This research has been supported in part by European Commission
FP6 IYTE-Wireless Project (Contract No: 017442)

Approximate Best Linear Unbiased Channel Estimation for Frequency Selective Multipath Channels with Long Delay Spreads

Serdar Özen*, Christopher Pladdy*, Mark J. Fimoff*, S. M. Nerayanuru* and Michael D. Zoltowski†
*Zenith Electronics Corp., R & D
Lincolnshire, IL 60069 - USA
Phone: (847) 941 8797
{serdar.ozen, christopher.pladdy, mark.fimoff, snerayanuru}@zenith.com
†School of Electrical & Computer Engineering,
Purdue University
W. Lafayette, IN 47907-1285 - USA
Phone: (765) 494-3512
mikedz@ecn.purdue.edu

***Abstract*— We provide an iterative and a non-iterative channel impulse response (CIR) estimation algorithm for communication systems which utilize a periodically transmitted training sequence within a continuous stream of information symbols. The iterative procedure calculates the (semi-blind) *Best Linear Unbiased Estimate* (BLUE) of the CIR. The non-iterative version is an approximation to the BLUE CIR estimate, achieving almost similar performance, with much lower complexity. We first provide a formulation of the received data and correlation processing with the adjacent symbol correlation taken into account, and we then present the connections of the correlation based CIR estimation scheme to the ordinary least squares CIR estimation, and the BLUE CIR estimation. Simulation results are provided to demonstrate the performance of the novel algorithms for 8-VSB ATSC Digital TV system.**

I. Introduction

For the communications systems utilizing a periodically transmitted training sequence, *least-squares* (LS) based channel estimation or the *correlation* based channel estimation algorithms have been the most widely used two alternatives [1]. Both methods use a stored copy of the known transmitted training sequence at the receiver. The properties and the length of the training sequence are generally different depending on the particular communication system's standard specifications. However the accuracy of most channel estimation schemes is degraded due to the *baseline noise* term which occurs due to the correlation of the stored copy of the training sequence with the unknown symbols adjacent to transmitted training sequence, as well as the additive channel noise [1], [10]. In the sequel, we provide (semi-blind) *Best Linear Unbiased Estimate* (BLUE) and approximate BLUE (a-BLUE) channel estimators for communication systems using a periodically transmitted training sequence. Although the examples following the derivations of the BLUE and the a-BLUE channel estimators will be drawn from the ATSC digital TV 8-VSB system [2], to the best of our knowledge it could be applied with minor modifications to any digital communication system with linear modulation which employs a periodically transmitted training sequence. The novel algorithm presented in the sequel is targeted for the systems that are desired to work with channels having long delay spreads L_d; in particular we consider the case where $(NT+1)/2 < L_d < NT$, where NT is the duration of the available training sequence. For instance the 8-VSB digital TV system has 728 training symbols, whereas the delays spreads of the terrestrial channels have been observed to be at least 400-500 symbols long [6], [7], [8]. In addition to the iterative BLUE algorithm we provide approximate BLUE algorithm which can be used as an initializer to the BLUE iterations, or as a stand-alone alternative approach that produces results of nearly the same quality as the results produced by the BLUE algorithm while at the same time requiring much less computational complexity (i.e., requiring about the same number of multiplications necessary to implement ordinary least squares) and having storage requirements similar to that of ordinary least squares.

Our novel CIR estimation algorithms can be considered as *semi-blind* techniques since these methods take advantage of the statistics of the data [3], [8].

A. *Overview of Generalized Least Squares*

Consider the linear model

$$\boldsymbol{y} \;=\; \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\nu} \tag{1}$$

where $\boldsymbol{y}$ is the observation (or response) vector, $\boldsymbol{A}$ is the regression (or design) matrix, $\boldsymbol{x}$ is the vector of unknown parameters to be estimated, and $\boldsymbol{\nu}$ is the observation noise (or measurement error) vector. Assuming that it is known that the random noise vector $\boldsymbol{\nu}$ is zero mean, and is correlated, that is $\text{Cov}\{\boldsymbol{\nu}\} = \boldsymbol{K}_\nu \equiv \frac{1}{2}E\{\boldsymbol{\nu}\boldsymbol{\nu}^H\} \neq \sigma_{\boldsymbol{\nu}}^2\boldsymbol{I}$, we define the (generalized) objective function for the model of (1) by

$$J_{GLS}(\boldsymbol{x}) \;=\; (\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x})^H\boldsymbol{K}_\nu^{-1}(\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x}). \tag{2}$$

