MMSE-OPTIMAL TRAINING SEQUENCES FOR SPECTRALLY-EFFICIENT MULTI-USER MIMO-OFDM SYSTEMS

Yuejie Chi[1], *Ahmad Gomaa*[2], *Naofal Al-Dhahir*[2], *and Robert Calderbank*[1,3]

[1]Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA
[2]Department of Electrical Engineering, University of Texas at Dallas, Richardson, TX 75083, USA
[3]Department of Computer Science, Duke University, Durham, NC 27708, USA

ABSTRACT

This paper proposes a new family of optimal training sequences in terms of minimizing the mean-square channel estimation error for spectrally-efficient Multi-User MIMO-OFDM systems with an arbitrary number of transmit antennas and an arbitrary number of training symbols. It addresses uplink transmission scenarios where the users overlap in time and frequency and are separated using spatial processing at the base station. In particular, optimal training sequences can be obtained easily from standard signal constellations such as QPSK with desired low PAPR, making it appealing for practical use.

1. INTRODUCTION

Orthogonal Frequency Division Multiplexing (OFDM) [1] is widely adopted in broadband communications standards for its efficient implementation, high spectral efficiency, and robustness to Inter-Symbol Interference (ISI). OFDM in combination with Multiple-Input-Multiple-Output (MIMO) can support multiple users by assigning each time/frequency slot to one user [2]. This combination greatly increases throughput. In OFDMA systems which are adopted by WiMAX and LTE standards, different users are assigned different subcarriers within the same OFDMA symbol. A different method of separating users is through the random-access CSMA/CA Medium Access Control (MAC) protocol used in WLAN standards, e.g. IEEE 802.11n. Both methods require that users do not overlap in either time or frequency and this restriction results in a significant loss in spectral efficiency. The introduction of multiple receive antennas at the base station makes it possible to improve spectral efficiency by allowing users to overlap while maintaining decodability, as in the recently-proposed Coordinated MultiPoint transmission (CoMP) techniques in the LTE-Advanced standard [3].

Accurate Channel State Information (CSI) is required at the receiver for coherent detection and is typically acquired by sending known training sequences from the transmit antennas and inferring channel parameters from the received signals. It is more challenging in MIMO-OFDM systems because there are more link parameters to calculate, and their estimation is complicated by interference between different transmissions. The direct approach is to invert a large matrix that describes cross-antenna interference at each OFDM tone [4]. Complexity can be reduced by exploiting the correlation between adjacent subchannels [5].

The work of Y. Chi and R. Calderbank was supported by ONR under Grant N00014-08-1-1110, by AFOSR under Grant FA 9550-09-1-0643, and by NSF under Grants NSF CCF-0915299 and NSF CCF-1017431. The work of N.Al-Dhahir and A. Gomaa was supported by QNRF.

Linear Least-Squares (LLS) channel estimation is of great practical importance since it does not require prior knowledge of the channel statistics and enjoys low implementation complexity. We consider frequency-selective block-fading channels where Time Domain (TD) representation usually requires much fewer parameters than Frequency Domain (FD) representation. The design of optimal training sequences for single-user MIMO-OFDM systems is investigated in [6] and [7]. The construction of optimal training sequences for multi-user MIMO-OFDM systems has been studied in both TD [8] and FD [9], but these designs do not directly extend to multiple OFDM training symbols. The unitary filter bank developed from Instantaneous Radar Polarimetry [10] supports FD LLS channel estimation in a 2×2 MIMO-OFDM system [11] and is able to suppress interference over two OFDM symbols with linear complexity. A limitation of this method is that the number of OFDM training symbols required is a power of 2 and at least the number of transmit antennas.

We focus on the design of training sequences for Multi-User MIMO OFDM systems that minimize the Mean Squared Error (MSE) of time-domain LLS channel estimation. Our framework supports the design of optimal training sequences for an arbitrary number of transmit antennas and an arbitrary number of training symbols. It provides a general family of MMSE-optimal training sequences for Multi-User MIMO-OFDM systems where Spatial Division Multiple Access (SDMA) is employed to increase the spectral efficiency. The optimality of our designs holds irrespective of the number of transmit antennas per user, the number of OFDM sub-carriers, the channel delay spread, and the number of users *provided* that the number of tones dedicated to estimation exceeds the product of the number of transmit antennas and the worst case delay spread. Moreover, individual training sequences from standard signal constellations with low Peak-to-Average Power Ratio (PAPR) can be derived from our design, making it very attractive from implementation perspectives.

