MIMO Frequency Selective Channel Estimation Using Aperiodic Complementary Sets of Sequences

Shuangquan Wang and Ali Abdi

Center for Wireless Communications and Signal Processing Research
Department of Electrical and Computer Engineering
New Jersey Institute of Technology, Newark, NJ 07102 Email: {sw27, ali.abdi}@njit.edu

***Abstract*— Accurate estimation of MIMO frequency selective fading channels is important for reliable communication. In this paper, a new channel estimation scheme which relies on uncorrelated aperiodic complementary sets of sequences is proposed. Theoretical analysis and Monte-Carlo simulation show that the estimator achieves the minimum possible Cramér-Rao lower bound (CRLB). Furthermore, low-complexity algorithms for both ASIC/FPGA and DSP implementations are provided, using the special structure of the uncorrelated aperiodic complementary sets of sequences.**

I. Introduction

In general, there are two classes of methods applicable to estimate the channel state information (CSI): blind estimation and training-based estimation. For quasi-static or slowly-varying fading channels, training-based channel estimation at the receiver is common in practice and offers less complexity and better performance, compared to the blind approach. The pilot-aided transmission consumes more bandwidth for transmitting the training symbols, but this overhead is negligible, when compared with long sequence of transmitted data.

Some information theoretical aspects of training sequence design for multiple-input multiple-output (MIMO) flat fading channels are discussed in [1]. Based on a parametric representation of MIMO frequency selective fading channels, a lower bound on estimation error of training-based techniques is derived in [2]. The design of the probing signal for system identification of a single-input single-output (SISO) tapped delay line model is addressed in [3, pp. 90-95]. The optimal training design for MIMO ISI channels is given in [4], and the training symbols are delta pulses.

For MIMO flat fading channels, one can easily design training sequences that satisfy the condition given in [1] and [5]. However, due to the block-Toeplitz structure of the training matrix shown in (4), finding such sequences for MIMO frequency selective channel estimation is a difficult task [5]. To overcome this problem, we propose a two-sided (preamble and postamble) training structure shown in Fig. 1 using uncorrelated aperiodic complementary sets of sequences, based on the SISO method of [6]. In Fig. 1, the gaps are filled by L 0's to separate the data and training symbols.

The rest of the paper is organized as follows. Definition and construction of uncorrelated aperiodic complementary sets of sequences are given in the Sec. II. MIMO system and channel models and the criterion for optimal training design are addressed in Sec. III and IV, respectively. The construction of optimal training symbols is stated in Sec. V. The fast implementation of low hardware and computation complexity is presented in Sec. VI and the simulation results are put in Sec. VII. At last, the conclusion remarks are summarized in Sec. VIII.

Notation: We reserve upper bold letters for matrices, lower bold letters for vectors, sequences or sets, $(\cdot)^T$ for the matrix transpose, $(\cdot)^{-1}$ for the matrix inverse, $\otimes$ for the Kronecker product, $\mathbb{E}(\cdot)$ for the ensemble average, tr$(\cdot)$ for the trace of a matrix, $\mathbf{I}_N$ for the $N \times N$ identity matrix, $\mathbf{0}_{m,n}$ for the $m \times n$ zero matrix, $\|\cdot\|_F$ for the Frobenius norm, $\overline{[\cdot]}$ for the sample average, $\lceil x \rceil$ for the minimum integer greater than or equal to x, $t \in [m, n]$ for the case t is an integer such that $m \leq t \leq n$, and diag$(d_0, d_1, \cdots, d_n)$ for a diagonal matrix, whose diagonal elements are given by $d_0, d_1, \cdots, d_n$.

II. Definition and Construction of Aperiodic Complementary Sets of Sequences

Let $\mathbf{a}_i = \{a_{i,0}, a_{i,1}, \cdots, a_{i,(N-1)}\}$ be a sequence of ± 1's, and $r_{\mathbf{a}_i,\mathbf{a}_i}(k) = \sum_{n=0}^{N-1-|k|} a_{i,n} a_{i,(n+|k|)}$, $|k| \in [0, N-1]$ is the aperiodic autocorrelation of the sequence $\mathbf{a}_i$ [7] [8]. A set of p sequences $\{\mathbf{a}_i\}_{i=0}^{p-1}$ is complementary if and only if $\sum_{i=0}^{p-1} r_{\mathbf{a}_i,\mathbf{a}_i}(k) = 0, k \neq 0$. In this paper, only sequences with the equal length N are considered, therefore, $\sum_{i=0}^{p-1} r_{\mathbf{a}_i,\mathbf{a}_i}(0) = pN$.

If another set of p sequences $\{\mathbf{b}_i\}_{i=0}^{p-1}$ is complementary and $\sum_{i=0}^{p-1} r_{\mathbf{a}_i,\mathbf{b}_i}(k) = 0, \forall k$, where $r_{\mathbf{a}_i,\mathbf{b}_i}(k) = \sum_{n=0}^{N-1-k} a_{i,n} b_{i,(n+k)}$, $k \in [0, N-1]$ and $r_{\mathbf{a}_i,\mathbf{b}_i}(k) = r_{\mathbf{b}_i,\mathbf{a}_i}(-k), k \in [-N+1, -1]$, then we call $\{\mathbf{b}_i\}_{i=0}^{p-1}$ is a mate of $\{\mathbf{a}_i\}_{i=0}^{p-1}$. Of course, $\{\mathbf{a}_i\}_{i=0}^{p-1}$ is also a mate of $\{\mathbf{b}_i\}_{i=0}^{p-1}$, more elaborate discussion can be found in [7].