The least squares estimate that minimizes Equation (2) is

$$\hat{\boldsymbol{x}}_{gls} = (\boldsymbol{A}^H \boldsymbol{K}_\nu^{-1} \boldsymbol{A})^{-1} \boldsymbol{A}^H \boldsymbol{K}_\nu^{-1} \boldsymbol{y}, \quad (3)$$

The estimator of (3) is called the *best linear unbiased estimate* (BLUE) [9] among all *linear* unbiased estimators if the noise covariance matrix is *known* to be $\text{Cov}\{\boldsymbol{\nu}\} = \boldsymbol{K}_\nu$. If the noise $\boldsymbol{\nu}$ is *known* to be *Gaussian* with zero mean and with covariance matrix $\boldsymbol{K}_\nu$, that is if it is known that $\boldsymbol{\nu} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{K}_\nu)$, then the estimator of (3) is called the *minimum variance unbiased estimator* (MVUE) among *all* unbiased estimators (not only linear).

II. Overview of Data Transmission Model

The baseband symbol rate sampled receiver pulse-matched filter output is given by

$$\begin{aligned} y[n] &\equiv y(t)|_{t=nT} = \sum_k I_k h[n-k] + \nu[n] \\ &= \sum_k I_k h[n-k] + \sum_k \eta[k] q^*[-n+k], \end{aligned} \quad (4)$$

where

$$I_k = \left\{ \begin{array}{ll} a_k, & 0 \le k \le N-1 \\ d_k, & N \le n \le N'-1, \end{array} \right\} \in \mathcal{A} \equiv \{\alpha_1, \cdots, \alpha_M\} \quad (5)$$

is the M-ary complex valued transmitted sequence, $\mathcal{A} \subset \mathbb{C}^1$, and $\{a_k\} \in \mathbb{C}^1$ denote the first N symbols within a *frame* of length N' to indicate that they are the known training symbols; $\nu[n] = \eta[n] * q^*[-n]$ denotes the complex (colored) noise process after the (pulse) matched filter, with $\eta[n]$ being a zero-mean white Gaussian noise process with variance σ_η^2 per real and imaginary part;

$$h(t) = q(t) * c(t) * q^*(-t) = \sum_{k=-K}^{L} c_k p(t - \tau_k)$$

is the complex valued impulse response of the composite channel, including pulse shaping transmit filter $q(t)$, the physical channel impulse response $c(t)$, and the receive filter $q^*(-t)$, and $p(t) = q(t) * q^*(-t)$ is the convolution of the transmit and receive filters where $q(t)$ has a finite support of $[-T_q/2, T_q/2]$, and the span of the transmit and receive filters, T_q, is an even multiple of the symbol period, T; that is $T_q = N_q T$, $N_q = 2L_q \in \mathbb{Z}^+$. $\{c_k\} \subset \mathbb{C}^1$ denote complex valued physical channel gains, and $\{\tau_k\}$ denote the multipath delays, or the Time-Of-Arrivals (TOA). It is assumed that the time-variations of the channel is slow enough that $c(t)$ can be assumed to be a static inter-symbol interference (ISI) channel, at least throughout the training period. Without loss of generality, the symbol rate sampled composite CIR, $h[n]$, can be written as a finite dimensional vector

$$\boldsymbol{h} = [h[-N_a], \cdots, h[-1], h[0], h[1], \cdots, h[N_c]]^T \quad (6)$$

where N_a and N_c denote the number of anti-causal and causal taps of the channel, respectively, and $L_d = (N_a + N_c + 1)T$ is the delay spread of the channel (including the pulse tails). The pulse matched filter output which includes *all* the contributions from the known training symbols (which includes the adjacent random data as well) can be written as

$$\begin{aligned} \boldsymbol{y}_{[-N_a:N+N_c-1]} &= (\boldsymbol{A} + \boldsymbol{D})\,\boldsymbol{h} + \boldsymbol{\nu}_{[-N_a:N+N_c-1]} \\ &= \boldsymbol{A}\boldsymbol{h} + \boldsymbol{D}\boldsymbol{h} + \boldsymbol{Q}\boldsymbol{\eta}_{[-N_a-L_q:N+N_c-1+L_q]} \quad (7) \\ &= \boldsymbol{A}\boldsymbol{h} + \boldsymbol{H}\boldsymbol{d} + \boldsymbol{Q}\boldsymbol{\eta}_{[-N_a-L_q:N+N_c-1+L_q]} \quad (8) \end{aligned}$$

