

An Accurate and Efficient Two-Stage Channel Estimation Method Utilizing Training Sequences with Closed Form Expressions

Çağatay Candan

Abstract—A novel two-stage frequency domain channel estimation method especially suitable for the estimation of long channels such as ultra wide band channels is proposed. The proposed method can efficiently use the sequences with closed form analytical expressions such as the Legendre sequences. (The suggested method does not require a computationally intense search for good training sequences which is infeasible for long training sequences.) The method is shown to present a minor improvement in the total estimation error variance when compared with the conventional single stage frequency domain channel estimation. In addition, the proposed method has a very efficient time domain implementation requiring at most $2N$ multiplications, where N is the training sequence length, in comparison to $O(N \log N)$ multiplications required for the conventional method.

Index Terms—Channel estimation, frequency domain channel estimation, Legendre sequences, ultra wide band channel estimation.

I. INTRODUCTION

AN accurate knowledge of the channel is required to establish reliable and high rate communication. In the absence of any a-priori information on the channel coefficients, the maximum likelihood (ML) method is utilized for the estimation of coefficients, [1, p. 490]. The design objective of this approach is to improve the estimation performance, that is to reduce the total estimation error variance, with a proper training sequence selection, [2]–[6]. It is well known that the optimal sequences to this aim (also called perfect codes) have an impulsive auto-correlation, that is a flat energy spectrum, under additive white Gaussian noise conditions.

Some well known codes such as the maximum length sequences (m-sequences) and the Legendre sequences almost satisfy the requirement for perfect codes. As an example, the m-sequence of length N ($N = 2^k - 1$, k is an integer) has the Discrete Fourier Transform (DFT) spectrum of

$$|C[k]|^2 = \begin{cases} 1, & k = 0 \\ N + 1, & k = \{1, 2, \dots, N - 1\} \end{cases}$$

The equation given above shows that m-sequences spread the signal energy equally across the AC spectrum ($k > 0$) and distribute the remaining energy ($1/N^2$ of the total energy) to the DC frequency ($k = 0$)¹. We can say that the constraint of

Paper approved by C. Tellambura, the Editor for Multicarrier Systems of the IEEE Communications Society. Manuscript received July 2, 2010; revised March 28, 2011 and May 29, 2011.

The author is with the Department of Electrical and Electronics Engineering, Middle East Technical University (METU), Ankara, Turkey (e-mail: ccandan@metu.edu.tr).

Digital Object Identifier 10.1109/TCOMM.2011.081711.100155

¹AC spectrum is a term inherited from electrical engineering literature and indicates the components of a signal having non-zero frequencies. The DC component of a sequence refers to its mean value or the zero'th DFT coefficient.

spectrum flatness is indeed satisfied, when the DC component is not considered.

The contribution of the present paper is the description of a method to improve the performance of almost perfect sequences in the channel estimation problem. To that aim, we propose a novel two-stage frequency domain channel estimation method in which the AC components of the channel are estimated first and then the remaining unknown (DC value of the channel) is determined. The separation of AC and DC components of the channel is shown to improve the estimation performance in comparison to the conventional method. The method operates in the DFT domain and can be categorized as a frequency domain channel estimation method. In spite of the mentioned performance improvement over the conventional method (frequency domain channel estimation), we consider the major advantage of the proposed method is the availability and abundance of the codes with closed form expressions to achieve the improved performance.

Good training sequences for the conventional single stage method are found through a computationally intensive search, [2], [7]. Specifically for the frequency domain channel estimation, Tellambura *et al.* have proposed a reduced complexity search technique in [2]. This procedure is based on dividing the search space into equivalence classes and searching the select classes in which the optimal code can reside. This procedure, even with the significant reduction of search complexity, is not feasible to implement for large length training sequences [2]. A family of efficient binary training sequences with a closed form expression for the channel estimation problem is not known for the conventional method.

In this paper, instead of presenting good codes for the conventional method, we present a novel two-stage method to improve the estimation performance of almost perfect codes. The suggested method is specific for almost perfect codes, that is m-sequences and Legendre sequences, but we emphasize the utilization of the Legendre sequences due to their abundance in number. (Legendre sequences are defined for every odd prime valued lengths.)

The paper is organized as follows: First the procedure for the estimation of AC part of the channel is described. This section can also be interpreted as an independent work on the estimation of AC coupled channels, that is the channels with a null at DC frequency such as powerline, capacitive coupled channels. The following section presents the two-stage method and the Legendre sequences. The paper concludes after the discussions on the estimation performance, comparisons with alternative methods and the description of an efficient implementation specific for the Legendre sequences.

II. ESTIMATION OF THE AC COMPONENTS

We present a brief review of frequency domain channel estimation and then proceed with the estimation of the AC components of the channel. The estimation of AC components is the first stage of the proposed method. The first stage can be solely utilized in the estimation of AC coupled channels, that is the channels having a null at the DC frequency due to the placed transformers or capacitors in the communication line.

A training sequence $c[n]$ with the cyclic prefix is shown below:

$$\underbrace{c[N-T], \dots, c[N-1]}_{\text{prefix}}, \underbrace{c[0], c[1], \dots, c[N-1]}_{\text{training sequence}}$$

We assume that $c[n]$ is of length N and binary valued. The cyclic-prefix, appended to the beginning of the training sequence, is of length T where $T \geq L-1$ and L is the maximum length of the channel to be estimated.