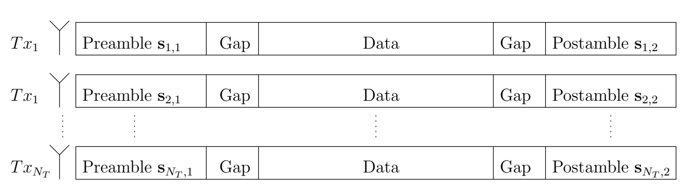

Fig. 1. Frame structure in a MIMO System.

A collection of uncorrelated aperiodic complementary sets of sequences $\{\mathbf{a}_i\}_{i=0}^{p-1}$, $\{\mathbf{b}_i\}_{i=0}^{p-1}$, $\cdots$, $\{\mathbf{z}_i\}_{i=0}^{p-1}$ are *mutually uncorrelated* if every two uncorrelated aperiodic complementary sets of sequences in the collection are mates of each other.

As an example, let $p = 2$. According to property 9) of [9], the aperiodic Golay complementary sequence pair $\{\mathbf{a}_0, \mathbf{a}_1\}$ of length $N = 2^M, M \geqslant 1$, can be constructed by the following recursive method [10]

$$\begin{aligned} a_{0,k}^{(m)} &= a_{0,k}^{(m-1)} + a_{1,(k-D_m)}^{(m-1)} \\ a_{1,k}^{(m)} &= a_{0,k}^{(m-1)} - a_{1,(k-D_m)}^{(m-1)} \end{aligned} \tag{1}$$

with $a_{0,k}^{(0)} = a_{1,k}^{(0)} = \delta_k$, where $\delta_0 = 1, \delta_k = 0, k \neq 0$, $\{D_m, 1 \leq m \leq M\}$ is a permutation of $\{1, 2, \cdots, 2^{M-1}\}$. After M iterations, we get a pair of aperiodic Golay complementary sequences $\mathbf{a}_0$ and $\mathbf{a}_1$, each of length N.

Based on property 3) of [9], $\mathbf{a}_0$ and $\overleftarrow{\mathbf{a}_1}$ are also aperiodic complementary sequences, where $\overleftarrow{\mathbf{b}}$ is the reverse of the sequence $\mathbf{b}$, i.e. $\overleftarrow{b}_k = b_{N-1-k}, k \in [0, N-1]$. Moreover, according to Theorem 11 in [7], if $\{\mathbf{x}, \mathbf{y}\}$ are a aperiodic complementary set of two sequences of length N, then $\{\overleftarrow{\mathbf{y}}, -\overleftarrow{\mathbf{x}}\}$ is its mate. So, $\{\mathbf{a}_1, -\overleftarrow{\mathbf{a}_0}\}$ is the mate of $\{\mathbf{a}_0, \overleftarrow{\mathbf{a}_1}\}$. These properties will be used for optimal training sequences construction in Sec. V.

For example, if $N = 16$ and $D_m = 2^{m-1}$, then $\mathbf{a}_0 = \{+++-++-++++---+-\}$ and $\mathbf{a}_1 = \{+++-++-+----+++-+\}$ are a pair of aperiodic complementary sequences, where $+$ denotes $+1$ and $-$ for -1.

III. System and Channel Models

We consider the block fading model for frequency selective MIMO channels, where the channel matrix is fixed within one block and random changes from one block to another. This is a suitable model for indoor MIMO channels, due to the low mobility [11].

Let $\mathbf{H} = [\mathbf{H}_0 \ \mathbf{H}_1 \ \cdots \ \mathbf{H}_L]$ be the discrete-time channel impulse response (CIR) matrix of the MIMO frequency selective channel, where $\mathbf{H}_l, l \in [0, L]$, is the l^{th} tap of the MIMO CIR matrix, defined by

$$\mathbf{H}_l = \begin{bmatrix} h_{1,1}(l) & \cdots & h_{1,N_T}(l) \\ \vdots & \ddots & \vdots \\ h_{N_R,1}(l) & \cdots & h_{N_R,N_T}(l) \end{bmatrix}, \tag{2}$$

where N_R and N_T are the number of receive and transmit antenna, respectively, $h_{n_r,n_t}(l)$ is the l^{th} tap of the CIR between the n_r^{th} receive antenna and the n_t^{th} transmit antenna. For perfect separating the data and training symbols as well as avoiding inter-frame interference, the gaps between them are filled by L 0's. Using matrix notation, the signals received by N_R antennas, corresponding to the training symbols transmitted from N_T antennas can be written as

$$\mathbf{Y} = \sqrt{\frac{\rho}{N_T}} \mathbf{H}\mathbf{S} + \mathbf{E}, \tag{3}$$

where the $N_T(L+1) \times (N_s + L)$ training matrix $\mathbf{S}$ is defined by

$$\mathbf{S} = \begin{bmatrix} \mathbf{s}(0) & \mathbf{s}(1) & \cdots & \mathbf{s}(N_s-1) & \mathbf{0}_{N_T,1} & \cdots & \mathbf{0}_{N_T,1} \\ \mathbf{0}_{N_T,1} & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \mathbf{0}_{N_T,1} \\ \mathbf{0}_{N_T,1} & \cdots & \mathbf{0}_{N_T,1} & \mathbf{s}(0) & \mathbf{s}(1) & \cdots & \mathbf{s}(N_s-1) \end{bmatrix}, \tag{4}$$