where

$$\boldsymbol{A} = \mathcal{T}\left\{ [a_0, \cdots, a_{N-1}, \underbrace{0, \cdots, 0}_{N_a+N_c}]^T, [a_0, \underbrace{0, \cdots, 0}_{N_a+N_c}] \right\}, \quad (9)$$

$$\begin{aligned} \boldsymbol{D} = \mathcal{T}\{&[\underbrace{0, \cdots, 0}_{N}, d_N, \cdots, d_{N_c+N_a+N-1}]^T, \\ &[0, d_{-1}, \cdots, d_{-N_c-N_a}]\}, \end{aligned} \quad (10)$$

where $\boldsymbol{A}$ is a Toeplitz matrix of dimension $(N + N_a + N_c) \times (N_a + N_c + 1)$ with first column $[a_0, a_1, \cdots, a_{N-1}, 0, \cdots, 0]^T$, and first row $[a_0, 0, \cdots, 0]$, and $\boldsymbol{D}$ is a Toeplitz matrix which includes the adjacent unknown symbols, prior to and after the training sequence. The data sequence $[d_{-1}, \cdots, d_{-N_c-N_a}]$ is the unknown information symbols transmitted at the end of the frame prior to the current frame being transmitted. $\boldsymbol{Q}$ is of dimension $(N + N_a + N_c) \times (N + N_a + N_c + N_q)$ and is given by

$$\boldsymbol{Q} = \begin{bmatrix} \boldsymbol{q}^T & 0 & \cdots & 0 \\ 0 & \boldsymbol{q}^T & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{q}^T \end{bmatrix} \quad (11)$$

and

$$\boldsymbol{q} = [q[+L_q], \cdots, q[0], \cdots, q[-L_q]]^T$$

is the receiver pulse matched filter, and

$$\boldsymbol{H} = \boldsymbol{\mathcal{H}}\boldsymbol{S}^T, \quad (12)$$

$$\bar{\boldsymbol{h}} = [h[N_c], \cdots, h[0], \cdots, h[-N_a]]^T = \boldsymbol{J}\boldsymbol{h}, \quad (13)$$

$$\boldsymbol{J} = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & \vdots & & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix}_{(N_a+N_c+1)\times(N_a+N_c+1)} \quad (14)$$

$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} \bar{\boldsymbol{h}}^T & 0 & \cdots & 0 \\ 0 & \bar{\boldsymbol{h}}^T & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \bar{\boldsymbol{h}}^T \end{bmatrix}_{(N+N_c+N_a)\times(N+2(N_a+N_c))} \quad (15)$$

and $\boldsymbol{d} = \boldsymbol{S}\tilde{\boldsymbol{d}}$, or equivalently $\tilde{\boldsymbol{d}} = \boldsymbol{S}^T\boldsymbol{d}$, where

$$\tilde{\boldsymbol{d}} = [d_{-N_c-N_a}, \cdots, d_{-1}, \boldsymbol{0}_{1\times N}, d_N, \cdots, d_{N+N_c+N_a-1}]^T \quad (16)$$

$$\boldsymbol{d} = [d_{-N-N_a}, \cdots, d_{-1}, d_N, \cdots, d_{N+N_c+N_a-1}]^T \quad (17)$$

$$\boldsymbol{S} = \begin{bmatrix} \boldsymbol{I}_{N_a+N_c} & \boldsymbol{0}_{(N_a+N_c)\times N} & \boldsymbol{0}_{(N_a+N_c)} \\ \boldsymbol{0}_{(N_a+N_c)} & \boldsymbol{0}_{(N_a+N_c)\times N} & \boldsymbol{I}_{N_a+N_c} \end{bmatrix} \quad (18)$$

where $\boldsymbol{S}$ is $(2(N_c + N_a)) \times (N + 2(N_a + N_c))$ dimensional *selection* matrix which retains the random data, eliminates the N zeros in the middle of the vector $\tilde{\boldsymbol{d}}$. where $\bar{\boldsymbol{h}}$ is the time reversed version of $\boldsymbol{h}$ (re-ordering is accomplished by the permutation matrix $\boldsymbol{J}$), and $\boldsymbol{H}$ is of dimension $(N + N_a + N_c) \times (2(N_c+N_a))$ with a "hole" inside which is created by the selection matrix $\boldsymbol{S}$.