The rest of this paper is organized as follows. The uplink Multi-User MIMO-OFDM communication system model is described in Section 2. The design of optimal training sequences is given separately for one and for multiple training symbol scenarios in Section 3. The properties and possible candidates for training sequences are discussed in Section 4. Simulation results are presented in Section 5. Finally, conclusions are drawn in Section 6.

2. SYSTEM MODEL

We consider the uplink of a Multi-User MIMO-OFDM system where the ith user is equipped with M_i transmit antennas,

$0 \le i \le L-1$, and $L \ge 1$ is the number of users. Therefore, the total number of transmit antennas among all users is given by $M = \sum_{i=0}^{L-1} M_i$.

We assume that the channel is quasi-static and remains constant over K successive OFDM training symbols. The channel from the jth transmit antenna of the ith user to the Base Station (BST) can be represented either in TD or FD. Let the size of Discrete Fourier Transform (DFT) be N, and the Channel Frequency Response (CFR) be

$$\mathbf{H}_{i,j} = [H_{i,j}(0), \cdots, H_{i,j}(N-1)]^T,$$

where $H_{i,j}(k)$ is the frequency response at the kth subcarrier. However, the Channel Impulse Response (CIR) in TD is represented by a much smaller number of parameters. We assume that the maximal memory over all CIRs is $\nu_{\max}$, and write the CIR as

$$\mathbf{h}_{i,j} = [h_{i,j}(0), \cdots, h_{i,j}(\nu_{\max})]^T.$$

Let $\nu = \nu_{\max} + 1$. Estimating the CIR instead of the CFR leads to the reduction of the number of unknowns from MN to $M\nu$. Hence, a more accurate channel estimate is attainable using the same amount of training. Furthermore, the CFR can be reconstructed from the CIR as follows

$$H_{i,j}(k) = \frac{1}{\sqrt{N}} \sum_{t=0}^{\nu_{\max}} h_{i,j}(t) e^{-j\frac{2\pi}{N}tk}. \tag{1}$$

At the jth ($0 \le j \le M_i - 1$) transmit antenna of the ith user, an OFDM symbol $\mathbf{X}_{i,j}$ of size N is given by

$$\mathbf{X}_{i,j} = [X_{i,j}(0), \cdots, X_{i,j}(N-1)]^T.$$

A Cyclic-Prefix (CP) of length L_p is used for the guard interval in the OFDM system where L_p is chosen to be greater than the channel memory, and all users are assumed to be synchronized with the BST. The received OFDM symbol $\mathbf{Y} = [Y(0), \cdots, Y(N-1)]^T$ at the BST in one symbol time can be written as

$$\mathbf{Y} = \sum_{i=0}^{L-1} \sum_{j=0}^{M_i-1} \mathrm{diag}(\mathbf{H}_{i,j}) \mathbf{X}_{i,j} + \mathbf{N}, \tag{2}$$

where $\mathbf{N} \sim \mathcal{N}(\mathbf{0}_{N\times 1}, \sigma^2 \mathbf{I}_N)$ is assumed to be Additive White Gaussian Noise (AWGN). We consider the mapping $(i,j) \mapsto m : m = \sum_{s=0}^{i} M_s + j - M_i$, $0 \le m \le M-1$, and relabel $\mathbf{H}_{i,j}$ and $\mathbf{X}_{i,j}$ as $\mathbf{H}_m$ and $\mathbf{X}_m$, respectively. Then, equation (2) can be written as

$$\mathbf{Y} = \sum_{m=0}^{M-1} \mathrm{diag}(\mathbf{H}_m) \mathbf{X}_m + \mathbf{N}. \tag{3}$$

The mapping is used by the BST in the assignment of training sequences for each user.

3. MMSE-OPTIMAL TRAINING SEQUENCES

3.1 One OFDM Training Symbol

We apply the IDFT of size N to Eq. (3), and get

$$\mathbf{y} = \sum_{m=0}^{M-1} \mathbf{S}_m \mathbf{h}_m + \mathbf{n} \triangleq \mathbf{S}\mathbf{h} + \mathbf{n} \tag{4}$$