in which $\mathbf{Y} = [\mathbf{y}(0) \ \mathbf{y}(1) \ \cdots \ \mathbf{y}(N_s + L - 1)]$, $\mathbf{E} = [\mathbf{e}(0) \ \mathbf{e}(1) \ \cdots \ \mathbf{e}(N_s + L - 1)]$, $\mathbf{s}(n) = [s_1(n) \ s_2(n) \ \cdots \ s_{N_T}(n)]^T$, $n \in [0, N_s - 1]$, $\mathbf{y}(n) = [y_1(n) \ y_2(n) \ \cdots, y_{N_R}(n)]^T$, $\mathbf{e}(n) = [e_1(n) \ e_2(n) \ \cdots \ e_{N_R}(n)]^T$, $n \in [0, N_s + L - 1]$.

Clearly, $s_{n_t}(n)$ is the training symbol transmitted by the n_t^{th} transmit antenna at time n, $y_{n_r}(n)$ is the symbol received by the n_r^{th} receive antenna at time n, polluted by the additive complex Gaussian noise component $e_{n_r}(n)$ of zero-mean unit power, ρ is the expected received signal-noise-ratio (SNR) at each receive antenna and N_s is the length of the training block. Moreover, $\mathbf{H}$ and $\mathbf{E}$ are independent.

IV. Criterion for Optimal Training Design

In this paper, we consider the total mean square error (TMSE) of the estimation only as the optimal design criterion. In addition, we treat the channel matrix $\mathbf{H}$ in two different ways. In the first approach, $\mathbf{H}$ is an unknown deterministic matrix. In the second setup, $\mathbf{H}$ is random and independent of the additive spatio-temporal white Gaussian noise $\mathbf{E}$, elements of $\mathbf{H}$ are Gaussian and independent with zero mean, and each subchannel $\mathbf{h}_{n_r,n_t} = [h_{n_r,n_t}(0), h_{n_r,n_t}(1), \cdots, h_{n_r,n_t}(L)]$ has unit power, i.e., $\sum_{l=0}^{L} \mathbb{E}(|h_{n_r,n_t}(l)|^2) = 1$. Moreover, the l^{th} taps of all the subchannels have the same power P_l, i.e., $\mathbb{E}(|h_{n_r,n_t}(l)|^2) = P_l, l \in [0, L], \forall n_r, n_t$. Obviously $\sum_{l=0}^{L} P_l = 1$. We also define $\mathbf{C_q} = \mathbb{E}(\mathbf{h}'_{n_r,n_t} \mathbf{h}^\star_{n_r,n_t})$ as the covariance matrix among the $L+1$ taps between the n_t^{th} transmit and n_r^{th} receive antennas, given by $\mathbf{C_q} = \text{diag}(P_0, P_1, \cdots, P_L), \forall n_r, n_t$.

Following the terminology of [3], the best estimators for deterministic and random channel representations are the maximum likelihood estimator (MLE) and the minimum mean square error (MMSE) estimator, respectively, which are presented in the following proposition.

Proposition 1: For the system model in (3), the MLE and MMSE estimator of $\mathbf{H}$ are given by[1]

$$\widehat{\mathbf{H}} = \sqrt{\frac{N_T}{\rho}} \mathbf{Y}\mathbf{S}^T \left(\mathbf{S}\mathbf{S}^T + \alpha \frac{N_T}{\rho} \mathbf{C}_\mathbf{q}^{-1} \otimes \mathbf{I}_{N_T} \right)^{-1}, \tag{5}$$

with the following TMSE

$$\varepsilon_\alpha = \frac{N_R N_T}{\rho} \text{tr} \left[\left(\mathbf{S}\mathbf{S}^T + \alpha \frac{N_T}{\rho} \mathbf{C}_\mathbf{q}^{-1} \otimes \mathbf{I}_{N_T} \right)^{-1} \right], \tag{6}$$

where $\alpha = 0$ and 1 correspond to the MLE and MMSE estimator, respectively. Furthermore, ε_0 and ε_1 are the classical

[1] To obtain a meaningful estimate of $\mathbf{H}$, we need at least as many measurements as unknowns [1], which implies $N_s + L \geqslant N_T(L+1)$.

CRLB and Bayesian CRLB[2], respectively, since elements of $\mathbf{E}$ and $\mathbf{H}$ are Gaussian and independent.

For $\alpha = 0$ in (5) and (6), they correspond to MLE, and are the same as the least square estimation (LSE), given in [5]. With $\alpha = 1$ in (5) and (6), we get the MMSE estimator. They can be easily proved by using Theorem 11.1 in [3], combined with the equality $\text{vec}(\mathbf{AZB}) = (\mathbf{B}' \otimes \mathbf{A})\text{vec}(\mathbf{Z})$, and rewriting (3) as $\text{vec}(\mathbf{Y}) = \sqrt{\rho/N_T}(\mathbf{S}^T \otimes \mathbf{I}_{N_R})\text{vec}(\mathbf{H}) + \text{vec}(\mathbf{E})$, where $\text{vec}(\cdot)$ stacks all of the columns of its argument into one long column vector.

