# Finite-Alphabet Constant-Amplitude Training Sequence for Multiple-Antenna Broadband Transmissions

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

***Abstract* – We propose a method to identify training sequences for multiple-antenna transmissions over quasi-static frequency-selective channels. These sequences are constructed to belong to a standard constant-amplitude $2^m$-PSK constellation (such as BPSK, QPSK etc) to simplify the transmitter/receiver implementation. Many practical systems use training of predetermined length. Optimal sequences do not exist for all training sequence lengths and constellation alphabets. The proposed method allows us to identify training sequences that belong to a standard constellation for an arbitrary training sequence length and an arbitrary number of unknown channel taps. Performance bounds derived indicate that these sequences achieve near-optimum performance.**

## I. INTRODUCTION

For quasi–static or slowly–varying fading channels, current wireless communication systems employ training–based channel estimation at the receiver. For implementation purposes (e.g. to avoid nonlinear amplifier distortion), it is desirable to use training sequences with constant amplitude. The construction of training sequences with constant amplitude can be classified in two main categories according to the training symbol alphabet size $N$.

The first approach [1] constructs optimal sequences[1] from an $N^{th}$ root-of-unity alphabet $A_N = \{exp(\frac{i2\pi k}{N}) : k = 0 \dots N-1\}$, without constraining the alphabet size $N$.

Such sequences are the perfect root-of-unity sequences (PRUS) or polyphase sequences that have been proposed in the literature for different applications (see [2] and the references therein). For any training sequence length $N_t$, there exist optimal training sequences that belong to an $N^{th}$ root-of-unity alphabet. The training sequence length $N_t$ determines the smallest possible alphabet size. Chu [3] shows that for any length $N_t$ there exists a PRUS with alphabet size $N = 2N_t$, and Mow [4] shows that for some $N_t$, smaller alphabet sizes are possible.

The second approach in the literature constrains the training sequence symbols to belong to a specific constellation, typically BSPK or QPSK, to simplify the transmitter/receiver implementation [5]. In this case, optimal sequences do not exist for all training sequence lengths $N_t$. Instead, exhaustive searches can identify suboptimal sequences according to some performance criteria. Tables of such sequences from a BPSK alphabet are provided, for example, in [5].

The training sequence best suited to a particular application depends on the training sequence length $N_t$, which for standardized systems is predetermined, the number of channels taps to estimate, and the signal constellation used. A PRUS of a predetermined length may not belong to a standard constellation, while exhaustive searches are in many cases computationally prohibitive. For example, for a system with $M$ transmit antennas, training sequence length $N_t$, and constellation of size $N$, the search space involves $N^{N_t \times M}$ sequences.

This paper proposes a method to easily identify training sequences that enjoy the following attractive properties

1. They belong to a standard constant-amplitude signal constellation of size $2^m$ for $m = 1, 2, 3, \dots$ such as BPSK, QPSK, 8-PSK etc.
2. They can be easily identified or constructed for an arbitrary training sequence length $N_t$ and an arbitrary number of unknown channel taps $L$.
3. They result in negligible channel estimation mean square error increase from its theoretical lower bound.

This paper is organized as follows. Section II defines the optimal training sequences. Section III introduces the "L-perfect" sequences. Section IV investigates a method (alternative to exhaustive search) to identify suboptimal training sequences when optimal training sequences do not exist, and derives bounds on the performance loss. Section V presents simulation results and the paper is concluded in Section VI.

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

0-7803-7400-2/02/$17.00 © 2002 IEEE

## II. Channel model and optimal training sequences

Consider a single–antenna transmitting the training signal $\mathbf{s}$ of length $N_t$ over a frequency-selective channel $\mathbf{h} = [h(0) \ \ldots \ h(L-1)]^T$ with $L$ taps. We assume that the channel $\mathbf{h}$ remains constant over the transmission of a block and varies independently from block to block.

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

$$\mathbf{y} = \mathbf{S}\mathbf{h} + \mathbf{z},$$

where $\mathbf{y}$ and $\mathbf{z}$ are of dimension $(N_t - L + 1) \times 1$ and $\mathbf{S}$ is a Toeplitz matrix of dimension $(N_t - L + 1) \times L$

$$\mathbf{S}(L, N_t) = \begin{bmatrix} s(L-1) & \ldots & s(0) \\ s(L) & \ldots & s(1) \\ \vdots & & \vdots \\ s(N_t - 1) & \ldots & s(N_t - L) \end{bmatrix}. \quad (2)$$

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 [6]

$$\hat{\mathbf{h}} = (\mathbf{S}^H\mathbf{S})^{-1}\mathbf{S}^H\mathbf{y}, \quad (3)$$

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 Mean Square Error (MMSE) is equal to

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

and is achieved if and only if [1]

$$\mathbf{S}^H\mathbf{S} = (N_t - L + 1)\mathbf{I}. \quad (6)$$

A sequence $\mathbf{s}$ that satisfies (6) is henceforth called an *optimal* sequence.