where $\mathbf{y} \in \mathbb{C}^N$, $\mathbf{h} = [\mathbf{h}_0^H, \cdots, \mathbf{h}_{M-1}^H]^H$, $\mathbf{h}_m \in \mathbb{C}^\nu$, and $\mathbf{S} = [\mathbf{S}_0, \cdots, \mathbf{S}_{M-1}] \in \mathbb{C}^{N\times M\nu}$, $\mathbf{S}_m \in \mathbb{C}^{N\nu}$ is the circulant training matrix constructed from the corresponding training sequence transmitted over the mth antenna, $0 \le m \le M-1$. Let $\mathbf{F} = [\mathbf{f}_0, \cdots, \mathbf{f}_{N-1}]$ be the DFT matrix of size N with $\mathbf{f}_i$ denoting its ith column, and let $\mathbf{F}_0 = [\mathbf{f}_0, \cdots, \mathbf{f}_{\nu_{\max}}]$ be composed of the first ν columns of $\mathbf{F}$. Then, $\mathbf{S}_m$ can be written as

$$\mathbf{S}_m = \mathbf{F}^H \mathbf{D}_m \mathbf{F}_0, \tag{5}$$

where $\mathbf{D}_m = \mathrm{diag}\,(X_m(0), \cdots, X_m(N-1))$. To enable LLS channel estimation, the following condition on dimensionality of the matrix $\mathbf{S}$ has to be satisfied [12]

$$N \ge M\nu, \quad \text{or alternatively,} \quad M \le \frac{N}{\nu}. \tag{6}$$

To minimize the MSE of the channel estimation error, the matrix $\mathbf{S}$ is required to satisfy [6]

$$\mathbf{S}^H \mathbf{S} = c\mathbf{I}_{M\nu} \tag{7}$$

and this requires that

$$\mathbf{S}_m^H \mathbf{S}_n = c\delta_{mn}\mathbf{I}_\nu, \quad 0 \le m,n \le M-1. \tag{8}$$

Given (5), the optimality condition becomes

$$\mathbf{F}_0^H \mathbf{D}_m^H \mathbf{D}_n \mathbf{F}_0 = c\delta_{mn}\mathbf{I}_\nu, \quad 0 \le m,n \le M-1. \tag{9}$$

Now we present a general approach which gives a family of optimal training sequences. As a starting point, we choose the FD training sequence as an arbitrary constant-amplitude sequence $\mathbf{X}$. Define

$$\mathbf{D} = \mathrm{diag}\,(X(0), \cdots, X(N-1)), \tag{10}$$

then $\mathbf{D}^H\mathbf{D} = c\mathbf{I}_N$ where c is determined by the signal constellation and/or transmit power constraints. The FD training sequence at the mth transmit antenna is given by

$$\mathbf{X}_m = \mathbf{\Lambda}_m \mathbf{X}, \ 0 \le m \le M-1. \tag{11}$$

where

$$\mathbf{\Lambda}_m = \mathrm{diag}\left(1, e^{j\frac{2\pi\nu}{N}m}, \cdots, e^{j\frac{2\pi\nu(N-1)}{N}m}\right). \tag{12}$$

Equivalently, $\mathbf{D}_m = \mathbf{\Lambda}_m\mathbf{D} = \mathbf{D}\mathbf{\Lambda}_m$, $0 \le m \le M-1$.

Let $\mathbf{F}_m$ be composed of ν consecutive columns of $\mathbf{F}$ starting at index $m\nu$, i.e.

$$\mathbf{F}_m = \left[\mathbf{f}_{m\nu}, \cdots, \mathbf{f}_{(m+1)\nu-1}\right] = \mathbf{\Lambda}_m \mathbf{F}_0. \tag{13}$$

It is straightforward that $\mathbf{F}_m^H \mathbf{F}_n = \delta_{mn}\mathbf{I}_\nu$. We have the following theorem.

Theorem 1. *The choice of FD training sequences in Eq.* (11) *is optimal for a single training OFDM symbol.*

Proof. Since $\mathbf{D}_n\mathbf{F}_0 = \mathbf{\Lambda}_n\mathbf{D}\mathbf{F}_0 = \mathbf{D}\mathbf{\Lambda}_n\mathbf{F}_0 = \mathbf{D}\mathbf{F}_n$, it follows

$$\mathbf{F}_0^H \mathbf{D}_m^H \mathbf{D}_n \mathbf{F}_0 = \mathbf{F}_m^H \mathbf{D}^H \mathbf{D} \mathbf{F}_n = c\mathbf{F}_m^H \mathbf{F}_n = c\delta_{mn}\mathbf{I}_\nu.$$

Therefore Eq. (9) holds. □

Therefore, the least-square estimates of $\mathbf{h}$ and $\mathbf{h}_m$ are given as

$$\hat{\mathbf{h}} = \frac{1}{c}\mathbf{S}^H\mathbf{y}, \quad \text{and} \quad \hat{\mathbf{h}}_m = \frac{1}{c}\mathbf{S}_m^H\mathbf{y}.$$