The signal received after discarding the cyclic prefix can be represented as $\mathbf{r} = \mathbf{C}\mathbf{h} + \mathbf{w}$. Here \mathbf{C} is the circular convolution matrix of size $N \times L$ and \mathbf{h} is the unknown $L \times 1$ channel coefficient vector. The first column of \mathbf{C} is the training sequence (code) and it is denoted as $[c[0], \dots, c[N-1]]$. The vector \mathbf{w} is zero-mean circularly symmetric complex Gaussian vector with covariance matrix $\sigma_w^2 \mathbf{I}$. The ML estimate of the channel coefficients is written as $\hat{\mathbf{h}} = (\mathbf{C}^H \mathbf{C})^{-1} \mathbf{C}^H \mathbf{r}$ and the resulting error covariance matrix is $\sigma_w^2 (\mathbf{C}^H \mathbf{C})^{-1}$, [1]. The design of optimal training sequences aims to minimize the total estimation error variance, that is $\sigma_w^2 \text{trace}\{(\mathbf{C}^H \mathbf{C})^{-1}\}$, under the total energy constraint, $\text{trace}\{\mathbf{C}^H \mathbf{C}\} \leq NL$. It is known that the optimal solution to this problem requires $\mathbf{C}^H \mathbf{C}$ to be proportional with the identity matrix, that is $\mathbf{C}^H \mathbf{C} \propto \mathbf{I}$, [8, p.176]. When the optimality condition is satisfied, the total error variance, which is the lower bound on the error variance, is $\sigma_w^2 L/N$.

The first step of the frequency domain channel estimation is the application of DFT to both sides of the equation $\mathbf{r} = \mathbf{C}\mathbf{h} + \mathbf{w}$. Here $\bar{\mathbf{h}}$ is the zero-padded version of \mathbf{h} , that is $N-L$ zeros are padded to \mathbf{h} . $\bar{\mathbf{C}}$ matrix is a circular convolution matrix of dimensions $N \times N$. The $N \times L$ submatrix of $\bar{\mathbf{C}}$ comprised of its first L columns is identical to the \mathbf{C} matrix, i.e. \mathbf{C} is embedded in $\bar{\mathbf{C}}$. With the application of DFT on both sides of $\mathbf{r} = \mathbf{C}\mathbf{h} + \mathbf{w}$, we get the following:

$$\underbrace{\mathbf{F}\mathbf{r}}_{\mathbf{r}_F} = \underbrace{\mathbf{F}\bar{\mathbf{C}}\mathbf{F}^{-1}}_{\mathbf{C}_F} \underbrace{\mathbf{F}\bar{\mathbf{h}}}_{\mathbf{h}_F} + \underbrace{\mathbf{F}\mathbf{w}}_{\mathbf{w}_F} \quad (1)$$

Here \mathbf{F} represents $N \times N$ DFT matrix and the subscript $(\cdot)_F$ (as in \mathbf{r}_F) denotes the DFT domain representation of the vectors and matrices.

The ML estimation of \mathbf{h}_F from (1) exactly follows from the discussion given for $\mathbf{r} = \mathbf{C}\mathbf{h} + \mathbf{w}$, but with an important interpretation difference due to DFT and time domains. The estimation goal in the frequency domain is \mathbf{h}_F which is a vector of length N representing the channel coefficients in DFT domain. In the time domain, the same channel vector is represented with a vector of length L where L is the channel length and $L < N$. The optimality condition for the frequency domain channel estimation remains the same, that is $\mathbf{C}_F^H \mathbf{C}_F \propto \mathbf{I}$. Once the frequency domain components are estimated, the

time domain representation of the channel can be retrieved by applying the inverse DFT on $\hat{\mathbf{h}}_F$. The resulting mean square error (MSE) is $N\sigma_w^2 \text{trace}\{(\mathbf{F}_L^{-1})(\mathbf{C}_F^H \mathbf{C}_F)^{-1}(\mathbf{F}_L^{-1})^H\}$, where \mathbf{F}_L^{-1} is the first L rows of inverse DFT matrix. The total error variance for the optimal training sequences (when $\mathbf{C}_F^H \mathbf{C}_F \propto \mathbf{I}$) is $\sigma_w^2 L/N$ which is identical to the lower bound in the time domain, [2].

The estimation of AC coupled channels or the AC part of an arbitrary channel can be posed similar to the conventional channel estimation problem along with an additional constraint of $\sum_{k=0}^{L-1} h[k] = 0$. To describe the estimation procedure under the given constraint, we first partition the $N \times N$ DFT and inverse DFT matrices as follows:

$$\mathbf{F} = \begin{bmatrix} \mathbf{1}^H \\ \mathbf{U}_{N-1}^H \end{bmatrix}; \quad \mathbf{F}^{-1} = \frac{1}{N} \begin{bmatrix} \mathbf{1} & \mathbf{U}_{N-1} \end{bmatrix} \quad (2)$$

