Optimal Training Sequences for Channel Estimation in Cyclic-Prefix-Based Single-Carrier Systems With Transmit Diversity

Justin Coon, Student Member, IEEE, Mark Beach, Associate Member, IEEE, and Joe McGeehan

Abstract—In this paper, we investigate a new class of training sequences that are optimal for least squares (LS) channel estimation in systems employing transmit diversity and single-carrier (SC) modulation with a cyclic prefix (CP) extension. The sequences have a constant envelope in the time domain and are orthogonal in the frequency domain. Transmission of these sequences facilitates optimal (in the LS sense) estimation of the channel impulse response at the receiver while precluding the peak-to-average power ratio problem that is inherent in other CP-based architectures such as orthogonal frequency division multiplexing.

Index Terms—Diversity methods, optimal training.

I. INTRODUCTION

In the past decade, the majority of research in the area of frequency-domain equalization (FDE) techniques has been focused primarily on orthogonal frequency division multiplexing (OFDM) due to its elegance and low complexity. Recently, researchers have taken interest in cyclic-prefix-based single-carrier (CP-SC) transmission with FDE at the receiver and have shown that this technique is a promising alternative to OFDM [1], [2]. While channel estimation in OFDM systems with transmit diversity is a well-researched area (e.g., see [3]–[7]), channel estimation that lends itself to CP-SC systems has seen relatively little attention in the literature. In [8], a recursive reconstructive algorithm for channel estimation in multiple-input multiple-output (MIMO) CP-SC systems was presented and novel training sequences for use in this algorithm were given. In this paper, we investigate these training sequences further in the context of the more general problem of least squares (LS) channel estimation in CP-SC systems with transmit diversity. We prove that these sequences are optimal in the sense that their exploitation leads to the minimum mean-square error (MSE) channel estimate. Furthermore, we derive a simple expression for the maximum number of transmit antennas that can be employed in the LS channel estimation problem and show that these sequences support this number.

Notation: We use a bold uppercase (lowercase) font to denote matrices (column vectors). \( \mathbf{F} \) is the normalized \( K \times K \) DFT matrix and \( \mathbf{F}_L \) is a matrix comprising the first \( L \) columns of \( \sqrt{K} \mathbf{F} \); \( \mathbf{I}_L \) is the \( L \times L \) identity matrix; \( \mathbf{0}_{L \times L} \) is the \( L \times L \) all-zero matrix; \( (\cdot)^*, (\cdot)^{-1} \), \( (\cdot)^T \), \( (\cdot)^H \), and \( [\cdot] \) denote the complex conjugate, inverse, transpose, conjugate transpose, and absolute value operations, respectively; \( \mathbb{E}\{\cdot\} \) is the expectation operator; \( \mathfrak{r}\{\cdot\} \) is the trace operator; \( \text{diag}\{\mathbf{a}\} \) is a diagonal matrix with the elements of \( \mathbf{a} \) on the diagonal; and \( \delta(\cdot) \) is the Kronecker delta function.

II. LEAST SQUARES CHANNEL ESTIMATION

As previously mentioned, equalization of CP-SC transmissions is typically performed in the frequency domain\(^1\). Consequently, we use the frequency domain to formulate the LS estimation problem. It should be noted, however, that once the channel has been estimated, any suitable time-domain or frequency-domain detection technique can be used for data recovery.

We begin by formulating an LS channel estimator for a CP-SC system with \( n_T \) transmit antennas. Define a block of \( K \) training symbols transmitted from the \( q \)-th antenna by \( \mathbf{t}_q = (t_q(0),\ldots,t_q(K-1))^T \). A CP of \( Q \) symbols is added to \( \mathbf{t}_q \) prior to transmission and removed at the receiver to mitigate interblock interference (IBI). Assuming a CP of adequate length is implemented (i.e., \( Q \geq L - 1 \) where \( L \) is the total number of taps in the discrete channel impulse response (CIR)), we can write the received baseband symbol vector after the removal of the CP as

\[
\mathbf{y} = \sum_{q=1}^{n_T} \mathbf{G}_q \mathbf{t}_q + \mathbf{\eta} \tag{1}
\]