The goal is to minimize the TMSE of both MLE and MMSE estimators, given by (6), under the power constraint of the training symbols

$$\min_{\frac{1}{N(L+1)}\text{tr}(\mathbf{SS}^T)\leq N_T} \frac{N_R N_T}{\rho} \text{tr}\left\{\left[\mathbf{SS}^T + \alpha\frac{N_T}{\rho}\mathbf{C}_{\mathbf{q}}^{-1} \otimes \mathbf{I}_{N_T}\right]^{-1}\right\}.$$

It has been shown that the training sequences should satisfy the condition of $\mathbf{SS}^T \propto \mathbf{I}$ [1][4][5], where $\propto$ means proportional to. If the condition is satisfied, it implies that the training sequence from each antenna is not only orthogonal to its temporal shifts within L taps, but also orthogonal to the training sequences from other antennas and their shifts within L taps, i.e.

- The aperiodic auto-correlation of the training sequence from each antenna is zero within L tap shifts, except at the zero shift, $r_{\mathbf{s}_i,\mathbf{s}_i}(k) = 0, |k| \in [1, L], \forall i$,
- The aperiodic cross-correlation of any two sequences is zero within L tap shifts, $r_{\mathbf{s}_i,\mathbf{s}_j}(k) = 0, |k| \leq L, \forall i \neq j$.

Since $\mathbf{S}$ has a block-Toeplitz structure, it is hard to find such training sequences [5][3], if only one training block is used. However, for two or more training blocks, we can build such optimal training sequences to satisfy $\mathbf{S}_1\mathbf{S}_1^T + \mathbf{S}_2\mathbf{S}_2^T \propto \mathbf{I}$ by using uncorrelated aperiodic complementary sets of sequences in Sec. V, where the indices 1 and 2 refer to preamble and postamble, respectively.

V. Construction of Optimal Training Sequences

Let $\{\mathbf{a}_0, \mathbf{a}_1\}$ be a pair of aperiodic complementary sequences given in Sec. II, each of length N. It is clear that the complementary property will not change if we append some 0's to the beginning or end or both. Therefore, we define a new pair of aperiodic complementary sequences as follows

$$\begin{aligned} \mathbf{u} &= [\mathbf{a}_0, \mathbf{o}], \quad & \breve{\mathbf{u}} &= [\overleftarrow{\mathbf{a}}_0, \mathbf{o}], \\ \mathbf{v} &= [\mathbf{a}_1, \mathbf{o}], \quad & \breve{\mathbf{v}} &= [\overleftarrow{\mathbf{a}}_1, \mathbf{o}], \end{aligned} \tag{7}$$

where $\mathbf{a}_0$ and $\mathbf{a}_1$ are generated by the recursion in (1) and $\mathbf{o} = \mathbf{0}_{1,(\lceil \frac{N_T}{2} \rceil - 1)(L+1)}$, then $N_s = N + (\lceil \frac{N_T}{2} \rceil - 1)(L + 1)$.

Based on the properties discussed in Sec. II, $\{\mathbf{u}, \breve{\mathbf{v}}\}$ and $\{\mathbf{v}, -\breve{\mathbf{u}}\}$ are mutually uncorrelated. If N_T is even, the training symbols assignment is given in Table I, where $\mathbf{s}_{n_t,1}$ and $\mathbf{s}_{n_t,2}$ are shown in Fig. 1, $\mathbf{\Pi}$ is the forward shift permutation matrix of order N_s [19, p. 27], and $\mathbf{x}\mathbf{\Pi}^p$ shifts the sequence $\mathbf{x}$ cyclically to the right by p elements. The similar assignment applies to the case where N_T is odd, which is omitted due to the limited space.

TABLE I
ASSIGNMENT OF TRAINING SYMBOLS TO TRANSMIT ANTENNAS

T_x	$\mathbf{s}_{n_t,1}$	$\mathbf{s}_{n_t,2}$
1	$\mathbf{u}$	$\breve{\mathbf{v}}$
2	$\mathbf{v}$	$-\breve{\mathbf{u}}$
3	$\mathbf{u}\mathbf{\Pi}^{L+1}$	$\breve{\mathbf{v}}\mathbf{\Pi}^{L+1}$
4	$\mathbf{v}\mathbf{\Pi}^{L+1}$	$-\breve{\mathbf{u}}\mathbf{\Pi}^{L+1}$
$\vdots$	$\vdots$	$\vdots$
$N_T - 1$	$\mathbf{u}\mathbf{\Pi}^{(\lceil \frac{N_T}{2} \rceil - 1)(L+1)}$	$\breve{\mathbf{v}}\mathbf{\Pi}^{(\lceil \frac{N_T}{2} \rceil - 1)(L+1)}$
N_T	$\mathbf{v}\mathbf{\Pi}^{(\lceil \frac{N_T}{2} \rceil - 1)(L+1)}$	$-\breve{\mathbf{u}}\mathbf{\Pi}^{(\lceil \frac{N_T}{2} \rceil - 1)(L+1)}$