The row vector $\mathbf{1}^H$ appearing above is the vector of all 1's. When \mathbf{F} and \mathbf{F}^{-1} is substituted in (1), we get the following:

$$\begin{bmatrix} \mathbf{1}^H \\ \mathbf{U}_{N-1}^H \end{bmatrix} \mathbf{r} = \frac{1}{N} \begin{bmatrix} \mathbf{1}^H \bar{\mathbf{C}} \mathbf{1} & \mathbf{1}^H \bar{\mathbf{C}} \mathbf{U}_{N-1} \\ \mathbf{U}_{N-1}^H \bar{\mathbf{C}} \mathbf{1} & \mathbf{U}_{N-1}^H \bar{\mathbf{C}} \mathbf{U}_{N-1} \end{bmatrix} \mathbf{h}_F + \mathbf{w}_F \quad (3)$$

Since $\bar{\mathbf{C}}$ is a square circular convolution matrix; $\mathbf{1}^H \bar{\mathbf{C}}$ is then equal to $C[0] \mathbf{1}^H$ where $C[0] = \sum_{n=0}^{N-1} c[n]$ and $C[0]$ is the DC value of the training sequence $c[n]$. We note that $\mathbf{1}^H \bar{\mathbf{C}} \mathbf{1} = NC[0]$ and $\mathbf{1}^H \bar{\mathbf{C}} \mathbf{U}_{N-1} = \mathbf{0}^H$ where $\mathbf{0}^H$ is the row vector of all zeros. Given these, the equation (3) can be simplified as follows:

$$r_F[0] = C[0]h_F[0] + w_F[0] \quad (4)$$

$$\tilde{\mathbf{r}}_F = \frac{1}{N} \underbrace{\mathbf{U}_{N-1}^H \bar{\mathbf{C}} \mathbf{U}_{N-1}}_{\tilde{\mathbf{C}}_F} \tilde{\mathbf{h}}_F + \tilde{\mathbf{w}}_F \quad (5)$$

Here $r_F[0]$ is the DC value received vector \mathbf{r} and $\tilde{\mathbf{r}}_F$ is the vector of AC components, that is $\tilde{\mathbf{r}}_F = \mathbf{U}_{N-1}^H \mathbf{r}$. (The other vectors with tilde on top should also be interpreted as the vector of AC components.) Since $h_F[0]$ is known to be zero for AC coupled channels, the equation (4) is redundant and can not be utilized further². It should be noted that $C[0]$ should be set to zero under finite energy constraint since the equation where $C[0]$ appears does not contain any useful information for the estimation of AC components.

The AC components of the channel can be estimated from equation (5). As noted before, the optimal training sequence for the given structure satisfies the relation $(\bar{\mathbf{C}}_F)^H \tilde{\mathbf{C}}_F \propto \mathbf{I}$, that is

$$\tilde{\mathbf{C}}_F^H \tilde{\mathbf{C}}_F = \frac{1}{N^2} \mathbf{U}_{N-1}^H (\bar{\mathbf{C}})^H \mathbf{U}_{N-1} \mathbf{U}_{N-1}^H \bar{\mathbf{C}} \mathbf{U}_{N-1} \propto \mathbf{I} \quad (6)$$

Since $\mathbf{U}_{N-1} \mathbf{U}_{N-1}^H \bar{\mathbf{C}} = N(\bar{\mathbf{C}} - C[0] \mathbf{1} \mathbf{1}^H)$ and $C[0] = 0$ for the optimal code, the constraint for optimal codes given in (6) reduces to the following:

$$\tilde{\mathbf{C}}_F^H \tilde{\mathbf{C}}_F = \frac{1}{N} \mathbf{U}_{N-1}^H (\bar{\mathbf{C}})^H \bar{\mathbf{C}} \mathbf{U}_{N-1} \propto \mathbf{I} \quad (7)$$

Since $\bar{\mathbf{C}}$ is a square circular convolution matrix, the matrix $(\bar{\mathbf{C}})^H \bar{\mathbf{C}}$ is another circular convolution matrix and

²An application for equation (4) can be the estimation of the noise variance, as noted in [2].

$\mathbf{U}_{N-1}^H (\overline{\mathbf{C}})^H \overline{\mathbf{C}} \mathbf{U}_{N-1}$ is a diagonal matrix, since the columns of DFT matrix are the eigenvectors of the circulant matrices. Hence the right hand side of the equality given in (7) is a diagonal matrix. To satisfy the optimality condition, the diagonal matrix $\mathbf{U}_{N-1}^H (\overline{\mathbf{C}})^H \overline{\mathbf{C}} \mathbf{U}_{N-1}$ should have the same value along its diagonal.

Since the matrix $(\overline{\mathbf{C}})^H \overline{\mathbf{C}}$ is the circular correlation matrix of the code $c[n]$, the diagonal entries of the matrix $\mathbf{U}_{N-1}^H (\overline{\mathbf{C}})^H \overline{\mathbf{C}} \mathbf{U}_{N-1}$ are $\{|C[1]|^2, |C[2]|^2, \dots, |C[N-1]|^2\}$ where $C[k]$ is the DFT of $c[n]$.

By Parseval's relation, we can write the following equation for $|C[k]|^2$:

$$\frac{1}{N} \sum_{k=0}^{N-1} |C[k]|^2 = \sum_{n=0}^{N-1} |c[n]|^2 = E_c \quad (8)$$

Since the optimal code is shown to satisfy $C[0] = 0$ and $|C[1]| = |C[2]| = \dots = |C[N-1]|$, the value for $|C[k]|^2$ can be expressed as follows:

$$|C[k]|^2 = \begin{cases} 0, & k = 0 \\ \frac{N}{N-1} E_c, & k \neq 0 \end{cases} \quad (9)$$

where E_c is the total energy of the training sequence and defined in (8).