Then, the CFR estimate is given by

$$\hat{\mathbf{H}}_m = \frac{1}{c}\mathbf{F}\mathbf{S}_m^H\mathbf{y} = \frac{1}{c}(\mathbf{F}\mathbf{F}_0^H)\mathbf{D}_m^H\mathbf{F}\mathbf{y}, \tag{14}$$

and the MMSE of TD channel estimation is given by

$$\sigma_e^2 = \sigma^2 \mathrm{Tr}\left((\mathbf{S}^H\mathbf{S})^{-1}\right) = \frac{M\nu}{c}\sigma^2. \tag{15}$$

3.2 K OFDM Training Symbols with $K \geq 2$

The major limitation of using only one training OFDM symbol is that the total number of transmit antennas is limited by $\frac{N}{\nu}$. When the channel is quasi-static over $K \geq 2$ OFDM training symbols it is possible to increase the number of transmit antennas and reduce MMSE by a factor of K.

Denoting the received TD OFDM symbol in the tth symbol time by $\mathbf{y}_t$, $0 \leq t \leq K-1$, we write

$$\begin{aligned}\begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{K-1} \end{bmatrix} &= \begin{bmatrix} \mathbf{S}_{00} & \mathbf{S}_{01} & \cdots & \mathbf{S}_{0,M-1} \\ \mathbf{S}_{10} & \mathbf{S}_{11} & \cdots & \mathbf{S}_{1,M-1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{S}_{K-1,0} & \mathbf{S}_{K-1,1} & \cdots & \mathbf{S}_{K-1,M-1} \end{bmatrix} \mathbf{h} + \mathbf{n} \\ &\triangleq \begin{bmatrix} \mathbf{S}_0 \\ \mathbf{S}_1 \\ \vdots \\ \mathbf{S}_{K-1} \end{bmatrix} \mathbf{h} + \mathbf{n} = \mathbf{S}\mathbf{h} + \mathbf{n}, \end{aligned} \tag{16}$$

where $\mathbf{S}_{tm} = \mathbf{F}^H \mathbf{D}_{tm} \mathbf{F}_0$, and the matrices $\mathbf{D}_{tm}$'s are diagonal matrices with the FD training sequences of the mth transmit antenna at the tth OFDM training symbol on their diagonals. Let $\mathbf{S}_t = [\mathbf{S}_{t1}, \cdots, \mathbf{S}_{t,M-1}]$ be the training matrix at the tth OFDM training symbol. Least-square estimation is possible when the following dimensionality condition for $\mathbf{S} \in \mathbb{C}^{KN \times M\nu}$ holds

$$KN \geq M\nu, \quad \text{or equivalently,} \quad M \leq \frac{KN}{\nu}. \tag{17}$$

The training matrix is required to satisfy

$$\mathbf{S}^H \mathbf{S} = \tilde{c} \mathbf{I}_{M\nu}$$

for some $\tilde{c}$ to be MMSE optimal [6]. We propose the following design of optimal training sequences. For $0 \leq m \leq M-1$, let

$$p = \left\lfloor \frac{m}{K} \right\rfloor \in \left\{ 0, \cdots, \left\lfloor \frac{M-1}{K} \right\rfloor \right\}, \tag{18}$$

$$q = m - Kp \in \{0, \cdots, K-1\}. \tag{19}$$

Let $\mathbf{U} = [U_{tq}] \in \mathbb{C}^{K \times K}$ be a scalar unitary matrix with unitary entries satisfying $|U_{tq}| = 1$ and $\mathbf{U}^H \mathbf{U} = K\mathbf{I}_K$. Possible candidates are a DFT matrix of size K or a perfect Space-Time Block Code (STBC) matrix. Let $\mathbf{X}$ and $\mathbf{D}$ be defined in (10), for the mth transmit antenna, its FD training sequence at the tth OFDM training symbol is given by

$$\mathbf{X}_{tm} = U_{tq} \mathbf{\Lambda}_p \mathbf{X}, \quad \text{if } m = Kp + q,\ 0 \leq m \leq M-1, \tag{20}$$

where $\mathbf{D}_{tm} = \text{diag}\{\mathbf{X}_{tm}(0), \cdots, \mathbf{X}_{tm}(N-1)\} = U_{tq} \mathbf{\Lambda}_p \mathbf{D}$ and $\mathbf{\Lambda}_p$ are defined in (12).