For multiple transmit antennas, to simplify the presentation, consider the system of Fig. 1 that employs two-transmit and one-receive antennas. The observed training sequence output in this case can be expressed as

$$\mathbf{y} = \mathbf{S}\mathbf{h} + \mathbf{z} = [\mathbf{S}_1(L, N_t) \ \ \mathbf{S}_2(L, N_t)] \begin{bmatrix} \mathbf{h}_1(L) \\ \mathbf{h}_2(L) \end{bmatrix} + \mathbf{z}.$$

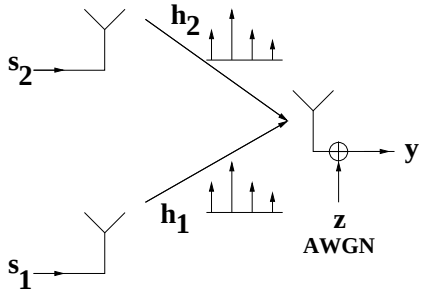


Fig. 1. System with two-transmit and one-receive antennas.

The MMSE is achieved if and only if [1]

$$\mathbf{S}^H\mathbf{S} = \begin{bmatrix} \mathbf{S}_1^H\mathbf{S}_1 & \mathbf{S}_2^H\mathbf{S}_1 \\ \mathbf{S}_1^H\mathbf{S}_2 & \mathbf{S}_2^H\mathbf{S}_2 \end{bmatrix} = (N_t - L + 1)\mathbf{I}. \quad (8)$$

Equation (8) implies that the optimal sequences have an impulse-like auto-correlation sequence and zero cross-correlation.

A straightforward method to design two optimal training sequences $\mathbf{s}_1$ and $\mathbf{s}_2$ of length $N_t$ to estimate two channels each of $L$ taps, is to design instead a single training sequence $\mathbf{s}$ of length $N_t' = N_t + L$ to estimate a single channel with $L' = 2L$ taps [7], [8] and construct the sequences $\mathbf{s}_1$ and $\mathbf{s}_2$ as

$$\mathbf{s}_1 = [s(0) \ \ldots \ s(N_t)] \text{ and } \mathbf{s}_2 = [s(L) \ \ldots \ s(N_t + L)].$$

Thus, the multiple-training sequence design problem is now reduced to designing a single, but longer, optimal sequence $\mathbf{s}$ that achieves the MMSE when estimating the longer channel impulse response with $L'$ taps. A similar approach can be followed for more than two transmit antennas.

Different methods to reduce the problem of designing multiple training sequences to the problem of designing a single training sequence with impulse-like auto-correlation are also possible. For example, the method in [7] encodes 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.

In the following, we address the design of a single training sequence of length $N_t$ used to estimate $L$ channel taps without restricting its use to multiple or single antenna systems.

## III. Perfect and L-perfect training sequences

A root-of-unity sequence with alphabet size $N$ has complex roots of unity elements of the form

$\{exp(\frac{i2\pi x}{N})\}$, with $x = 0 \ldots N-1$. The $N$-roots of unity define a constant-amplitude finite-size constellation.

A sequence is said to be $perfect$ if all of its out-of-phase periodic auto-correlation terms are equal to zero [4]. The periodic auto-correlation of a sequence $\mathbf{s}$ of length $N_t$ at shift $t$ can be calculated as

$$\theta(t) = \sum_{n=0}^{N_t-1} s^*(n)s((n+t) \mod N_t).$$

A unified construction method in [2], [4] constructs a PRUS of any length $N_{PRUS}$ but with alphabet size $N$ determined by $N_{PRUS}$. More specifically, let $N_{PRUS} = pt^2$ where $p$ and $t$ are both positive integers and $p$ is not a square of an integer greater than 1, then (see [4], Corollary 3)

$$N = \begin{cases} 2pt & \text{for p even and t odd,} \\ pt & \text{otherwise.} \end{cases} \quad (11)$$

We are only interested in alphabets of size $N = 2^m$, $m \geq 1$. From (11), this implies that $2pt = 2^m$ for $p$ even and $t$ odd (or $t = 2^{m-1}$ otherwise), which in turn implies that the construction method in [4], when the alphabet size is a power of two, produces PRUS of length $N_{PRUS}$ which is also a power of two.