III. Overview of the Proposed CIR Estimator

For comparison purposes we first provide the well known correlation and ordinary least squares based estimators, where correlations based estimation is denoted $\hat{\boldsymbol{h}}_u$ (the subscript $_u$ stands for the *uncleaned* CIR estimate) and is given by

$$\hat{\boldsymbol{h}}_u = \frac{1}{r_a[0]} \boldsymbol{A}^H \boldsymbol{y}_{[-N_a:N+N_c-1]}, \tag{19}$$

with $r_a[0] = \sum_{k=0}^{N-1} \|a_k\|^2$, and the ordinary least squares CIR estimate is denoted by $\hat{\boldsymbol{h}}_c$ (the subscript $_c$ stands for the *cleaned* CIR estimate) and is given by

$$\hat{\boldsymbol{h}}_c = (\boldsymbol{A}^H \boldsymbol{A})^{-1} \boldsymbol{A}^H \boldsymbol{y}_{[-N_a:N+N_c-1]}, \tag{20}$$

where "cleaning" is accomplished by removing the known sidelobes of the aperiodic correlation operation which is accomplished in (19).

We can denote the two terms on the right side of Equation (8) by $\boldsymbol{v} = \boldsymbol{H}\boldsymbol{d} + \boldsymbol{Q}\boldsymbol{\eta}_{[-N_a-L_q:N+N_c-1+L_q]}$. Hence we rewrite (8) as

$$\boldsymbol{y}_{[-N_a:N+N_c-1]} = \boldsymbol{A}\boldsymbol{h} + \boldsymbol{v}. \tag{21}$$

By noting the statistical independence of the random vectors $\boldsymbol{d}$ and $\boldsymbol{\eta}$, and also noting that both vectors are zero mean, the covariance matrix, $\boldsymbol{K_v}$ of $\boldsymbol{v}$ is given by

$$\text{Cov}\{\boldsymbol{v}\} = \boldsymbol{K_v} \equiv \frac{1}{2}E\{\boldsymbol{v}\boldsymbol{v}^H\} = \frac{\mathcal{E}_d}{2}\boldsymbol{H}\boldsymbol{H}^H + \sigma_\eta^2 \boldsymbol{Q}\boldsymbol{Q}^H, \tag{22}$$

where $\mathcal{E}_d$ is the energy of the transmitted information symbols, and equals to 21 if the symbols $\{d_k\}$ are chosen from the set $\{\pm1, \pm3, \pm5, \pm7\}$. For the model of (21) the generalized least squares objective function to be minimized is

$$J_{GLS}(\boldsymbol{h}) = \left(\boldsymbol{y}_{[-N_a:N+N_c-1]} - \boldsymbol{A}\boldsymbol{h}\right)^H \boldsymbol{K_v}^{-1} \left(\boldsymbol{y}_{[-N_a:N+N_c-1]} - \boldsymbol{A}\boldsymbol{h}\right) \tag{23}$$

Then the generalized least-squares solution to the model of Equation (21) which minimizes the objective function of $J_{GLS}(\boldsymbol{y})$ is given by

$$\hat{\boldsymbol{h}}_K = (\boldsymbol{A}^H \boldsymbol{K_v}^{-1} \boldsymbol{A})^{-1} \boldsymbol{A}^H \boldsymbol{K_v}^{-1} \boldsymbol{y}_{[-N_a:N+N_c-1]}. \tag{24}$$

The problem with Equation (24) is that the channel estimate $\hat{\boldsymbol{h}}_K$ is based on the covariance matrix $\boldsymbol{K_v}$, which is a function of the true channel impulse response vector $\boldsymbol{h}$ as well as the channel noise variance σ_η^2. In actual applications the BLUE channel estimate of Equation (24) can not be exactly obtained. Hence we need an *iterative* technique to calculate generalized least squares estimate of (24) where every iteration produces an updated estimate of the covariance matrix as well as the noise variance. Without going into the details, a simplified version of the iterations, which yield a closer approximation to the exact BLUE CIR estimate after each step, is provided in Algorithm 1. In the intermediate steps noise variance is estimated by $\sigma_\eta^2 = \frac{1}{2\mathcal{E}_q(N-N_a-N_c)}\|\hat{\boldsymbol{y}}_{[N_c:N-N_a]} - \boldsymbol{y}_{[N_c:N-N_a]}\|^2$, where $\mathcal{E}_q = \|\boldsymbol{q}\|^2$ and $\hat{\boldsymbol{y}}_{[N_c:N-N_a]} = \tilde{\boldsymbol{A}}\hat{\boldsymbol{h}}_{th}$, $\tilde{\boldsymbol{A}} = \mathcal{T}\left\{[a_{N_c+N_a}, \cdots, a_{N-1}]^T, [a_{N_c+N_a}, \cdots, a_0]\right\}$.