When using two training blocks, i.e., preamble and postamble, (3) can be written as

$$[\mathbf{Y}_1, \mathbf{Y}_2] = \sqrt{\frac{\gamma}{N_T}}\mathbf{H}[\mathbf{S}_1, \mathbf{S}_2] + [\mathbf{E}_1, \mathbf{E}_2], \tag{8}$$

where γ is set as $\frac{N_s}{N}\rho$ by considering the power compensation due to the insertions of $N_s - N$ 0's into the original binary aperiodic complementary sequences, and the estimator in (5) simplifies to

$$\widehat{\mathbf{H}} = \sqrt{\frac{N_T}{\gamma}}\left(\sum_{t=1}^{2}\mathbf{Y}_t\mathbf{S}_t^T\right)\left(\sum_{t=1}^{2}\mathbf{S}_t\mathbf{S}_t^T + \alpha\frac{N_T}{\gamma}\mathbf{C}_{\mathbf{q}}^{-1} \otimes \mathbf{I}_{N_T}\right)^{-1}. \tag{9}$$

Combining (4), (7) and Table I, it is easy to show that $\mathbf{S}_1\mathbf{S}_1^T + \mathbf{S}_2\mathbf{S}_2^T = 2N\mathbf{I}_{N_T(L+1)}$, based on the complementary properties presented in Sec. II. This demonstrates the optimality of our design and simplifies (9) to

$$\widehat{\mathbf{H}} = \left(\sum_{t=1}^{2}\mathbf{Y}_t\mathbf{S}_t^T\right)\left[\left(2N\sqrt{\frac{\gamma}{N_T}}\mathbf{I}_{L+1} + \alpha\sqrt{\frac{N_T}{\gamma}}\mathbf{C}_{\mathbf{q}}^{-1}\right)^{-1} \otimes \mathbf{I}_{N_T}\right],$$

i.e.,

$$\left[\widehat{\mathbf{H}}_l\right]_{n_r,n_t} = \frac{P_l\sqrt{\gamma N_T}}{2\gamma N P_l + \alpha N_T}\left[\sum_{t=1}^{2}\mathbf{Y}_t\mathbf{S}_t^T\right]_{n_r,n_t'}, \tag{10}$$

where $n_t' = lN_T + n_t$, $n_r \in [1, N_R]$, $n_t \in [1, N_T]$, $l \in [0, L]$, and $[\mathbf{X}]_{m,n}$ is the $\{m, n\}^{\text{th}}$ element of the matrix $\mathbf{X}$.

VI. Fast Implementation of the Estimator

Without loss of generality, we assume the number of transmit antennas is even, i.e., N_T is even. By defining

$$^{q}_{n_r}\mathbf{H} = \begin{bmatrix} h_{n_r,2q-1}(0) & h_{n_r,2q-1}(1) & \cdots & h_{n_r,2q-1}(L) \\ h_{n_r,2q}(0) & h_{n_r,2q}(1) & \cdots & h_{n_r,2q}(L) \end{bmatrix}, \tag{11}$$

where $q = \lceil \frac{n_t}{2} \rceil$, and

$$_{n_r}\mathbf{H} = \begin{bmatrix} ^{1}_{n_r}\mathbf{H} & ^{2}_{n_r}\mathbf{H} & \cdots & ^{\frac{N_T}{2}}_{n_r}\mathbf{H} \end{bmatrix}, \tag{12}$$

[2]The concept was introduced in [12], and named later in [13] and [14].

[3]Note that ZCZ [15][16], Frank [17] and Chu [18] sequences can not satisfy the above condition, since they are developed based on the periodic correlation properties. Moreover, their structures are not suitable for fast hardware and software implementations.

whose dimension is $2 \times \frac{N_T}{2}(L+1)$. Based on (11) and (12), the estimator in (10) that corresponds to the n_r^{th} antenna, can be written as

$$ {}_{n_r}\widehat{\mathbf{H}} = \left(\sum_{t=1}^{2} \mathbf{X}_t \mathbf{Y}_{n_r,t}\right)\left(\mathbf{I}_{\frac{N_T}{2}} \otimes \mathbf{\Phi}\right), \tag{13} $$

where $\mathbf{X}_1 = \left[\begin{smallmatrix}\mathbf{a}_0\\ \mathbf{a}_1\end{smallmatrix}\right]$, $\mathbf{X}_2 = \left[\begin{smallmatrix}\overleftarrow{\mathbf{a}_1}\\ -\overleftarrow{\mathbf{a}_0}\end{smallmatrix}\right]$ and $\mathbf{\Phi} = \mathrm{diag}\Big(\frac{P_0\sqrt{\gamma N_T}}{2\gamma N P_0+\alpha N_T}, \frac{P_1\sqrt{\gamma N_T}}{2\gamma N P_1+\alpha N_T}, \cdots, \frac{P_L\sqrt{\gamma N_T}}{2\gamma N P_L+\alpha N_T}\Big)$. Furthermore, $\mathbf{Y}_{n_r,t}$ is a Hankel matrix whose dimension is $N \times \frac{N_T}{2}(L+1)$ with the $\{i,j\}^{\text{th}}$ element given by $y_{n_r,t}(i+j)$, i.e.,

$$ \mathbf{Y}_{n_r,t} = \begin{bmatrix} y_{n_r,t}(0) & y_{n_r,t}(1) & \cdots & y_{n_r,t}(N_s{+}L{-}N) \\ y_{n_r,t}(1) & \ddots & \ddots & y_{n_r,t}(N_s{+}L{-}N{+}1) \\ \vdots & \ddots & \ddots & \vdots \\ y_{n_r,t}(N{-}1) & y_{n_r,t}(N) & \cdots & y_{n_r,t}(N_s{+}L{-}1) \end{bmatrix}, \tag{14} $$

where $t \in \{1,2\}$.