The lower bound on the total estimation error variance is then

$$\text{Minimum Error Variance} : \sigma_w^2 \frac{(N-1)^2}{N^2 E_c} L. \quad (10)$$

A. Legendre Sequences

We briefly describe the Legendre sequences and show that these sequences exactly satisfy the optimality condition given in (9).

The Legendre symbol, from which the sequence is derived, is defined as follows, [9]–[11]:

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & a \in \mathcal{R} \\ -1 & a \notin \mathcal{R} \\ 0 & a = 0 \end{cases} \quad (11)$$

Here p appearing in $\left(\frac{a}{p}\right)$ is an odd prime number and a is defined in modulo p . The set \mathcal{R} indicates the quadratic residues for modulo p . A non-zero number a is called a quadratic residue if a solution for the following equation exists $a \equiv b^2 \pmod{p}$. One may interpret the quadratic residues as the numbers whose "square root exists" in modulo p . As an example, under modulo 7 we have $\{0^2, 1^2, 2^2, 3^2, 4^2, 5^2, 6^2\} \equiv \{0, 1, 4, 2, 2, 4, 1\} \pmod{7}$. From this list, we can express the set of quadratic residues as $\mathcal{R} = \{1, 2, 4\}$ and the set of quadratic non-residues as $\{3, 5, 6\}$. Readers can examine [9] for more information on quadratic residues which is a topic intricately connected with many results in the number theory.

The Legendre sequence is defined as follows:

$$b[n] = \left(\frac{n}{p}\right), \quad n = \{0, \dots, p-1\} \quad (12)$$

The Legendre sequence can be interpreted as a sequence of numbers in modulo p indicating their quadratic residue property. The Legendre sequence for $p = 7$ is then $[0, 1, 1, -1, 1, -1, -1]$ and its energy is $E_c = p - 1$.

A relatively less known property of the Legendre sequences is their invariance under DFT operation. Stated differently, the Legendre sequences are the eigenvectors of the DFT matrix, [12, p.185]:

$$\text{DFT}\{b[n]\} = B[k] = \begin{cases} \sqrt{p} b[k] & p \equiv 1 \pmod{4} \\ -i\sqrt{p} b[k] & p \equiv 3 \pmod{4} \end{cases} \quad (13)$$

For $p = 7$, the DFT of the Legendre sequence is then $[0, -i\sqrt{7}, -i\sqrt{7}, i\sqrt{7}, -i\sqrt{7}, i\sqrt{7}, i\sqrt{7}]$. Comparing with (9), we can immediately note the optimality of Legendre sequences for the estimation of AC coupled channels. This result establishes the foundation of the proposed method.

We would like note that due to the eigen-relation with DFT, the estimate for AC components can be implemented very efficiently even without calculating forward and inverse DFT's. To estimate the l th tap, we only need to implement the circular convolution of $r[n]$ and $b[(n-l)_N]$ where $(n-l)_N$ denotes a circular shift of l units:

$$\hat{h}_{AC}[l] = \left(\frac{-1}{N}\right) \frac{1}{N} \sum_{n=0}^{N-1} r[n] b[(l-n)_N] \quad (14)$$

In the last equation, $l = \{0, 1, \dots, L-1\}$ and $\left(\frac{-1}{N}\right)$ is the Legendre symbol of -1 which is equivalent to $(-1)^{\frac{N-1}{2}}$. Using this relation, the estimate for the AC part of the channel can be realized with a total of LN additions and no multiplications (apart from a constant scaling of $\frac{1}{N}$ which may not be necessary to implement in many systems).

III. TWO STAGE FREQUENCY DOMAIN CHANNEL ESTIMATION

The channel coefficient vector $\overline{\mathbf{h}}$, given in $\mathbf{r} = \overline{\mathbf{C}} \overline{\mathbf{h}} + \mathbf{w}$, can be written as a combination AC and DC components, $\overline{\mathbf{h}} = \overline{\mathbf{h}}_{AC} + h_{DC} \mathbf{1}$. The vector $\overline{\mathbf{h}}_{AC}$ is the AC part of $\overline{\mathbf{h}}$, that is the projection of $\overline{\mathbf{h}}$ onto the sub-space spanned by the vectors associated with the AC terms of the DFT matrix (or the mean subtracted version of $\overline{\mathbf{h}}$) and can be written as $\overline{\mathbf{h}}_{AC} = (\mathbf{I} - \frac{1}{N} \mathbf{1} \mathbf{1}^H) \overline{\mathbf{h}}$. The remaining unknown component is the average value of the channel coefficients, $h_{DC} = \frac{1}{N} \mathbf{1}^H \overline{\mathbf{h}} = \frac{1}{N} \sum_{n=0}^{L-1} h[n]$.