A. Approximate BLUE CIR estimation

An alternative approach may be used to produce results of nearly the same quality as the results produced by the algorithm described in Algorithm 1 while at the same time requiring much less computational complexity (i.e., requiring about the same number of multiplications necessary to implement Equation (20)) and having storage requirements similar to that of Equation (20). According to this alternative, the initial least squares estimation error can be reduced by seeking an approximation in which it is assumed that the baseband representation of the physical channel $c(t)$ is a distortion-free (no multipath) channel; that is $c(t) = \delta(t)$ which implies

$$h(t) = p(t) * c(t) = p(t). \tag{25}$$

Thus we can assume that our finite length channel impulse response vector can be (initially) approximated by

$$\tilde{\boldsymbol{h}} = [\underbrace{0, \cdots, 0}_{N_a - N_q}, \underbrace{p[-N_q], \cdots, p[0], \cdots, p[N_q]}_{\text{raised cosine pulse}}, \underbrace{0, \cdots, 0}_{N_c - N_q}]^T \tag{26}$$

with the assumptions of $N_a \geq N_q$ and $N_c \geq N_q$, that is the tail span of the composite pulse shape is well confined to within the assumed delay spread of $[-N_aT, N_cT]$. Then the approximation of (26) can be substituted into Equations (12-18) to yield an initial (approximate) channel convolution matrix $\tilde{\boldsymbol{H}}$ and is given by $\tilde{\boldsymbol{H}} = \tilde{\boldsymbol{\mathcal{H}}}\boldsymbol{S}^T$ where $\tilde{\boldsymbol{\mathcal{H}}}$ is formed as in Equation (15) with $\bar{\tilde{\boldsymbol{h}}} = \boldsymbol{J}\tilde{\boldsymbol{h}}$. We can also assume a reasonable received Signal-to-Noise (SNR) ratio measured at the input to the matched filter which is given by

$$\text{SNR} = \frac{\mathcal{E}_d \|(c(t) * q(t))|_{t=nT}\|^2}{\sigma_\eta^2} = \frac{\mathcal{E}_d \|\boldsymbol{q}\|^2}{\sigma_\eta^2}. \tag{27}$$

For instance we can assume an approximate SNR of 20dB yielding an initial noise variance of

$$\tilde{\sigma}_\eta^2 = \frac{\mathcal{E}_d \|\boldsymbol{q}\|^2}{100}. \tag{28}$$

Then combining $\tilde{\boldsymbol{H}}$ and $\tilde{\sigma}_\eta^2$ we can pre-calculate the initial approximate covariance matrix where the covariance matrix of the approximate channel is given by

$$\tilde{\boldsymbol{K}}_{\boldsymbol{v}}(\tilde{\boldsymbol{H}}) = \frac{1}{2}\mathcal{E}_d \tilde{\boldsymbol{H}}\tilde{\boldsymbol{H}}^H + \tilde{\sigma}_\eta^2 \boldsymbol{Q}\boldsymbol{Q}^H, \tag{29}$$

which further leads to the initial channel estimate of

$$\hat{\boldsymbol{h}}_{\tilde{K}} = \underbrace{\left(\boldsymbol{A}^H [\tilde{\boldsymbol{K}}_{\boldsymbol{v}}(\tilde{\boldsymbol{H}})]^{-1} \boldsymbol{A}\right)^{-1} \boldsymbol{A}^H [\tilde{\boldsymbol{K}}_{\boldsymbol{v}}(\tilde{\boldsymbol{H}})]^{-1}}_{\text{pre-computed and stored}} \boldsymbol{y}_{[-N_a:N+N_c-1]}. \tag{30}$$