A. Hardware (ASIC/FPGA) Implementation

First we consider the fast implementation of the preamble part, i.e. $\mathbf{X}_1\mathbf{Y}_{n_r,1}$. From the definition of $\mathbf{X}_1$ and (14), it is easy to see that $\mathbf{X}_1\mathbf{Y}_{n_r,1}$ does the correlation between the received signal $\mathbf{y}_{n_r,1}$ and the training sequences $\mathbf{a}_0$ and $\mathbf{a}_1$, respectively, with different delays. The same reasoning applies to $\mathbf{X}_2\mathbf{Y}_{n_r,2}$, which does the correlation between $\mathbf{y}_{n_r,2}$ and $\overleftarrow{\mathbf{a}_0}$, $-\overleftarrow{\mathbf{a}_1}$, respectively.

It is well known that the correlation can be implemented by the convolution, so we can put a filter structure shown in Fig. 2 for each receive antenna and do the parallel processing to boost the estimation speed with low complexity. The fast Golay correlator (FGC) in Fig. 2, which is introduced in [10], can be obtained by taking the Z transform of (1) with respect to k

$$ \begin{aligned} A_0^{(m)}(z) &= A_0^{(m-1)}(z) + A_1^{(m-1)}(z)z^{-D_m} \\ A_1^{(m)}(z) &= A_0^{(m-1)}(z) - A_1^{(m-1)}(z)z^{-D_m} \end{aligned} \tag{15} $$

with $A_0^{(0)}(z) = A_1^{(0)}(z) = 1$. And the Efficient Golay Correlator (EGC) in Fig. 2 is described in [20].

In Fig. 2, $\widehat{\mathbf{h}}_{n_r,t}$, $t \in \{1,2\}$, is the t^{th} row of ${}_{n_r}\widehat{\mathbf{H}}$. The "Extractor" unit discards the first $N-1$ values, takes the following $\frac{N_T}{2}(L+1)$ values, and discards the remaining $N-1$ values, the "$\bigotimes$" unit after "Extractor" means the l^{th} CIR of all subchannels is multiplied by the factor $\frac{P_l\sqrt{\gamma N_T}}{2\gamma N P_l+\alpha N_T}$. Note: some "$\bigoplus$" units have one "+" branch and one "−" branch, while some have two "+" branches.

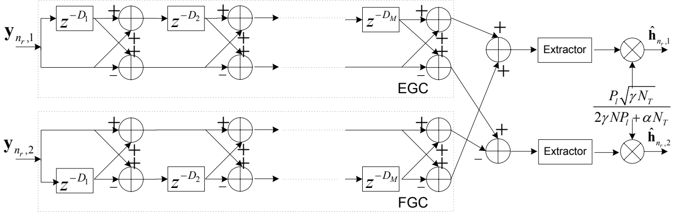

Fig. 2. Filter structure of the fast MLE and MMSE estimators on the n_r^{th} receive antenna.

About the hardware complexity, the total number of addition units on each receive antenna is $4\log_2 N + 2$ [21]. However, upon the straightforward implementation of (13) with non-binary symbols by 4 different correlators are used ($\mathbf{a}_0$, $\mathbf{a}_1$, $\overleftarrow{\mathbf{a}_0}$ and $-\overleftarrow{\mathbf{a}_1}$), the total number of required units for addition and multiplication are $4(N-1)+2$ and $4N$, respectively.

B. Software (DSP) Implementation

In Subsec. VI-A, the fast filter structure is introduced, which is suitable for the AISC or FPGA implementation. Nowadays, DSP processors have enough power to implement complex estimation and detection algorithms, and are widely used in both base stations and mobile terminals. In this subsection, we introduce the fast DSP implementation of (13).

First, we focus on computing $\mathbf{X}_1\mathbf{Y}_{n_r,1}$ in an efficient way. Without loss of the generality, we assume $D_m = 2^{m-1}$, based on (1) in Sec. II. In addition, we set $\mathbf{Y}_{n_r,1} = \left[\begin{smallmatrix}\mathbf{P}^{(M)}\\ \mathbf{Q}^{(M)}\end{smallmatrix}\right]$, where $\mathbf{P}^{(M)}$ contains the first 2^{M-1} rows of $\mathbf{Y}_{n_r,1}$, and $\mathbf{Q}^{(M)}$ contains the last 2^{M-1} rows of $\mathbf{Y}_{n_r,1}$. Furthermore, we define $\mathbf{P}^{(M)} = \left[\begin{smallmatrix}\mathbf{P}_1^{(M)}\\ \mathbf{P}_2^{(M)}\end{smallmatrix}\right]$ and $\mathbf{Q}^{(M)} = \left[\begin{smallmatrix}\mathbf{Q}_1^{(M)}\\ \mathbf{Q}_2^{(M)}\end{smallmatrix}\right]$. Now $\mathbf{X}_1\mathbf{Y}_{n_r,1}$ in (13) can be written as