Let the matrix $\mathbf{\Sigma} \in \mathbb{C}^{KN \times KN}$ be constructed as a Kronecker product $\mathbf{\Sigma} = \mathbf{U} \otimes \mathbf{D}$, where $\mathbf{\Sigma}_{tq} = U_{tq} \mathbf{D}$ is the $N \times N$ diagonal matrix located at the (t,q) block of $\mathbf{\Sigma}$. Therefore the matrix $\mathbf{\Sigma}$ satisfies

$$\mathbf{\Sigma}^H \mathbf{\Sigma} = \mathbf{U}^H \mathbf{U} \otimes \mathbf{D}^H \mathbf{D} = \tilde{c} \mathbf{I}_{KN}, \tag{21}$$

where $\tilde{c} = Kc$. We have the following optimality result.

Theorem 2. *The training sequences in* (20) *are optimal for* $K \geq 2$ *training OFDM symbols.*

Proof. It is enough to show that

$$\sum_{t=0}^{K-1} \mathbf{S}_{tm}^H \mathbf{S}_{tn} = \mathbf{F}_0^H \left(\sum_{t=0}^{K-1} \mathbf{D}_{tm}^H \mathbf{D}_{tn} \right) \mathbf{F}_0 = \tilde{c} \delta_{mn} \mathbf{I}_{(\nu_{\max}+1)} \tag{22}$$

It is obvious that when $m = n$, the above Eq. (22) holds. When $m \neq n$, we write $m = Kp_1 + q_1$ and $n = Kp_2 + q_2$ and split the proof into two cases:

- $q_1 = q_2 = q \in \{0, \cdots, K-1\}$ but $p_1 \neq p_2$. Then,

$$\begin{aligned}\mathbf{S}_{tm}^H \mathbf{S}_{tn} &= \mathbf{F}_{p_1}^H \mathbf{\Sigma}_{tq}^H \mathbf{\Sigma}_{tq} \mathbf{F}_{p_2} \\ &= c|U_{tq}|^2 \mathbf{F}_{p_1}^H \mathbf{F}_{p_2} = \mathbf{0}_\nu.\end{aligned}$$

- $q_1 \neq q_2$. Then we have

$$\sum_{t=0}^{K-1} \mathbf{D}_{tm}^H \mathbf{D}_{tn} = \mathbf{\Lambda}_{p_1}^H \left(\sum_{t=0}^{K-1} \mathbf{\Sigma}_{t,q_1}^H \mathbf{\Sigma}_{t,q_2} \right) \mathbf{\Lambda}_{p_2} = \mathbf{0}_N.$$

Now, Eq. (22) follows trivially. □

The bijection $\pi : m \mapsto \{p, q\}$ groups the antennas into K classes depending on the equivalence of the residue q. For two antennas *not* in the same class, the orthogonality between their training sequences can be proved over *any* OFDM training symbol similar to Theorem 1. For two antennas in the same class, their training sequences can be proved orthogonal over *all* K OFDM training symbols. We give the detailed proof below.

Finally, the least-square estimate of $\mathbf{h}$ and $\mathbf{h}_m$ are given by

$$\hat{\mathbf{h}} = \frac{1}{\tilde{c}} \sum_{t=0}^{K-1} \mathbf{S}_t^H \mathbf{y}_t, \quad \text{and} \quad \hat{\mathbf{h}}_m = \frac{1}{\tilde{c}} \sum_{t=0}^{K-1} \mathbf{S}_{tm}^H \mathbf{y}_t.$$

Then, the CFR is given by

$$\hat{\mathbf{H}}_m = \frac{1}{\tilde{c}} \mathbf{F} \hat{\mathbf{h}}_m = \frac{1}{\tilde{c}} (\mathbf{F}\mathbf{F}_0^H) \sum_{t=0}^{K-1} \mathbf{D}_{tm}^H \mathbf{F} \mathbf{y}_t.$$

The resulting MMSE of TD channel estimation is improved by a factor of K, given by

$$\sigma_e^2 = \sigma^2 \text{Tr} \left(\sum_{t=0}^{K-1} \mathbf{S}_t^H \mathbf{S}_t \right)^{-1} = \frac{M(\nu_{\max}+1)}{Kc} \sigma^2.$$

3.3 The case $K = 2$

When K=2, there is a special construction using Hamilton's Biquaternions that is similar in spirit to the Alamouti STBC [13]. We will choose two FD training sequences $\mathbf{X}$ and $\mathbf{Z}$ where the sum of their squared amplitudes is constant, i.e.

$$\mathbf{D}_X^H \mathbf{D}_X + \mathbf{D}_Z^H \mathbf{D}_Z = 2c\mathbf{I}_N. \tag{23}$$

where $\mathbf{D}_X = \text{diag}\,(X(0), \cdots, X(N-1))$, and $\mathbf{D}_Z = \text{diag}\,(Z(0), \cdots, Z(N-1))$.