Equation (30) is the resulting a-BLUE CIR estimate. The key advantage of the a-BLUE method is that the matrix $\left(\boldsymbol{A}^H[\tilde{\boldsymbol{K}}_{\boldsymbol{v}}(\tilde{\boldsymbol{H}})]^{-1}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^H[\tilde{\boldsymbol{K}}_{\boldsymbol{v}}(\tilde{\boldsymbol{H}})]^{-1}$ is constructed based on the initial assumptions that the receiver is expected to operate, and can be *pre-computed* and *stored* in the receiver. By using the pre-stored matrix $\left(\boldsymbol{A}^H[\tilde{\boldsymbol{K}}_{\boldsymbol{v}}(\tilde{\boldsymbol{H}})]^{-1}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^H[\tilde{\boldsymbol{K}}_{\boldsymbol{v}}(\tilde{\boldsymbol{H}})]^{-1}$ as in Equation (30) we obtain a CIR estimate with much lower computational complexity than the BLUE algorithm. We also note that a-BLUE CIR estimate can be used either as a stand-alone CIR estimator, or as an initial estimate which can be used by the BLUE algorithm.

Algorithm 1 Iterative Algorithm to obtain a CIR estimate via Generalized Least-Squares

[1] Get an initial CIR estimate using one of (19), (20), or (30), and denote it by $\hat{\boldsymbol{h}}[0]$;
[2] Threshold the initial CIR estimate, and denote it by $\hat{\boldsymbol{h}}^{(th)}[0]$;
[3] Estimate the noise variance $\widehat{\sigma^2}_\eta[0]$
[4]
for $k = 1, \ldots, N_{iter}$ **do**
  [4-a] Calculate the inverse of the (estimated) covariance matrix $\widehat{\boldsymbol{K}_{\boldsymbol{v}}}^{-1}[k] = \left[\frac{\mathcal{E}_d}{2}\boldsymbol{H}(\hat{\boldsymbol{h}}^{(th)}[k-1])\boldsymbol{H}^H(\hat{\boldsymbol{h}}^{(th)}[k-1]) + \widehat{\sigma^2}_\eta[k-1]\boldsymbol{Q}\boldsymbol{Q}^H\right]^{-1}$;
  [4-b] $\hat{\boldsymbol{h}}_K[k] = (\boldsymbol{A}^H\widehat{\boldsymbol{K}_{\boldsymbol{v}}}^{-1}[k]\boldsymbol{A})^{-1}\boldsymbol{A}^H\widehat{\boldsymbol{K}_{\boldsymbol{v}}}^{-1}[k]\boldsymbol{y}_{[-N_a:N+N_c-1]}$;
  [4-c] Threshold the CIR estimate $\hat{\boldsymbol{h}}_K[k]$, and denote it by $\hat{\boldsymbol{h}}^{(th)}[k]$;
  [4-d] Estimate the noise variance $\widehat{\sigma^2}_\eta[k]$.
end for

B. Analysis of Baseline Noise and CFAR Thresholding

The channel estimates $\hat{\boldsymbol{h}}_c$ or $\hat{\boldsymbol{h}}_{\tilde{K}}$ have contributions due to unknown symbols prior to and after the training sequence, which are elements of the vector $\boldsymbol{d}$, as well as the additive channel noise. These contributions due to unknown symbols and channel noise is called *baseline noise*, and we can give an expression which summarizes the baseline noise for two different estimators of Equations (20), and (30). The general channel estimate can be written in the form

$$\hat{\boldsymbol{h}} = \boldsymbol{h} + \boldsymbol{\xi} = \boldsymbol{h} + \boldsymbol{B}\Big(\boldsymbol{H}\boldsymbol{d} + \boldsymbol{Q}\boldsymbol{\eta}_{[-N_a-L_q:N+N_c-1+L_q]}\Big) \quad (31)$$

where the baseline noise vector $\boldsymbol{\xi}$ is defined by

$$\boldsymbol{\xi} \;=\; \boldsymbol{B}\left(\boldsymbol{H}\boldsymbol{d} + \boldsymbol{Q}\boldsymbol{\eta}_{[-N_a-L_q:N+N_c-1+L_q]}\right) \quad (32)$$

and the matrix $\boldsymbol{B}$ takes one of the two following different forms depending on the estimator used:

$$\boldsymbol{B} = \begin{cases} \left(\boldsymbol{A}^H\boldsymbol{A}\right)^{-1}\boldsymbol{A}^H, \text{ for } \hat{\boldsymbol{h}}_c \\ \left(\boldsymbol{A}^H[\tilde{\boldsymbol{K}}_{\boldsymbol{v}}(\tilde{\boldsymbol{H}})]^{-1}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^H[\tilde{\boldsymbol{K}}_{\boldsymbol{v}}(\tilde{\boldsymbol{H}})]^{-1}, \text{ for } \hat{\boldsymbol{h}}_{\tilde{K}} \end{cases} \quad (33)$$