When $\overline{\mathbf{h}}$ is substituted with $\overline{\mathbf{h}} = \overline{\mathbf{h}}_{AC} + h_{DC} \mathbf{1}$ in the equation $\mathbf{r} = \overline{\mathbf{C}} \overline{\mathbf{h}} + \mathbf{w}$, we get the following equation system:

$$\mathbf{r} = \overline{\mathbf{C}} (\overline{\mathbf{h}}_{AC} + h_{DC} \mathbf{1}) + \mathbf{w} \quad (15)$$

Estimation of $\overline{\mathbf{h}}_{AC}$: The estimation of the AC part is identical to the process given in Section II. The equation (16), given on the next page, shows the estimator.

In equation (16), the matrix \mathbf{U}_{N-1} is defined in (2), $C[k]$ is the DFT of the training sequence. Similar to the conventional technique, the estimate for $\hat{h}_{AC}[n]$ for $n \geq L$ is nulled, that is $\hat{h}_{AC}[n] = 0$ for $n \geq L$. We denote the estimate after the nulling as $\hat{\mathbf{h}}_{AC}$. A computationally simpler implementation valid for the Legendre training sequences using the relation in (14) is shown in Figure 1.

Estimation of h_{DC} : The first stage of the proposed systems estimates the AC components of the channel. The first stage meets the lower bound on the estimation error variance, hence it is the optimal estimator. In the second stage, the optimal

$$\begin{aligned}\hat{\mathbf{h}}_{AC} &= \mathbf{U}_{N-1}(\mathbf{U}_{N-1}^H \bar{\mathbf{C}} \bar{\mathbf{C}}^H \mathbf{U}_{N-1})^{-1}(\mathbf{U}_{N-1}^H \bar{\mathbf{C}}^H) \underbrace{\mathbf{U}_{N-1} \mathbf{U}_{N-1}^H}_{\mathbf{I}} \mathbf{r} \\ &= \frac{1}{N} \mathbf{U}_{N-1} \text{diag}\left(\frac{1}{C[1]}, \frac{1}{C[2]}, \dots, \frac{1}{C[N-1]}\right) \mathbf{U}_{N-1}^H \mathbf{r}\end{aligned}\quad (16)$$

estimate is inserted in (15) and the remaining unknown is solved to minimize the least square error as shown below:

$$\underbrace{\mathbf{r} - \mathbf{C}\hat{\mathbf{h}}_{AC}}_{\mathbf{r}_{DC}} = \underbrace{(\mathbf{C}\mathbf{1}_L)}_{\mathbf{v}_{DC}} h_{DC} + \mathbf{w} \quad (17)$$

Here \mathbf{r}_{DC} is the residual of the observation vector once the contribution of the AC part is subtracted. The least square (LS) solution of the resultant equation system can be written as follows:

$$\hat{h}_{DC} = \frac{\mathbf{v}_{DC}^H \mathbf{r}_{DC}}{\mathbf{v}_{DC}^H \mathbf{v}_{DC}} \quad (18)$$

Briefly postponing the discussion of the optimality of the least squares solution, we note that the final estimate for the channel coefficients is written as $\hat{h}[n] = \hat{h}_{AC}[n] + \hat{h}_{DC}$, $n = \{0, 1, \dots, L-1\}$.

Total Error Variance of the Proposed Method: The estimate $\hat{\mathbf{h}}_{AC}$ formed at the end of the first stage is the minimum variance unbiased estimate of \mathbf{h}_{AC} . If we let $\boldsymbol{\epsilon}_{AC} = \hat{\mathbf{h}}_{AC} - \mathbf{h}_{AC}$, then $\boldsymbol{\epsilon}_{AC} = \underbrace{\mathbf{F}_L^{-1} \boldsymbol{\Lambda}_C \mathbf{F}}_{\mathbf{M}} \mathbf{w} = \mathbf{M}\mathbf{w}$ where \mathbf{F} is $N \times N$

DFT matrix and \mathbf{F}_L^{-1} is the first L rows of the inverse DFT matrix and $\boldsymbol{\Lambda}_C = \text{diag}(0, 1/C[1], 1/C[2], \dots, 1/C[N-1])$.

The error on the DC term can be expressed as a combination of two sources. The first contribution is due to \mathbf{w} , which is the white noise on \mathbf{r} . The second contribution is due to the estimation error on $\hat{\mathbf{h}}_{AC}$, which can be written as $\boldsymbol{\epsilon}_{AC} = \mathbf{M}\mathbf{w}$. Then the error on the residual observation vector \mathbf{r}_{DC} , defined in (17), can be written as $\mathbf{w} - \mathbf{C}\boldsymbol{\epsilon}_{AC} = (\mathbf{I} - \mathbf{C}\mathbf{M})\mathbf{w}$.

The estimation error for the DC term, $\epsilon_{DC} = \hat{h}_{DC} - h_{DC}$, can be written as $\epsilon_{DC} = \underbrace{\frac{\mathbf{v}_{DC}^H}{\|\mathbf{v}_{DC}\|^2} (\mathbf{I} - \mathbf{C}\mathbf{M})}_{\mathbf{k}^H} \mathbf{w} = \mathbf{k}^H \mathbf{w}$

where \mathbf{v}_{DC} is defined in (17).

