}}_2^*]$ and $[\mathbf{s}_2 \quad \tilde{\mathbf{s}}_1^*]$ to be transmitted from the two antennas, as illustrated in Fig. 6. The operation denoted by $\tilde{(.)}$ refers to time-reversing a sequence, that is, if $\mathbf{s} = [s(0) \ \ s(1) \ \ \ldots \ \ s(N_t-1)]$ then $\tilde{\mathbf{s}} = [s(N_t-1) \ \ \ldots \ \ s(1) \ \ s(0)]$, and $(.)^*$ refers to component-wise complex conjugation. Assume that the block code is applied to the training symbols and that the channels $\mathbf{h}_1$ and $\mathbf{h}_2$ to be estimated remain constant over two blocks. The received signals during the first and second blocks denoted by $\mathbf{y}_1$ and $\mathbf{y}_2$ respectively, can be expressed in matrix notation as

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \underbrace{\begin{bmatrix} -\tilde{\mathbf{S}}_2^* & \tilde{\mathbf{S}}_1^* \\ \mathbf{S}_1 & \mathbf{S}_2 \end{bmatrix}}_{\mathbf{S}} \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix}, \qquad (20)$$

where the matrices $\mathbf{S}_i$ and $\tilde{\mathbf{S}}_i$ (for $i = 1,2$) are of dimension $(N_t - L + 1) \times L$. Then,

$$\mathbf{S}^*\mathbf{S} = \begin{bmatrix} \tilde{\mathbf{S}}_2^T\tilde{\mathbf{S}}_2^* + (\mathbf{S}_1^*)^T\mathbf{S}_1 & -\tilde{\mathbf{S}}_2^T\tilde{\mathbf{S}}_1^* + (\mathbf{S}_1^*)^T\mathbf{S}_2 \\ -\tilde{\mathbf{S}}_1^T\tilde{\mathbf{S}}_2^* + (\mathbf{S}_2^*)^T\mathbf{S}_1 & \tilde{\mathbf{S}}_1^T\tilde{\mathbf{S}}_1^* + (\mathbf{S}_2^*)^T\mathbf{S}_2 \end{bmatrix}.$$

From (6), a necessary condition to achieve the MSE lower bound is to have zero cross-correlation terms, i.e.

$$-\tilde{\mathbf{S}}_1^T\tilde{\mathbf{S}}_2^* + (\mathbf{S}_2^*)^T\mathbf{S}_1 = 0 = (-\tilde{\mathbf{S}}_2^T\tilde{\mathbf{S}}_1^* + (\mathbf{S}_1^*)^T\mathbf{S}_2)^T. \quad (22)$$

If $\mathbf{J}_x = [\delta_{i,x-i+1}]$ denotes the $x \times x$ square reversion ("backward identity") matrix, where $\delta_{i,j}$ is the delta function, then

$$\tilde{\mathbf{S}}_i = \mathbf{J}_{N_t-L+1}\mathbf{S}_i\mathbf{J}_L \ \text{ for } \ i = 1,2. \qquad (23)$$

Since $\mathbf{J}_x = \mathbf{J}_x^T = \mathbf{J}_x^{-1}$, by substituting (23) in (22) we have an equivalent condition

$$\mathbf{J}_L((\mathbf{S}_2^*)^T\mathbf{S}_1)^T\mathbf{J}_L = (\mathbf{S}_2^*)^T\mathbf{S}_1. \qquad (24)$$

Moreover, an additional requirement to achieve the MSE lower bound is that

$$\tilde{\mathbf{S}}_2^T\tilde{\mathbf{S}}_2^* + (\mathbf{S}_1^*)^T\mathbf{S}_1 = 2(N_t - L + 1)\mathbf{I}. \qquad (25)$$

Two simple choices that satisfy conditions (25) and (24) are

1. $(\mathbf{S}_1^*)^T\mathbf{S}_1 = (N_t - L + 1)\mathbf{I}$, $\tilde{\mathbf{S}}_1 = \mathbf{S}_1$, and $\mathbf{S}_2 = \mathbf{S}_1$. That is, identify a sequence $\mathbf{s}_1$ symmetric about its center with impulse-like auto-correlation and set $\mathbf{s}_2 = \mathbf{s}_1$.
2. $(\mathbf{S}_1^*)^T\mathbf{S}_1 = (N_t - L + 1)\mathbf{I}$ and $\mathbf{S}_2 = \tilde{\mathbf{S}}_1$. That is, identify a sequence $\mathbf{s}_1$ with impulse-like auto-correlation and set $\mathbf{s}_2 = \tilde{\mathbf{s}}_1$.