For $0 \leq m \leq M-1$, let $p = \left\lfloor \frac{m}{2} \right\rfloor$, $0 \leq p \leq \left\lfloor \frac{M-1}{2} \right\rfloor$ and $q = m - 2p \in \{0, 1\}$. Let $\mathbf{X}_p = \mathbf{\Lambda}_p \mathbf{X}$ and $\mathbf{Z}_p = \mathbf{\Lambda}_p \mathbf{Z}$, $0 \leq p \leq \left\lfloor \frac{M-1}{2} \right\rfloor$, where $\mathbf{\Lambda}_p$ is defined in Eq. (12).

The diagonal FD training matrices of the mth antenna in the 0th and 1st training symbols are given by $\mathbf{D}_{0m}$ and $\mathbf{D}_{1m}$ respectively:

$$\mathbf{D}_{0m} = \begin{cases} \mathbf{\Lambda}_p \mathbf{D}_X, & \text{if } q=0, m=2p \\ \mathbf{\Lambda}_p^H \mathbf{D}_Z^H, & \text{if } q=1, m=2p+1 \end{cases},$$

$$\text{and} \quad \mathbf{D}_{1m} = \begin{cases} \mathbf{\Lambda}_p \mathbf{D}_Z, & \text{if } q=0, m=2p \\ -\mathbf{\Lambda}_p^H \mathbf{D}_X^H, & \text{if } q=1, m=2p+1 \end{cases}. \quad (24)$$

Without proof we state the following optimality result.

Theorem 3. *The FD training sequences in Eq.* (24) *are optimal for two training OFDM symbols.*

If all the users employ two transmit antennas and Alamouti STBC, their training sequences in two symbol intervals are assigned according to Eq. (24), which can be generated simply using the same Alamouti code generator, greatly reducing the training assignment complexity.

4. DISCUSSIONS

4.1 PAPR of training sequences

The PAPR of a training sequence $S(n)$, $0 \le n \le N-1$, is given by

$$\text{PAPR}(S) = \frac{\max_n |S(n)|^2}{\frac{1}{N}\sum_{n=0}^{N-1} |S(n)|^2}. \quad (25)$$

The transform operator $\mathbf{\Lambda}_m$ between different FD training sequences can be viewed as a frequency modulation operator, which is equivalent to circulant shift of the training sequence in the TD. Hence, we have the following proposition.

Proposition 4. *All TD training sequences in* (11) *and* (20) *have the same PAPR.*

This property is important when designing the training sequences. As long as the PAPR of $\mathbf{X}$ is low, all training symbols will have same low PAPR. Another merit of our design is that if we choose ν such that $\frac{N}{\nu} = 2^k$ for some integer k, and choose $\mathbf{X}$ from a 2^k-Phase Shift Keying (PSK) constellation, then the transform $\mathbf{\Lambda}_m$ guarantees that all FD training sequences $\{\mathbf{X}_m,\ 0 \le m \le M-1\}$ belong to the same 2^k-PSK constellation, which is very convenient to generate.

4.2 Candidates

Once the initial sequence $\mathbf{X} = \{X(k)\}_{k=0}^{N-1}$ is chosen, the family of training sequences is completely determined. One possibility for $\mathbf{X}$ is a Constant-Amplitude-Zero-Auto-Correlation (CAZAC) sequence [14] with optimal PAPR (0dB), such as the chirp sequence given below:

$$X(k) = \begin{cases} \sqrt{c} e^{j\frac{\pi u k^2}{N}}, & \text{if } N \text{ is even} \\ \sqrt{c} e^{j\frac{\pi u k(k+1)}{N}}, & \text{if } N \text{ is odd} \end{cases}, \ 0 \le k \le N-1. \quad (26)$$

where u is any integer relatively prime to N. This sequence is also known as the Zardoff-Chu sequence, which is widely adopted in the LTE systems. A disadvantage of this and other CAZAC sequences is that the entries are not restricted to a standard signal constellation. An alternative is provided by Golay complementary sequences [15] which only assume values from $\{-\sqrt{c}, \sqrt{c}\}$ with a PAPR smaller than 3dB. A third possibility is the flat sequence (TD impulsive) $\{\mathbf{X} :\ X(k) = \sqrt{c},\ \text{for all } k\}$ with unbounded PAPR. The performance of these three sets of training sequences are evaluated in Section 5 by simulation.