Finally, the estimation error vector for the channel coefficients is $\boldsymbol{\epsilon} = \boldsymbol{\epsilon}_{AC} + \epsilon_{DC} \mathbf{1}_L = (\mathbf{M} + \mathbf{1}_L \mathbf{k}^H) \mathbf{w}$ and has the total error variance of

$$\begin{aligned}E\{\boldsymbol{\epsilon}^H \boldsymbol{\epsilon}\} &= \sigma_w^2 [\text{trace}\{\mathbf{M}\mathbf{M}^H\} + L\mathbf{k}^H \mathbf{k} \\ &\quad + 2\text{real}\{\mathbf{k}^H \mathbf{M}^H \mathbf{1}_L\}].\end{aligned}\quad (19)$$

An Efficient Implementation: The proposed method can be very efficiently implemented in time domain by utilizing the eigenrelation of the Legendre sequences with DFT. The filtering diagram in Figure 1 shows the suggested implementation of the proposed method for the Legendre sequences. The given implementation has at most $2N$ multiplication requirements. (The matrix $\bar{\mathbf{C}}$ is composed of 1, -1 and 0's and therefore does not have any multiplication requirements; N multiplications are required for the scaling with $1/N$; at most N multiplications are needed to calculate the inner-product required for \hat{h}_{DC} .) Hence the suggested implementation has

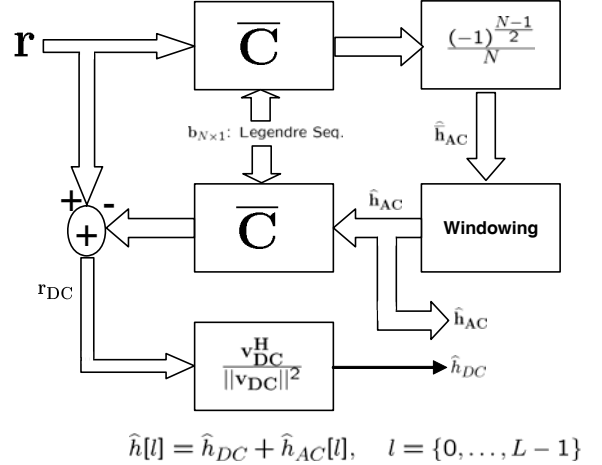


Fig. 1. Suggested implementation for the proposed method with the Legendre Sequences. \mathbf{r} is the input to channel estimation system after the removal of cyclic prefix. $\bar{\mathbf{C}}$ is $N \times N$ circular convolution matrix with the Legendre sequence \mathbf{b} . Windowing block multiplies the first L components of the input by 1 and others by 0. The vectors \mathbf{r}_{DC} , \mathbf{v}_{DC} are defined in (17). As an example, for $N = 5$ and $L = 3$, the vectors shown are $\mathbf{b} = [0, 1, -1, -1, 1]$ and $\mathbf{v}_{DC} = [0, 2, 0, -1, -1]$.

the multiplication requirement of $O(N)$ multipliers instead of $O(N \log N)$ multipliers as required for the DFT domain implementation.

On the optimality of proposed method: It should be noted that the estimates for \mathbf{h}_{AC} and h_{DC} are both unbiased and therefore $\hat{\mathbf{h}} = \hat{\mathbf{h}}_{AC} + \hat{h}_{DC} \mathbf{1}_L$ is also an unbiased estimate of \mathbf{h} . But different from $\hat{\mathbf{h}}_{AC}$, \hat{h}_{DC} is not a minimum variance unbiased estimate; but only the least squares estimate. The minimum variance unbiased estimate for h_{DC} is the weighted least squares estimate of \mathbf{r}_{DC} where the weighting is the covariance matrix of the error on \mathbf{r}_{DC} . We have chosen to use the simpler least squares estimate, since the resultant coloring of the noise on \mathbf{r}_{DC} is not significant due to good cyclic autocorrelation of the Legendre sequence [11].

IV. NUMERICAL RESULTS

Results on AC Coupled Channel Estimation: The performance of Legendre sequences, Schroeder sequences [12] and the optimal binary sequences for the conventional frequency domain channel estimation given in [2] are compared.

The Schroeder sequences of length N are defined as $b_n = 1 - 2 \left(\left\lfloor \frac{n^2}{2N} \right\rfloor \bmod 2 \right)$ for $n = \{1, \dots, N\}$. Here $\lfloor \cdot \rfloor$ denotes the rounding of the argument to the largest integer not exceeding the argument.

The optimal sequences of length $N = \{29, 31, 37, 41\}$ (which are listed in Table VI of [2]) are compared with the same length Legendre sequences. We would like to remind that the optimal sequences listed in [2] are optimal only for the

estimation arbitrary channels with the conventional method, not for the estimation of AC coupled channels. We denote the sequences described in [2] as Tellambura sequences in order not to cause confusion. It has been shown that the Legendre sequences are the optimal training sequences for the estimation of AC coupled channels. Our goal is to examine the loss incurred for another sequence.