where \( \mathbf{G}_q \) is a circulant matrix modeling the channel between the \( q \)-th transmit antenna and the receiver and \( \mathbf{\eta} \) is a vector of zero-mean, uncorrelated, complex Gaussian noise samples with variance \( \sigma^2 \) per dimension. Performing a DFT on the received samples \( \mathbf{y} \) gives

\[
\mathbf{z} = \mathbf{Fy} = \sum_{q=1}^{n_T} \mathbf{FG}_q \mathbf{F}^{-1} \mathbf{Ft}_q + \mathbf{F} \mathbf{\eta} = \sum_{q=1}^{n_T} \mathbf{H}_q \mathbf{t}_q + \mathbf{F} \mathbf{\eta} \tag{2}
\]

\(^1\)CP-SC systems can also utilize time-domain equalization (TDE); however, FDE benefits from lower complexity than TDE [1].
where \( \mathbf{u}_q = \mathbf{F} \mathbf{t}_q, \mathbf{H}_q = \mathbf{F} \mathbf{G}_q \mathbf{F}^{-1} = \text{diag}\{ \mathbf{h}_q \} \), and \( \mathbf{h}_q(k) \) is the frequency response of the \( q \)-th channel at the \( k \)-th tone for \( 0 \leq k \leq K - 1 \). We can rewrite (2) as

\[
\mathbf{z} = \sum_{q=1}^{n_T} \mathbf{U}_q \mathbf{h}_q + \mathbf{F} \eta
= \sum_{q=1}^{n_T} \mathbf{U}_q \mathbf{F} \mathbf{L} \mathbf{g}_q + \mathbf{F} \eta
\]

(3)

where \( \mathbf{U}_q = \text{diag}\{ \mathbf{u}_q \} \) and \( \mathbf{g}_q \) is the length-\( L \) vector of CIR taps for the \( q \)-th channel. Noting that

\[
\mathbf{z} = \left[ \mathbf{U}_1 \mathbf{F}_L \cdots \mathbf{U}_{n_T} \mathbf{F}_L \right] \mathbf{A}^{-1} \mathbf{g} + \mathbf{F} \eta
\]

(4)

we see that the resulting LS channel estimate is given by

\[
\hat{\mathbf{g}} = \mathbf{A}^\dagger \mathbf{z}
= \mathbf{g} + \mathbf{A}^\dagger \mathbf{F} \eta
\]

(5)

where \( \mathbf{A}^\dagger = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \) is the pseudoinverse of \( \mathbf{A} \). Note, the channel vector \( \mathbf{g} \) is identifiable only if \( \mathbf{A} \) has full column rank, which occurs when \( K \geq L n_T \). Consequently, the number of transmit antennas that can be employed is bounded by

\[
n_T \leq \frac{K}{L}.
\]

(6)

The MSE of the channel estimate \( \hat{\mathbf{g}} \) is easily derived as follows.

\[
\text{MSE} = \frac{1}{L n_T} \text{tr} \mathbb{E}\{ (\hat{\mathbf{g}} - \mathbf{g})(\hat{\mathbf{g}} - \mathbf{g})^H \}
= \frac{1}{L n_T} \text{tr} \mathbb{E}\{ (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{F} \eta \mathbf{F}^H \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \}
= \frac{\sigma_n^2}{L n_T} \text{tr}\{ (\mathbf{A}^H \mathbf{A})^{-1} \}.
\]

(7)

In [4], it was shown that the MSE is bounded by

\[
\text{MSE} \geq \frac{\sigma_n^2}{K}
\]

(8)

where the equality is met if and only if \( (\mathbf{A}^H \mathbf{A})^{-1} \) is a diagonal matrix and all of the elements on the diagonal of \( (\mathbf{A}^H \mathbf{A})^{-1} \) are equal. Rewriting \( \mathbf{D}^{-1} = (\mathbf{A}^H \mathbf{A})^{-1} \) as

\[
\mathbf{D}^{-1} = \begin{bmatrix}
\mathbf{F}_L^H \mathbf{U}_1^H \mathbf{U}_L & \cdots & \mathbf{F}_L^H \mathbf{U}_1^H \mathbf{U}_{n_T} \mathbf{F}_L \\
\vdots & \ddots & \vdots \\
\mathbf{F}_L^H \mathbf{U}_{n_T}^H \mathbf{U}_1 & \cdots & \mathbf{F}_L^H \mathbf{U}_{n_T}^H \mathbf{U}_{n_T} \mathbf{F}_L
\end{bmatrix}^{-1}
\]