IV. Simulation results

We investigate the performance achieved by the proposed reduction schemes in terms of BER and MSE.

8-state 8-PSK space-time trellis code

Table I gives sequences $\mathbf{s}_e$ of length $N_t/2 = 13$ which are applicable to the EDGE Typical Urban environment (where $N_t = 26$ and $L = 4$ [10], [11]) and the 8-state 8-PSK space-time trellis code in Section III-B.1. The corresponding $\mathbf{s}_o$ can be calculated from (19). The exhaustive search identified a total of 94 sequences that achieve the minimum MSE, which in this case was 0.0816. The lower bound for MSE achieved by PRUS was 0.0435.

Fig. 7 compares the BER of the scheme in Fig. 3 (and an optimal PRUS sequence) with the performance of the proposed scheme in Fig. 4 (and the first sequence in Table I). The joint space-time equalizer/decoder employs a prefilter to concentrate the channel energy in a smaller number of taps followed by a reduced-complexity Maximum Aposteriori (MAP) equalizer/decoder with M active trellis states as described in [12]. The BER results are for $M = 16$ and $M = 32$. The optimal and the suboptimal training sequences achieve similar performance.

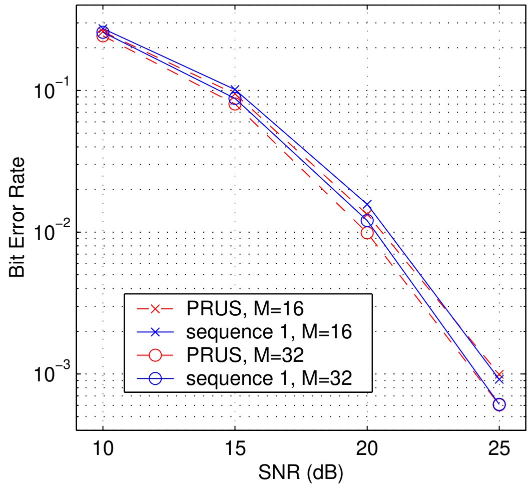

Fig. 7. Performance for the optimal PRUS and a proposed suboptimal training sequence. Sequence 1 refers to the first sequence in Table I, denoted in boldface.

TABLE I

TRAINING SEQUENCES FOR THE EDGE TU ENVIRONMENT ($N_t = 26$, $L = 4$) WITH THE 8-STATE 8-PSK SPACE-TIME TRELLIS CODE IN [17]. THE CORRESPONDING SEQUENCES $\mathbf{s}_o$ ARE CREATED ACCORDING TO (19).

sequence #	a	$\mathbf{s}_e$
1	$e^{\frac{5\pi}{4}}$	**-1 1 1 1 1 -1 -i -1 1 1 -1 1 1**
2	$e^{i\frac{3\pi}{4}}$	1 1 -1 1 i i 1 -i i -1 -1 -1 1
3	$e^{i\frac{\pi}{4}}$	1 -1 -1 -i i -i 1 1 1 -i -1 1 1
4	$e^{i\frac{\pi}{4}}$	1 -1 -1 -i 1 -1 1 -i -i -i -1 1 1
5	$e^{i\frac{\pi}{4}}$	1 i 1 1 i -1 -1 i 1 -1 1 i 1
6	$e^{i\frac{3\pi}{4}}$	1 i 1 i -1 -1 1 -1 -1 i 1 i 1
7	$e^{i\frac{7\pi}{4}}$	1 -i 1 1 -i -1 -1 -i 1 -1 1 -i 1
8	$e^{i\frac{5\pi}{4}}$	1 -i 1 -1 i 1 -1 -i 1 1 1 -i 1
9	$e^{i\frac{3\pi}{4}}$	-1 1 1 1 -1 -1 -i -1 -1 1 -1 1 1
10	$e^{i\frac{7\pi}{4}}$	-1 i -1 -i 1 -i i i 1 i 1 -i 1
11	$e^{i\frac{\pi}{4}}$	-1 -i -1 i 1 i -i -i 1 -i 1 i 1
12	$e^{i\frac{3\pi}{4}}$	-1 -i -1 i -1 i -i -i -1 -i 1 i 1