A sequence $\mathbf{s}$ of length $N_t$ is called $L-perfect$ if the corresponding training matrix $\mathbf{S}$ of dimension $(N_t - L + 1) \times L$ (constructed as in Equation (2)) satisfies $\mathbf{S}^H\mathbf{S} = (N_t - L + 1)\mathbf{I}_L$. Thus, an L-perfect sequence of length $N_t$ is optimal (i.e. achieves the MMSE in (5)) for a channel with $L$ taps.

**Proposition 1** The length $N_t$ of an L-perfect sequence from a $2^m$-constellation can only be equal to

$$N_t = \left\{ \begin{array}{ll} 2(L+i) & \text{for } L\text{=odd} \\ 2(L+i)-1 & \text{for } L\text{=even} \end{array} \right\}, \text{ for } i = 0, 1, \ldots \quad (12)$$

*Proof*: A necessary condition for the columns of $\mathbf{S}$ to be orthogonal is that $N_t - L + 1 \geq L$, that is, the number of rows is greater or equal to the number of columns. Moreover, it is easy to show that for $2^m$-roots of unity and any $x_1,\ x_2 \in \{2^m\text{-roots of unity}\}$, there does not exist $x_3 \in \{2^m\text{-roots of unity}\}$ such that $x_1 + x_2 = x_3$. This property implies that the number of rows of $\mathbf{S}$ which is $(N_t - L + 1)$ has to be an even number so that the inner-product of two columns involves an even number of terms to be canceled out. □

Equation (12) is a necessary (but not sufficient) condition for L-perfect sequences of length $N_t$ to exist.

From a perfect sequence of length $N_{PRUS}$, we can build L-perfect training sequences to estimate up to $L = N_{PRUS}$ unknowns. The L-perfect training sequence length $N_t$ can only be equal to $kN_{PRUS} + L - 1$ for any $k \geq 1$ integer. Let $\mathbf{s}_P(N_{PRUS}) = [s_p(0)\ \ldots\ s_p(N_{PRUS}-1)]$ denote the perfect sequence, then the L-perfect sequence of length $kN_{PRUS} + L - 1$ can be constructed by repeating $k$ times the perfect sequence and circularly extending it by $L-1$ symbols, for example

$$[s(0)\ s(1)\ \ldots s(kN_{PRUS}+L-2)] = \underbrace{[\underbrace{s_p(0)\ \ldots\ s_p(N_{PRUS}-1)}_{\mathbf{s}_P}\ \ldots s(0)\ s(1)\ \ldots s(L-2)]}_{\text{repeat } k \text{ times}}.$$

Table I shows exhaustive search results for $L$-perfect sequences for some small $L$ and $N_t$. The numerical entries denote the number of sequences found, "$*$" indicates that all or some of the existing sequences can be generated from perfect sequences, "$x$" denotes that such sequences exist but their exact number was not calculated, while "$-$" indicates that no such sequence exists. No entry denotes that no search was performed.

For our search range, perfect sequences could be used to construct most, but not all, of the L-perfect sequences. $L$-perfect sequences exist for a broader range of $N_t$ than perfect sequences can provide. For example, for $L = 3$ and sequence length $N_t = 8$, there exist L-perfect sequences that can not be constructed from perfect sequences.

## IV. Suboptimal training sequences to estimate $L$ unknowns

$L$-perfect sequences do not exist for all training sequence lengths $N_t$. For example, for a specific $L$ and alphabet size $N$, if there exists an L-perfect sequence of length $N_t$, then from (12), there does not exist an L-perfect sequence of length $N_t + 1$. For length $N_t + 1$, only suboptimal solutions are possible. Moreover, even if L-perfect sequences exist, if they cannot be constructed from perfect sequences, it might be computationally intensive to identify them through computer search. In the next subsection, we propose a method to construct suboptimal L-perfect sequences and give MSE bounds on their performance.