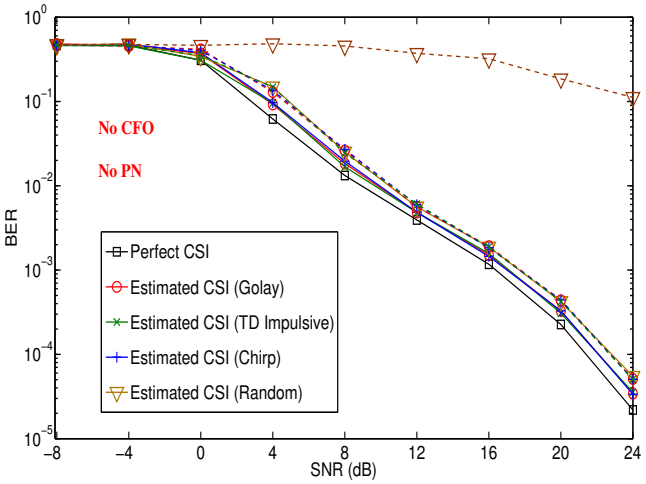

Figure 1: BER versus SNR for $K = 1$ (dashed) and 2 (solid) training OFDM symbols from optimal training sequences (generated from Golay, TD impulsive and chirp) and non-optimal random sequence compared with perfect CSI.

5. NUMERICAL RESULTS

We consider uplink transmission in a Multi-User MIMO-OFDM system with N = 64 and $\nu_{\max}$ = 15. Each BST is equipped with two co-located receive antennas and two users are each equipped with two transmit antennas over which the Alamouti STBC is employed. Each user employs a non-systematic rate-1/2 convolutional code with octal generator $(133, 171)$ and constraint length 7, which are further QPSK modulated. All channel paths are assumed to have uncorrelated and identically-distributed CIRs with 8 zero-mean complex Gaussian taps following an exponentially-decaying power delay profile with a 3 dB decay per tap. K OFDM training symbols are transmitted over each transmit antenna for channel estimation as described in Section 3. The CIR estimates are used for detection of the OFDM data symbols through the joint Linear Minimum-Mean-Square-Error (LMMSE) technique [16] where the received signals from the two receive antennas are processed jointly to separate the two users.

Using these parameters, the dimensionality condition in (17) is met with $K \ge 1$. In Fig. 1, the Bit Error Rate (BER) performances of three FD training sequences proposed in Section 3.4 (namely: Chirp, Golay, and TD Impulsive) with $K = 1$ and 2 are compared with the perfect CSI case. All training sequences can be generated from standard QPSK constellation except chirp sequences. In Fig. 1, all users are assumed to have perfect frequency synchronization with the receiver. All training sequences achieve roughly the same BER performance with SNR losses of 1.5 and 0.7 dB for $K = 1$ and 2, respectively compared with the perfect CSI case. The performance of a random BPSK sequence *not* satisfying the optimality condition is also shown for comparison. The performance of the random sequence is inferior to that of the other sequences satisfying the optimal-

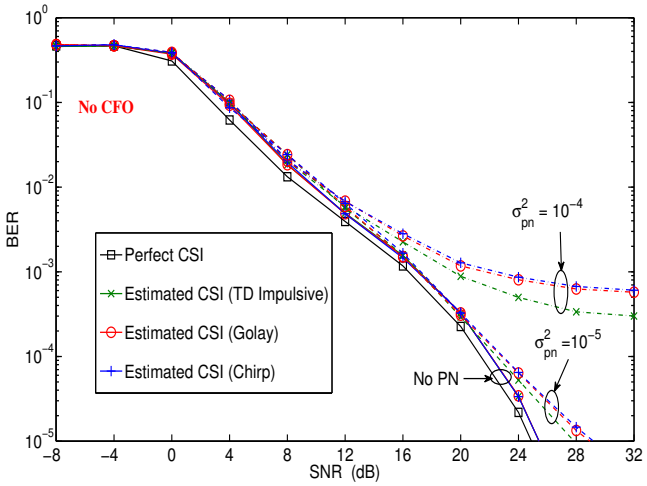

Figure 2: BER versus SNR for different PN levels $\sigma^2_{\text{pn}} = 0$ (solid), 10^{-5} (dashed), and 10^{-4} (dash-dotted) with $K = 2$ training OFDM symbols.

ity condition; especially with $K = 1$ training symbol where the number of equations equals the number of unknowns making the channel estimate unreliable when the optimality condition is not satisfied. From another perspective, our optimally-designed training sequences with $K = 1$ training symbol achieve comparable performance to that of the random sequence with $K = 2$ training symbols, i.e. with 50% less training overhead. This is besides the additional complexity needed to invert the matrix $\mathbf{S}^H\mathbf{S}$ which is not a scaled identity in the case of non-optimal sequences.