Space time block code

We now examine the MSE achieved, when we encode the training symbols with the block code in Section III-B.2. For the EDGE TU environment, the lower bound on MSE when using two consecutive training sequences to estimate the channels $\mathbf{h}_1$ and $\mathbf{h}_2$ is 0.1739. The second of the previously proposed training sequence choices ($\mathbf{s}_2 = \tilde{\mathbf{s}}_1$) leads to an MSE of 0.175, which is very close to MMSE.

BER results for space-time block codes with channel estimation on frequency-selective channels were given in [13].

V. CONCLUSIONS

This paper proposed training methods for systems employing multiple transmit antennas over frequency-selective channels.

One proposed method uses a space-time block code to simplify the channel estimation problem from designing multiple training sequences with impulse-like auto-correlation and zero cross-correlation to designing a single training sequence with impulse-like auto-correlation. A second method estimates the overall single-input single-output channel coefficients that incorporate the space time code. Simulation results show performance with training symbols belonging to a standard alphabet such as BPSK, QPSK, and 8-PSK are very close to the optimal PRUS channel estimation.

REFERENCES

[1] P. Spasojevic and C. Georghiades. Complementary sequences for ISI channel estimation. *IEEE Transactions on Info. Theory*, 47(3):1145–1152, March 2001.

[2] W. H. Mow. Sequence design for spread spectrum. *The Chinses Univerity press, Chinese University of Hong Kong*, 1995.

[3] S. N. Crozier, D. D. Falconer, and S. A. Mahmoud. Least sum of squared errors (LSSE) channel estimation. *IEEE Proc.-F*, 138(4):371–378, August 1992.

[4] A. Furuskar, S. Mazur, F. Muller, and H. Olofsson. EDGE: Enhanced Data Rates for GSM and TDMA/136 Evolution. *IEEE Personal Communications*, pages 56–66, June 1999.

[5] T. Kailath, A. H. Sayed, and B. Hassibi. Linear Estimation. *Prentice Hall*, 2000.

[6] G. Caire and U. Mitra. Training sequence design for adaptive equalization of multi-user systems. *Thirty-Second Asilomar Conference*, 2:1479–1483, 1998.

[7] V. Tarokh, N. Seshadri, and A. R. Calderbank. Space-time codes for high data rate wireless communications: performance criterion and code construction. *IEEE Transactions on Info. Theory*, 44(2):744–765, March 1998.

[8] E. Lindskog and A. Paulraj. A transmit diversity scheme for channels with intersymbol interference. *ICC*, 1:307 –311, 2000.

[9] S. Alamouti. A simple transmit diversity technique for wireless communications. *IEEE Journal on Selected Areas in Communications*, 16(8):1451–1458, October 1998.

[10] GSM-0.5.02. Digital cellular telecommunications system (Phase 2+):Multiplexing and Multiple access on the radio path. *ETSI, July 2000*, version 8.5.0 Release 1999.

[11] GSM-0.5.04. Digital cellular telecommunications system (Phase 2+):Modulation. *ETSI, De. 1999*, version 8.1.0 Release 1999.

[12] C. Fragouli, N. Al-Dhahir, S. N. Diggavi, and W. Turin. Prefiltered M-BCJR algorithm for MIMO frequency-selective channels. *CISS*, pages 841–845, March 2001.

[13] N. Al-Dhahir. Single-carrier frequency-domain equalization for space-time-coded transmissions over broadband wireless channels. *Personal, Indoor and Mobile Radio Communications, 2001 12th IEEE International Symposium on*, 1:143–146, September 2001.