The total estimation error variance for the AC coupled channels is defined as follows:

$$E\{||\mathbf{h} - \hat{\mathbf{h}}||^2\} = \sigma_w^2 \frac{L}{N} \sum_{k=1}^{N-1} \frac{1}{|B(k)|^2} \quad (20)$$

Note that the summation given in (20) starts from 1 for the AC coupled channel estimation. We use the gain-loss factor defined by Tellambura et al. in [2] for comparison. The gain-loss factor is the ratio of the error variance achieved by a training sequence and the lower bound on the variance. The gain-loss factor shows the amount of signal energy that should be increased to compensate the non-optimality of the code. Combining (20) and the lower bound given in (10), the gain-loss factor for AC coupled channels can be written as follows:

$$\mu = 10 \log_{10} \left(\frac{E_c N}{(N-1)^2} \sum_{k=1}^{N-1} \frac{1}{|B(k)|^2} \right) \text{ (dB)} \quad (21)$$

Table I shows the gain-loss factor for Schroeder, Tellambura and Legendre sequences each having the total energy of $E_c = N$. It should be noted that the gain-loss factors for Tellambura sequences given in Table I are higher than the ones reported in [2]. This is due to the change in the lower bound and the non-optimality of Tellambura sequences for the AC coupled channel estimation.

Results on Arbitrary Channel Estimation: The gain-loss factor of the two-stage method with the Legendre sequences is compared with the conventional method utilizing the optimal sequences. The gain-loss factor for arbitrary channel estimation is defined as follows,

$$\mu = 10 \log_{10} \left(\frac{E\{||\mathbf{h} - \hat{\mathbf{h}}||^2\}}{\sigma_w^2 L/N} \right) \text{ (dB)} \quad (22)$$

For the conventional frequency domain channel estimation method, the gain-loss factor can be expressed as $10 \log_{10} \left(\sum_{k=0}^{N-1} \frac{1}{|B(k)|^2} \right)$ (dB) and for the proposed scheme the gain-loss factor is equal to $10 \log_{10} \left(\frac{E\{\epsilon^H \epsilon\}}{\sigma_w^2 L/N} \right)$ (dB) where $E\{\epsilon^H \epsilon\}$ is defined in (19).

Table II shows the results for the conventional frequency domain channel estimation (single stage method) for Schroeder, Tellambura and the maximum length sequences. The Legendre sequences are not listed here due to their infinite gain-loss factor because of the null at the DC frequency. Table II presents the results for different training sequence lengths (N) and the channel lengths (L). Different from the conventional method, the error of the two-stage method depends on the channel length. This is due to the windowing stage implemented in between the estimation of AC and DC components (see Figure 1) and can be further explained with the change in the vector \mathbf{v}_{DC} given in (17) with L .

TABLE I
GAIN-LOSS FACTOR (μ) FOR THE AC COUPLED CHANNEL ESTIMATION

N	Schroeder	Tellambura	Legendre
	μ (dB)	μ (dB)	μ (dB)
29	1.81	0.72	0
31	1.75	0.43	0
37	1.7	0.47	0
41	2.42	0.42	0

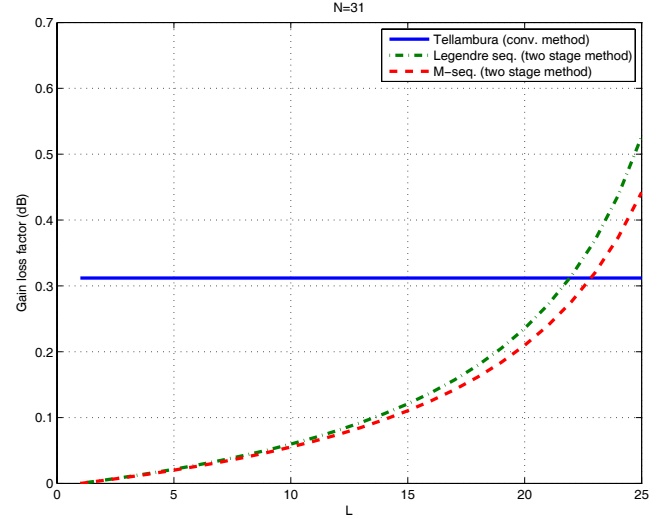


Fig. 2. The variation of the gain-loss factor for the conventional and suggested method. The channel length is L , the training sequence length is N .

Table II shows that the two-stage method with the Legendre sequences has a better gain-loss factor than the conventional method with the optimal sequences. In spite of some improvement in the estimation error variance, we believe that the major contribution of the proposed method is its ability to efficiently utilize the Legendre sequences. We would like to remind that the Legendre sequences have closed form formulas; hence can be immediately generated for any N , however large it is.

Figure 2 shows the gain-loss factor of the two-stage method with the Legendre, m-sequences and the gain-loss factor of the conventional method with optimal the Tellambura sequences for $N = 31$. It can be observed that two-stage method has a smaller estimation error variance than the conventional technique when L is sufficiently smaller than N , which is the case for the classical channel estimation problem. As expected, the performance of the two-stage method degrades as L approaches N . The m-sequences have a non-zero DC value, therefore operate at a marginally better overall MSE than the Legendre sequences. But m-sequences are defined only for $N = 2^k - 1$, where k is a positive integer, therefore they are quite rare in comparison to the Legendre sequences and do not have an efficient implementation. Therefore we suggest the usage of the Legendre sequences instead of m-sequences with the proposed method.