(9)

and using a similar argument as in [4], we see that in order to obtain the minimum MSE channel estimate, we must have

\[
\mathbf{D}_{pq} = \mathbf{F}_L^H \mathbf{U}_p^H \mathbf{U}_q \mathbf{F}_L = \begin{cases}
K \mathbf{I}_L, & p = q \\
0_{L \times L}, & p \neq q
\end{cases}
\]

(10)

In the next section, we present sequences that meet the criterion stated in (10) and discuss some of their properties.

### III. Optimal Training Sequences

In [8], a recursive reconstructive algorithm for channel estimation in MIMO CP-SC systems was presented and a new class of training sequences for use with this algorithm were given. In this work, we present a more general overview of the sequences and prove that these sequences are optimal in terms of channel estimation MSE and the number of antennas that they can support.

#### A. Sequence Construction

Before discussing the construction of the training sequences, we first review Chu sequences, which are polyphase sequences that have a constant magnitude in both the time domain and the frequency domain [9]. This interesting property stems from the fact that Chu sequences have perfect periodic autocorrelation properties (i.e., a Kronecker delta function). It is the property of constant frequency-domain magnitude that makes Chu sequences invaluable in the design of optimal training sequences since it leads to the equality of all of the diagonal elements in \( \mathbf{D} \), which will be shown later. The constant time-domain magnitude of Chu sequences precludes peak-to-average power ratio problems that plague many CP-based systems. The \( n \)-th element of a length-\( P \) Chu sequence is given by

\[
a(n) = \begin{cases}
\cos(\pi n^2/P), & \text{for even } P \\
\cos(\pi n(n+1)/P), & \text{for odd } P
\end{cases}
\]

(11)

where \( \ell \) and \( P \) are relatively prime [9].

To construct the training sequences for a CP-SC system with \( n_T \) transmit antennas, one base sequence is first designed from an arbitrary length-(\( K/n_T \)) Chu sequence by repeating it \( n_T \) times. This repetition in the time-domain causes a stretching of the frequency-domain characteristics of the sequence. Consequently, if \( \mathbf{t}_1 \triangleq (t_1(0), \ldots, t_1(K/n_T - 1))^T \) is the length-(\( K/n_T \)) Chu sequence and \( \mathbf{t}_1 \triangleq (t_1(0), \ldots, t_1(K - 1))^T \) is the repeated sequence, the \( k \)-th tone of the DFT of \( \mathbf{t}_1 \) is given by

\[
u_1(k) = \begin{cases}
u_1(k), & k \text{ mod } n_T = 0 \\
0, & \text{otherwise}
\end{cases}
\]

(12)

where \( \nu_1(k) \in \mathbb{C} \). Equation (12) follows from a property of repeated finite-length sequences [10].

The second part of (10) suggests that the transmitted training sequences must be orthogonal in the frequency domain. This orthogonality can easily be achieved by progressively rotating the phases of the base training symbols in the time domain, resulting in a constant shift of the frequency-domain sequence. This is just the frequency shift property of DFT’s [10]:

\[
t_3(n) = \mathbf{e}^{j2\pi(n-1)/K} \leftrightarrow u_3(k - q + 1).
\]

(13)
Thus, the sequence transmitted from the \( q \)th antenna for \( q = 2, 3, \ldots, n_T \) is given by
\[
t_q(n) = t_q(n)e^{2\pi in(q-1)/K},
\]
For brevity, we refer to these sequences as RPC (repeated, phase-rotated, Chu) sequences throughout the rest of this paper.

**B. Optimality of Sequences**

Now that we have presented a thorough explanation of the construction of RPC sequences, we prove that these sequences are optimal in the sense that they achieve the minimum channel estimation MSE. Returning to (10), we may write the \((r,s)\)th element of the matrix \( D_{pq}^{r,s} \) as
\[
D_{pq}^{r,s} = \sum_{k=0}^{K-1} u_p^*(k)u_q(k)e^{-j2\pi k(s-r)/K} \tag{15}
\]
for \(0 \leq r,s \leq L - 1\). From (10), we want
\[
D_{pq}^{r,s} = \begin{cases} K, & p = q, \quad r = s \\ 0, & \text{otherwise}. \end{cases} \tag{16}
\]
It is easy to see that for \( p \neq q, D_{pq}^{r,s} = 0 \) regardless of \( r \) and \( s \). This equality follows from the orthogonality of RPC sequences in the frequency domain. Therefore, RPC sequences satisfy the second part of (10).