We further examine the performance of the optimal training sequences when Phase Noise (PN) is present at the transmitter and receivers in Fig. 2 when its variance $\sigma^2_{pn} = 0$, 10^{-5} and 10^{-4}. Although all the optimal training sequences obtain MMSE, their PAPR and robustness to PN is different. This tradeoff is further explored in [17].

6. CONCLUSIONS

A family of MMSE-optimal training sequences is proposed for spectrally-efficient Multi-User MIMO-OFDM systems with an arbitrary number of transmit antennas and an arbitrary number of training symbols for LLS channel estimation. The capability to sense more channels with optimal MMSE is systematically obtained by modulating the cyclic stucture of the Fourier matrix with entries from a unitary matrix in the multiple OFDM training symbol case. Numerical results confirms the benefits of our design, in particular showing its advantage when the number of channel parameters is close to the number of available equations.

REFERENCES

[1] A. R. S. Bahai, B. R. Saltzberg, and M. Ergen, *Multi-carrier digital communications theory and applications of OFDM*. New York, NY: Springer, 2004.

[2] M. Jiang and L. Hanzo, "Multiuser MIMO-OFDM for next-generation wireless systems," *Proceedings of the IEEE*, vol. 95, no. 7, pp. 1430–1469, Jul. 2007.

[3] S. Parkvall, E. Dahlman, A. Furuskar, Y. Jading, M. Olsson, S. Wanstedt, and K. Zangi, "LTE-advanced - evolving LTE towards IMT-advanced," in *IEEE Vehicular Technology Conference*, 2008.

[4] Y. G. Li, N. Seshadri, and S. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels," *IEEE Journal on Selected Areas in Communications*, vol. 17, pp. 461 – 471, 1999.

[5] H. Minn, D. I. Kim, and V. Bhargava, "A reduced complexity channel estimation for OFDM systems with transmit diversity in mobile wireless channels," *IEEE Transactions on Communications*, 2002.

[6] H. Minn and N. Al-Dhahir, "Optimal training signals for MIMO OFDM channel estimation," *IEEE Trans. Wireless Comm.*, vol. 5, no. 5, pp. 1158–1168, April 2006.

[7] Y. G. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Transactions on Wireless Communications*, vol. 1, pp. 67 – 75, 2002.

[8] Y. Zeng, A. R. Leyman, S. Ma, and T.-S. Ng, "Optimal pilot and fast algorithm for MIMO-OFDM channel estimation," in *Proceeding of IEEE International Conference on Information, Communications and Signal Processing (ICICS)*, 2005.

[9] F. Horlin and L. V. der Perre, "Optimal training sequences for low complexity ML multi-channel estimation in multi-user MIMO OFDM-based communications," in *Proceeding of IEEE International Conference on Communications (ICC)*, vol. 4, Jun. 2004.

[10] S. D. Howard, A. R. Calderbank, and W. Moran, "A simple signal processing architecture for instantaneous radar polarimetry," *IEEE Trans. Inform. Theory*, vol. 53, no. 4, pp. 1282–1289, Apr. 2007.

[11] M. D. Zoltowski, T. R. Qureshi, and R. Calderbank, "Complementary codes based channel estimation for MIMO-OFDM systems," in *Proc. 46th Annual Allerton Conf. Communication, Control, and Computing*, Monticello, IL, Sep. 2008, pp. 133–138.

[12] S. M. Kay, *Fundamentals Of Statistical Signal Processing: Estimation Theory*. New Jersey: Prentice Hall, 1993.

[13] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 1451–1458, 1998.

[14] D. Chu, "Polyphase codes with good periodic correlation properties," *IEEE Trans. on Info. Theory*, vol. IT-18, 1972.

[15] M. J. E. Golay, "Complementary series," *IRE Trans. Inform. Theory*, vol. 7, no. 2, pp. 82–87, April 1961.

[16] A. Naguib, N. Seshadri, and A. Calderbank, "Increasing data rate over wireless channels," *IEEE Signal Processing Magazine*, vol. 17, 2000.

[17] Y. Chi, A. Gomaa, N. Al-Dhahir, and R. Calderbank, "Training signal design and tradeoffs for spectrally-efficient multi-user mimo-ofdm systems," *To Appear, IEEE Transactions on Wireless Communications*, 2011.