Figure 3 graphically illustrates the major contribution of the study. This figure shows the performance of the proposed method for $N = \{31, 41, 53, 61\}$ for various channel lengths L . It should be noted that a good training sequence for two-stage method for $N = 61$ is immediately available from (12);

TABLE II
GAIN-LOSS FACTOR (μ) FOR THE SINGLE STAGE (CONVENTIONAL) AND TWO-STAGE FREQUENCY DOMAIN CHANNEL ESTIMATION METHODS (N: TRAINING SEQUENCE LENGTH, L: CHANNEL LENGTH)

N	Single Stage Method			Two-Stage Method			
	Schroeder	Tellambura Independent of L	m-sequences	Legendre			
				L=5	L=10	L=15	L=20
29	1.57	0.50	-	0.02	0.07	0.15	0.30
31	1.80	0.31	2.87	0.02	0.06	0.12	0.23
37	1.79	0.32	-	0.01	0.04	0.08	0.13
41	2.23	0.29	-	0.01	0.03	0.05	0.10

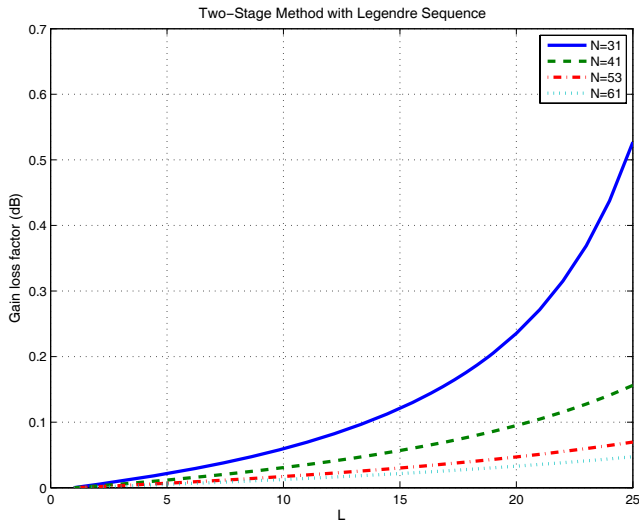


Fig. 3. The variation of the gain-loss factor for the two-stage method. As the training sequence length (N) increases, the error reduces monotonically for all channel lengths (L).

but it is unknown for the conventional method and requires a search of 2^{61} candidate points. The search complexity, even with the significant complexity reduction proposed by Tellambura *et al.*, is infeasible for $N \geq 43$ as reported in [2]. The availability of long training sequences can be especially important for the estimation long channels such as ultra wide band, underwater acoustic channels, etc.

V. CONCLUSIONS

The present work describes a novel two-stage channel estimation method. Different from the conventional method, the proposed method estimates AC and DC parts of the channel in succession. It has been shown that this approach leads to a reduction in the estimation error variance. The major advantage of the proposed method is the availability of the *optimal codes with analytical formulas* (Legendre sequences)

for the estimation of AC components of the channel. The conventional frequency domain channel estimation method requires a search of 2^N binary sequences to find an optimal training sequence of length N which is not feasible for large N . The proposed method also has an efficient time-domain implementation which requires at most $2N$ multipliers in comparison to $O(N \log N)$ multipliers for the conventional method. Hence the method enables efficient estimation of long channels, such as the ultra wide band channels, without a performance compromise (even with some gain) through a family of analytically expressible codes.

REFERENCES

- [1] S. M. Kay, *Fundamentals of Statistical Signal Processing, Vol. 1: Estimation Theory*. Prentice Hall, 1988.
- [2] C. Tellambura, M. Parker, Y. Guo, S. Shepherd, and S. Barton, "Optimal sequences for channel estimation using discrete Fourier transform techniques," *IEEE Trans. Commun.*, vol. 47, no. 2, pp. 230-238, Feb. 1999.
- [3] H. Minn, X. Fu, and V. Bhargava, "Optimal periodic training signal for frequency offset estimation in frequency-selective fading channels," *IEEE Trans. Commun.*, vol. 54, no. 6, pp. 1081-1096, June 2006.
- [4] H. Minn and N. Al-Dhahir, "Optimal training signals for MIMO OFDM channel estimation," *IEEE Trans. Wireless Commun.*, vol. 5, no. 5, pp. 1158-1168, May 2006.
- [5] Y. Wang and X. Dong, "Frequency-domain channel estimation for SC-FDE in UWB communications," *IEEE Trans. Commun.*, vol. 54, no. 12, pp. 2155-2163, 2006.
- [6] Z. Zhang, W. Zhang, and C. Tellambura, "MIMO-OFDM channel estimation in the presence of frequency offsets," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2329-2339, June 2008.
- [7] W. Chen and U. Mitra, "Training sequence optimization: Comparisons and an alternative criterion," *IEEE Trans. Commun.*, vol. 48, no. 12, pp. 1987-1991, Dec. 2000.
- [8] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*. Cambridge University Press, 2005.
- [9] R. Kumanduri and C. Romero, *Number Theory with Computer Applications*. Prentice Hall, 1997.
- [10] S.-C. Pei, C.-C. Wen, and J.-J. Ding, "Closed-form orthogonal DFT eigenvectors generated by complete generalized Legendre sequence," *IEEE Trans. Circuits Syst. I*, vol. 55, no. 11, pp. 3469-3479, 2008.
- [11] M. Su and A. Winterhof, "Autocorrelation of Legendre-Sidelnikov sequences," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1714-1718, 2010.
- [12] M. Schroeder, *Number Theory in Science and Communication*. Springer, 2008.