$$ \begin{aligned} \mathbf{X}_1\mathbf{Y}_{n_r,1} &= \begin{bmatrix}\mathbf{a}_0^{(M)}\\ \mathbf{a}_1^{(M)}\end{bmatrix}\begin{bmatrix}\mathbf{P}^{(M)}\\ \mathbf{Q}^{(M)}\end{bmatrix}, \\ &= \begin{bmatrix}\mathbf{a}_0^{(M-1)} & \mathbf{a}_1^{(M-1)}\\ \mathbf{a}_0^{(M-1)} & -\mathbf{a}_1^{(M-1)}\end{bmatrix}\begin{bmatrix}\mathbf{P}^{(M)}\\ \mathbf{Q}^{(M)}\end{bmatrix}, \\ &= \begin{bmatrix}\mathbf{a}_0^{(M-2)} & \mathbf{a}_1^{(M-2)} & \mathbf{a}_0^{(M-2)} & -\mathbf{a}_1^{(M-2)}\\ \mathbf{a}_0^{(M-2)} & \mathbf{a}_1^{(M-2)} & -\mathbf{a}_0^{(M-2)} & \mathbf{a}_1^{(M-2)}\end{bmatrix}\begin{bmatrix}\mathbf{P}_1^{(M)}\\ \mathbf{P}_2^{(M)}\\ \mathbf{Q}_1^{(M)}\\ \mathbf{Q}_2^{(M)}\end{bmatrix}, \\ &= \begin{bmatrix}\mathbf{a}_0^{(M-2)}\mathbf{P}^{(M-1)+} + \mathbf{a}_1^{(M-2)}\mathbf{Q}^{(M-1)-}\\ \mathbf{a}_0^{(M-2)}\mathbf{P}^{(M-1)-} + \mathbf{a}_1^{(M-2)}\mathbf{Q}^{(M-1)+}\end{bmatrix}, \end{aligned} \tag{16} $$

where

$$ \begin{aligned} \mathbf{P}^{(M-1)+} &= \mathbf{P}_1^{(M)} + \mathbf{Q}_1^{(M)}, \\ \mathbf{P}^{(M-1)-} &= \mathbf{P}_1^{(M)} - \mathbf{Q}_1^{(M)}, \\ \mathbf{Q}^{(M-1)+} &= \mathbf{P}_2^{(M)} + \mathbf{Q}_2^{(M)}, \\ \mathbf{Q}^{(M-1)-} &= \mathbf{P}_2^{(M)} - \mathbf{Q}_2^{(M)}, \end{aligned} \tag{17} $$

The number of additions (including substractions) in (17) is $2(N_s + L - 2^{M-1}) = N + 2(L'-1)$, where $L' = \frac{N_T}{2}(L+1)$. In addition, $\mathbf{a}_0^{(M-2)}\mathbf{P}^{(M-1)+} + \mathbf{a}_1^{(M-2)}\mathbf{Q}^{(M-1)-}$ and $\mathbf{a}_0^{(M-2)}\mathbf{P}^{(M-1)-} + \mathbf{a}_1^{(M-2)}\mathbf{Q}^{(M-1)+}$ can be calculated in a recursive way [22] with $2\left[2\sum_{m=0}^{M-3}(2^m + L' - 1) + L'\right]$ additions. The total number of additions for each $\mathbf{X}_1\mathbf{Y}_{n_r,1}$ and $\mathbf{X}_2\mathbf{Y}_{n_r,2}$ is $(4\log_2 N - 5)L' + 2(N - 2\log_2 N + 1)$. So, the total number of additions in $\sum_{t=1}^{2}\mathbf{X}_t\mathbf{Y}_{n_r,t}$ is $8(\log_2 N -$

$1)L' + 4(N - 2\log_2 N + 1)$ and the number of multiplications is L'.

If we calculate (13) directly, the number of additions is $2L'(2N-1)$ and the number of multiplications is L'. Clearly, direct computation of (13) requires much more operations. For example, if $N = 128$, $L = 31$, and $N_T = 4$, then the total number of additions in our fast recursive method is 3532, but is 32640 in the direct computation of (13).

VII. Simulation Results

In the simulations, we take $L = 15$, which means 16 taps in each subchannel, and $P_l = \frac{1}{L+1}, \forall l$. The two scenarios considered are $N_T = 4$, $N_R = 4$ and $N = 64$, as well as $N_T = 4$, $N_R = 4$ and $N = 128$.

Fig. 3 shows the theoretical minimum classical CRLB (when the channel is treated as unknown deterministic) and minimum Bayesian CRLB (when the channel is treated as random), given by $\frac{N_T(L+1)}{2\rho N_s}$ and $N_T \sum_{l=0}^{L} \frac{P_l}{2\rho N_s P_l + N_T}$, respectively, which are normalized by $\mathbb{E}[\|\mathbf{H}\|_F^2] = N_R N_T$, and the simulated TMSEs of both the MLE and MMSE estimators, $\overline{\|\widehat{\mathbf{H}} - \mathbf{H}\|_F^2} / \overline{\|\mathbf{H}\|_F^2}$, versus different SNR levels. Note that all the simulation and theoretical results perfectly match. From Fig. 3, it is easy to see that the longer the training, the better the estimation performance. Also, MMSE performs better than MLE in low-SNR region as it takes advantage of the prior statistical information of the channel.

Although the end-to-end bit-error-rate (BER) performance is the final result we care about, it is generally true that the less the estimation error, the better the BER performance, as long as other system aspects are fixed. Fig. 3 shows the proposed schemes achieve the minimum possible CRLB, which is the best we can do for the channel estimation. Therefore, the proposed scheme will achieve the best BER performance, compared with other channel estimators.