To prove that RPC sequences satisfy the first part of (10), we observe the case where \( p = q \) and rewrite (15) as
\[
D_{pq}^{r,s} = \sum_{k=0}^{K-1} |u_q(k)|^2 e^{-j2\pi k(s-r)/K}. \tag{17}
\]
Let \( \Delta = |s-r| \) and consider the elements on and below the diagonal of \( D_{pq}^{r,s} \) for which \( r \geq s \). We now have
\[
D_{pq}^{r,s} = \sum_{k=0}^{K-1} |u_q(k)|^2 e^{2\pi jk\Delta/K} \nonumber
\]
\[
= \frac{1}{K} \sum_{k=0}^{K-1} |\text{IDFT}(t_q)_k|^2 e^{2\pi jk\Delta/K} \nonumber
\]
\[
= |\text{IDFT}(\text{PSD}(t_q))_\Delta| \tag{18}
\]
for \(0 \leq \Delta \leq L - 1\) where \( \text{DFT}(\cdot)_i \) and \( \text{IDFT}(\cdot)_i \) denote the conventional DFT and inverse DFT operations, respectively, evaluated at \( i \) and \( \text{PSD}(t_q)_k = |\text{IDFT}(t_q)_k|^2 \) denotes the power spectral density of \( t_q \). It follows from the Wiener-Khinchin theorem that \( D_{pq}^{r,s} \) is the periodic autocorrelation function of \( t_q \) evaluated at time \( \Delta = 0, \ldots, L - 1 \) [11]. Recall from Section II that in order for the channel to be identifiable, \( L \leq K/n_T \) must be true. Consequently, we are most interested in the case where \( L = K/n_T \) and the periodic autocorrelation function of \( t_q \) is unique for \( \Delta = 0, \ldots, K/n_T - 1 \).

**Proposition 1:** The periodic autocorrelation function \( R(\Delta) \) of any length-\( K \) RPC sequence is given by
\[
R(\Delta) = K\delta(\Delta) \tag{19}
\]
for \(0 \leq \Delta \leq K/n_T - 1\). (The proof is given in the Appendix.)

From Proposition 1, \( D_{pq}^{r,s} = K\delta(\Delta) \) for all identifiable LS problems (i.e., \( L \leq K/n_T \)). Since \( D_{pq}^{r,s} \) is Hermitian symmetric, the elements on and above the diagonal are given by
\[
D_{pq}^{r,s} = D_{pq}^{s,r*} = K\delta(\Delta), \tag{20}
\]
It follows from Proposition 1 and (20) that RPC sequences satisfy the first part of (10) and are optimal in the MSE sense.

As a final note, we observe that RPC sequences are constructed from a single Chu sequence. Since at least one Chu sequence exists for any finite length, we conclude that RPC sequences support the maximum number of antennas, which is determined by the condition \( n_T \leq K/L \) as given in (6).

**IV. RESULTS AND DISCUSSION**

We verified the optimality of RPC sequences through computer simulations. All of the simulated systems used \( n_T = 4 \) transmit antennas and \( n_R = 4 \) receive antennas. For each system, blocks of \( K = 64, 16\)-QAM symbols were encoded according to one of two space-time processing techniques: spatial multiplexing (SM) [2] and space-time block codes (STBC) [12]. A CP was appended to each block at the transmitter and removed at the receiver to eliminate IBI. Each single-input single-output channel was modeled as having \( L = 11 \) i.i.d. complex Gaussian taps. We assumed that the channels were spatially uncorrelated and stationary for the transmission of one data block.

LS channel estimation was performed at the receiver using random BPSK training sequences and RPC sequences. As a benchmark, one system was simulated with perfect channel state information (CSI). A linear frequency-domain minimum mean-square error (MMSE) equalizer was employed to remove ISI. Fig. 1 shows the performance of the various simulated systems. Note that a substantial gain in performance can be obtained by employing RPC training sequences rather than random sequences.