TABLE I

L-PERFECT SEQUENCES THAT EXIST FOR THE MINIMUM LENGTH AND FROM SMALL-SIZE CONSTELLATIONS

| **L=2 taps** | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| const. $N_t$ | 3 | 2 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| BPSK | 4 | - | $^*12$ | - | 40 | - | $^*140$ | - | 504 | - | $^*x$ | - |
| QPSK | $^*16$ | - | $^*144$ | - | $^*1600$ | - | $^*x$ | - | $^*x$ | - | $^*x$ | - |
| 8PSK | $^*64$ | - | $^*1344$ | - | $^*x$ | - | $^*x$ | - | $^*x$ | - | $^*x$ | - |
| **L=3 taps** | | | | | | | | | | | | |
| const. $N_t$ | 3 | 2 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| BPSK | - | - | - | $^*8$ | - | - | - | $^*72$ | - | | - | $^*x$ |
| QPSK | - | - | - | $^*64$ | - | 192 | - | $^*3776$ | - | | - | $^*x$ |
| 8PSK | - | - | - | $^*512$ | - | 768 | - | $^*x$ | - | | - | $^*x$ |
| **L=4 taps** | | | | | | | | | | | | |
| const. $N_t$ | 3 | 2 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| BPSK | - | - | - | - | $^*8$ | - | - | - | $^*40$ | - | | - |
| QPSK | - | - | - | - | $^*64$ | - | - | - | $^*1216$ | - | | - |
| 8PSK | - | - | - | - | $^*512$ | - | | - | $^*x$ | - | | - |

### *A. Construction method*

Assume that $N_t$ symbols from a specific alphabet are available to estimate $L$ unknowns. Express $N_t$ as $N_t = kN_{PRUS}+L-1+M$ for a nonnegative integer $M$ and for a PRUS from the desired alphabet. Choose the value of $N_{PRUS} \geq L$ that minimizes $M$. Construct the L-perfect sequence of length $kN_{PRUS} + L - 1$ and extend it by adding $M$ symbols through exhaustive search. If $M = 0$, the solution is optimal. If $M = 1$, no search is needed as the following Proposition 2 states.

For AWGN with variance $2\sigma^2$, the MSE is equal to $2\sigma^2 tr((\mathbf{S}^H\mathbf{S})^{-1})$. In the following discussion we assume that $2\sigma^2 = 1$, i.e. we drop the term $2\sigma^2$ which is common for all different training matrices $\mathbf{S}$ and thus does not play any role in their performance evaluation.

**Proposition 2** Consider an L-optimal training sequence of length $N_t$ from an $2^m$-roots of unity alphabet. Adding one training symbol to create a training sequence of length $N_t + 1$ leads to MSE value (which is denoted by $MSE_1$) equal to

$$MSE_1 = \frac{L}{N_S} - \frac{L}{N_S(N_S + L)}, \tag{14}$$

where $N_S = N_t - L + 1$ is the number of rows of the $\mathbf{S}$ matrix. Note that $MSE_1$ does not depend on the added symbol.

The proof of Propositions 2, 3 and 4 are provided in [8]. The MSE achieved with the above extension method for suboptimal sequences of length $N_t + 1$ is not always the minimum MSE possible for this training sequence length and restricted alphabet.

**Proposition 3** Consider an L-optimal training sequence of length $N_t$ from a $2^m$-roots of unity alphabet. Adding two training symbols to create a training sequence of length $N_t + 2$ amounts to adding two rows $\mathbf{r}_1$ and $\mathbf{r}_2$ to matrix $\mathbf{S}_{N_S}$, i.e.

$$\mathbf{S}_{N_S+2} = \begin{bmatrix} \mathbf{S}_{N_S} \\ \mathbf{r}_1 \\ \mathbf{r}_2 \end{bmatrix}, \tag{15}$$

and leads to an MSE value, which we denote by $MSE_2$, equal to

$$MSE_2 = MSE_1 - \frac{(L + N_S)^2 L - a(L + 2N_S)}{N_S(L + N_S)((L + N_S)^2 - a)},$$

where $a = |\mathbf{r}_1\mathbf{r}_2^*|^2$ and $MSE_1$ is given by (14).

Continuing along the same lines, one could derive the $MSE_k$ achieved when extending an L-optimal sequence by $k$ symbols, but the calculations become tedious as $k$ increases. Instead, we give an upper bound on the MSE as a function of the extension length $k$. This bound allows us to evaluate the maximum possible MSE increase from MMSE.