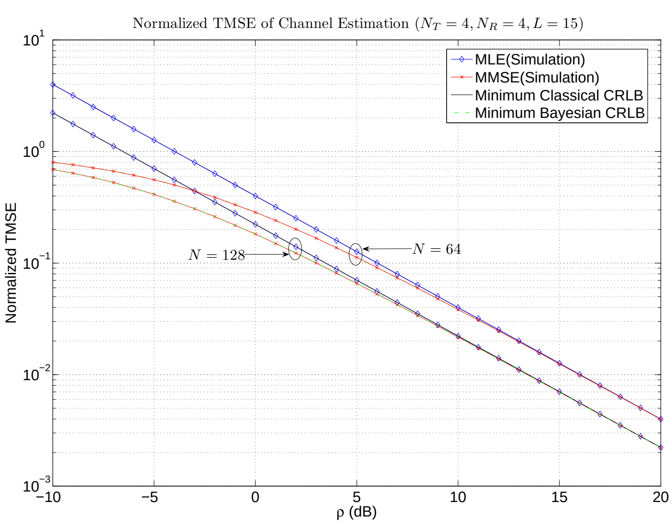

Fig. 3. The normalized TMSE of the MLE and MMSE estimators: $N_T = 4$, $N_R = 4$, $L = 15$, and $N = 64, 128$.

VIII. Conclusion

In this paper we have constructed optimal training sequences using uncorrelated aperiodic complementary sets of sequences, for estimating block-fading MIMO frequency selective channels. Theoretical and simulation results are provided to demonstrate the optimality of the proposed estimator. Compared with the one-side training scheme, the proposed two-sided scheme needs one more gap of length L for separating training and data symbols. However, this overhead is negligible if L is much smaller than the frame length, which is generally true in practice. Most importantly, both fast hardware and software implementations of the proposed estimators are also presented in this paper, which make them suitable and ready for real-world MIMO applications.

References

[1] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inform. Theory*, vol. 49, pp. 951–963, 2003.

[2] O. Simeone and U. Spagnolini, "Lower bound on training-based channel estimation error for frequency-selective block-fading Rayleigh MIMO channels," *IEEE Trans. Signal Processing*, vol. 52, pp. 3265–3277, 2004.

[3] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Upper Saddle River, NJ: Prentice Hall PTR, 1993.

[4] X. Ma, L. Yang, and G. B. Giannakis, "Optimal training for MIMO frequency-selective fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 453–466, 2005.

[5] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*. Cambridge, UK: Cambridge University Press, 2003.

[6] P. Spasojevic and C. N. Georghiades, "Complementary sequences for ISI channel estimation," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1145–1152, 2001.

[7] C. C. Tseng and C. L. Liu, "Complementary set of sequences," *IEEE Trans. Inform. Theory*, vol. 18, pp. 644–652, 1972.

[8] P. Fan and M. Darnell, *Sequence Design for Communications Applications*. Hertfordshire, UK: Research Studies Press, 1996.

[9] M. J. E. Golay, "Complementary series," *IEEE Trans. Inform. Theory*, vol. 7, pp. 82–87, 1961.

[10] S. Z. Budišin, "Efficient pulse compressor for Golay complementary sequences," *Electron. Lett.*, vol. 27, pp. 219–220, 1991.

[11] S. Wang, K. Raghukumar, A. Abdi, J. Wallace, and M. Jensen, "Indoor MIMO channels: A parametric correlation model and experimental results," in *Proc. IEEE Sarnoff Symp.*, Princeton, NJ, 2004, pp. 1–5.

[12] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Vol. I*. New York: Wiley, 1968.

[13] B. Z. Bobrovsky, E. Mayer-Wolf, and M. Zakai, "Some classes of global Cramér-Rao bounds," *Ann. Statist.*, vol. 15, pp. 1221–1438, 1987.

[14] R. D. Gill and B. Y. Levit, "Applications of the van Trees inequality: A Bayesian Cramér-Rao bound," *Bernoulli*, vol. 1, pp. 59–79, 1995.

[15] S.-A. Yang and J. Wu, "Optimal binary training sequence design for multiple-antenna systems over dispersive fading channels," *IEEE Trans. Veh. Technol.*, vol. 51, pp. 1271–1276, 2002.

[16] P. Fan and W. H. Mow, "On optimal training sequence design for multiple-antenna systems over dispersive fading channels and its extensions," *IEEE Trans. Veh. Technol.*, vol. 53, pp. 1623–1626, 2004.

[17] R. Frank, S. Zadoff, and R. Heimiller, "Phase shift pulse codes with good periodic correlation properties," *IEEE Trans. Inform. Theory*, vol. 8, pp. 381–382, 1962.

[18] D. Chu, "Polyphase codes with good periodic correlation properties," *IEEE Trans. Inform. Theory*, vol. 18, pp. 531–532, 1972.

[19] P. J. Davis, *Circulant Matrices*. New York: Wiley, 1979.

[20] B. M. Popović, "Efficient Golay correlator," *Electron. Lett.*, vol. 35, pp. 1427–1428, 1999.

[21] S. Wang and A. Abdi, "Aperiodic complementary sets of sequences-based MIMO frequency selective channel estimation," accepted for publication in *IEEE Commun. Lett.*, 2005.

[22] B. Xu and G. Bi, "Channel estimation using complementary sequence pairs for UWB/OFDM systems," *Electron. Lett.*, vol. 40, pp. 1196–1197, 2004.