Fig. 2 illustrates the MSE of the LS channel estimate when both random and RPC sequences are used. The lower bound on MSE, which is given by (8) and illustrated in Fig. 2, is met up to a scaling factor with the RPC sequences. The scaling factor, \( n_T \), is due to the normalization of total transmit power in the simulations. Therefore, RPC sequences are, indeed, optimal in the MSE sense, and the advantage of using RPC training sequences is clearly visible in Fig. 1 and 2.
Fig. 2. MSE of LS channel estimates for RPC and random BPSK training sequences. \( n_T = n_R = 4, K = 64 \).

V. CONCLUSION

In this paper, we investigated training sequences for channel estimation in multi-antenna CP-SC systems. We proved that these sequences are optimal in the sense that their implementation results in the minimum MSE channel estimate and they support the maximum number of transmit antennas. Computer simulations were used to verify their optimality where it was shown that they significantly outperform random sequences.

APPENDIX

PROOF OF PROPOSITION 1

First, consider the periodic autocorrelation function \( \mathcal{R}_q(\Delta) \) of the RPC sequence \( t_q \) at \( \Delta = 0 \) where

\[
\mathcal{R}_q(0) = \sum_{n=0}^{K-1} t_q(n) t_q^*(n) = \sum_{n=0}^{K-1} t_1(n) e^{j \frac{2 \pi n (q-1)}{K}} t_2^*(n) e^{-j \frac{2 \pi n (q-1)}{K}} = \sum_{n=0}^{K-1} |t_1(n)|^2 = K. \tag{A.21}
\]

Now, consider the periodic autocorrelation function \( \mathcal{R}_q(\Delta) \) of \( t_q \) for \( 1 \leq \Delta \leq \frac{K}{n_T} - 1 \) where

\[
\mathcal{R}_q(\Delta) = \sum_{n=0}^{K-1} t_q(n) \mathbf{t}_q^*(n + \Delta) = \mathcal{S}_q(\Delta) + \sum_{n=K-\Delta}^{K-1} t_q(n) \mathbf{t}_q^*(n + \Delta - K), \tag{A.22}
\]

We have

\[
\mathcal{S}_q(\Delta) = \sum_{n=0}^{K-\Delta-1} t_1(n) e^{j \frac{2 \pi n (q-1)}{K}} \cdot t_2^*(n + \Delta) e^{-j \frac{2 \pi n (q-1) \Delta}{K}} = e^{-j \frac{2 \pi n (q-1) \Delta}{K}} \sum_{n=0}^{K-\Delta-1} t_1(n) t_2^*(n + \Delta) \tag{A.23}
\]

and

\[
\mathcal{T}_q(\Delta) = \sum_{n=K-1}^{K-\Delta-1} t_1(n) e^{j \frac{2 \pi n (q-1)}{K}} \cdot t_2^*(n + \Delta - K) e^{-j \frac{2 \pi n (q-1) \Delta}{K}} = e^{-j \frac{2 \pi n (q-1) \Delta}{K}} \sum_{n=K-\Delta}^{K-1} t_1(n) t_2^*(n + \Delta - K) \tag{A.24}
\]

Therefore, (A.22) reduces to

\[
\mathcal{R}_q(\Delta) = e^{-j \frac{2 \pi n (q-1) \Delta}{K}} (\mathcal{S}_q(\Delta) + \mathcal{T}_q(\Delta)) = e^{-j \frac{2 \pi n (q-1) \Delta}{K}} \mathcal{R}_1(\Delta), \tag{A.25}
\]

But \( \mathcal{R}_1(\Delta) \) is just the periodic autocorrelation function of a length-\( (K/n_T) \) Chu sequence, which is zero for \( 1 \leq \Delta \leq K/n_T - 1 \). Thus \((19)\) is true for \( 0 \leq \Delta \leq K/n_T - 1 \), which concludes the proof.

\[\Box\]

ACKNOWLEDGMENT

The authors are indebted to Dr. M. Sandell, Dr. R. Piechocki, Mr. J. Siew, and the anonymous reviewers for their insightful comments regarding this work.

REFERENCES