**Proposition 4** Consider an L-optimal training sequence of length $N_t$ from a $2^m$-roots of unity alphabet. Extending the training sequence to length $N_t + k$ by adding $k$ symbols results in $MSE_k$ upper–bounded as follows

$$MSE_k \leq \frac{L}{N_S}(1 - \frac{k}{Lk + N_S}). \tag{17}$$

Thus, the maximum increase in MSE from not using an optimal sequence is upper-bounded by

$$MSE_k - MMSE \leq \frac{L}{N_S}(1 - \frac{k}{Lk + N_S}) - \frac{L}{k + N_S}.$$

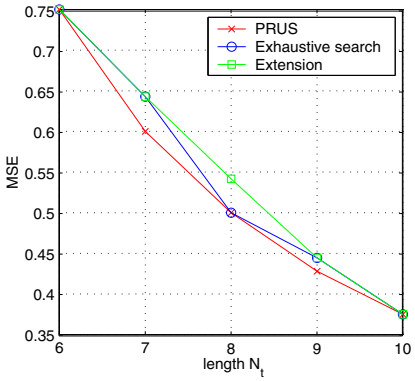


Fig. 2. MSE when using PRUS (any alphabet size), sequences from a QPSK-alphabet identified through exhaustive search, and sequences from a QPSK-alphabet identified through the proposed construction method, respectively, for a channel with L=3 taps.

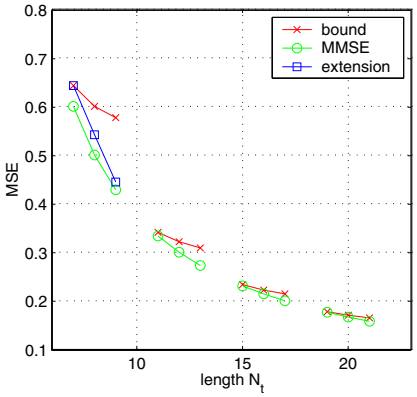


Fig. 3. MSE achieved with PRUS (any alphabet size) and upper bound on MSE achieved through extension. Assumed L=3 taps.

This bound does not depend on the constellation employed or the extension symbols. It upper-bounds the largest MSE we may get by extending an L-perfect sequence by $k$ randomly chosen symbols. For $k = 1$, the bound reduces to $(14)$.

## V. MSE PERFORMANCE

We examine the achievable MSE of the extension method proposed in Section IV for a specific example that employs QPSK-alphabet to estimate $L = 3$ channel taps.

Fig. 2 plots three curves versus training sequence length: the minimum MSE achieved with PRUS and no restriction on the alphabet size, the minimum MSE for training sequences with QPSK alphabet found through exhaustive search, and the MSE for training sequences with QPSK alphabet identified from the proposed construction method. Perfect sequences can be used to construct L-perfect sequences of length $N_t = 6$ and 10. Extending the L-perfect sequences by one symbol leads to the same minimum MSE achieved by an exhaustive search over all possible sequences of this length. However, there exist perfect sequences of length $N_t = 8$ which cannot be produced by extending the length $N_t = 6$ sequences by two symbols. On the other hand, there exist length $N_t = 6$ sequences which, when extended by four symbols produce perfect sequences of length $N_t + 4 = 10$.

Fig. 3 plots the bound in Equation (17) and the optimal MMSE. The bound predicts the largest MSE we may get by extending an L-perfect sequence by some randomly-chosen symbols, i.e. the worst-case scenario. Observe that the bound gets closer to the MMSE as the training length increases.

## VI. CONCLUSIONS

We proposed a method to identify suboptimal training sequences of arbitrary length for any number of channel taps to be estimated. Bounds on the MSE increase using the proposed method indicate that achievable performance is close to optimal. Our focus was on training symbols belonging in alphabets of size $2^m$ such as BPSK, QPSK, and 8-PSK, as they simplify the transmitter/receiver structure and result in negligible MSE increase from the case of training with optimal PRUS with unrestricted alphabet. Reduced–complexity channel estimation algorithms for multiple–transmit–antenna systems are discussed in [7].

## REFERENCES

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

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

[3] D. C. Chu. Polyphase codes with good correlation properties. *IEEE Transactions on Info. Theory*, 18:531–532, July 1972.

[4] W. H. Mow. A new unified construction of perfect root-of-unity sequences. *Spread Spectrum Techniques and Applications Proceedings*, 3:955–959, 1996.

[5] 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.

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

[7] C. Fragouli, N. Al-Dhahir, and W. Turin. Reduced-Complexity Training Schemes for Multiple-Antenna Broadband Transmission. *Accepted to IEEE Wireless Communications and Networking Conference (WCNC) 2002*, March 2002.

[8] C. Fragouli, N. Al-Dhahir, and W. Turin. Training-Based Channel Estimation for Multiple-Antenna Broadband Transmissions. *Accepted under modifications to J-SAC on Wireless Communications*, June 2001.