Although we can derive the exact probability distribution of the baseline noise term, we can alternatively make the assumption of *normality* (having Gaussian distribution) of the baseline noise. This assumption can be asserted by invoking the central limit theorem[4]. The baseline noise vector $\boldsymbol{\xi}$ has covariance matrix $\boldsymbol{K}_\xi = \text{Cov}\{\boldsymbol{\xi}\}$

$$\boldsymbol{K}_\xi = \boldsymbol{B}(\frac{\mathcal{E}_d}{2}\boldsymbol{H}\boldsymbol{H}^H + \sigma_\eta^2\boldsymbol{Q}\boldsymbol{Q}^H)\boldsymbol{B}^H = \boldsymbol{B}\boldsymbol{K}_{\boldsymbol{v}}\boldsymbol{B}^H \quad (34)$$

where $\boldsymbol{K}_{\boldsymbol{v}}$ is given in (22), and we make the approximation

$$\boldsymbol{\xi} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{B}(\frac{\mathcal{E}_d}{2}\boldsymbol{H}\boldsymbol{H}^H + \sigma_\eta^2\boldsymbol{Q}\boldsymbol{Q}^H)\boldsymbol{B}^H) = \mathcal{N}(\boldsymbol{0}, \boldsymbol{B}\boldsymbol{K}_{\boldsymbol{v}}\boldsymbol{B}^H) \quad (35)$$

by invoking the central limit theorem, where $\boldsymbol{B}$ takes one of the appropriate forms as displayed in Equation (33).

We also provide the probability distribution of $\|\xi_k\|^2$ where subscript k denotes the kth element of the baseline noise vector $\boldsymbol{\xi} = [\xi_1, \ldots, \xi_{N_a+N_c+1}]^T$. Based on (35) we can show that ξ_k has a Gaussian marginal distribution with zero mean and variance[4]

$$\sigma_{\xi_k}^2 \equiv \frac{1}{2}E\{\xi_k\xi_k^*\} = \boldsymbol{1}_k^T\boldsymbol{B}\boldsymbol{K}_{\boldsymbol{v}}\boldsymbol{B}^H\boldsymbol{1}_k \quad (36)$$

that is $\xi_k = \boldsymbol{1}_k^T\boldsymbol{B}(\boldsymbol{H}\boldsymbol{d} + \boldsymbol{Q}\boldsymbol{\eta}_{[-N_a-L_q:N+N_c-1+L_q]})$, and

$$\xi_k \sim \mathcal{N}(0, \underbrace{\boldsymbol{1}_k^T\boldsymbol{B}\boldsymbol{K}_{\boldsymbol{v}}\boldsymbol{B}^H\boldsymbol{1}_k}_{\sigma_{\xi_k}^2}), \quad (37)$$

where $\boldsymbol{B}$ takes one of the appropriate forms as displayed in Equation (33), and $\boldsymbol{1}_k = [\underbrace{0, \ldots, 0}_{k-1}, 1, 0, \ldots, 0]^T$ is the vector of zeros of appropriate dimension with a 1 at the kth position.

Now we state an important fact about the probability distribution of the square-norm of the complex Gaussian random variables [11]. Let $\xi = \xi_r + j\xi_q$ be a complex valued random variable, with statistically independent real and imaginary parts ξ_r and ξ_q. Given that ξ is Gaussian with 0 mean and variance $\sigma_\xi^2 = \sigma_{\xi_r}^2 = \sigma_{\xi_q}^2 = \frac{1}{2}E\{\xi\xi^*\}$, the random variable defined by $Z = \|\xi\|^2 = \xi_r^2 + \xi_q^2$ is exponentially distributed, and its density is given by

$$p_Z(z) \;=\; \begin{cases} \frac{1}{2\sigma_\xi^2}e^{-\frac{z}{2\sigma_\xi^2}}, & r \geq 0 \\ 0, & r < 0. \end{cases} \quad (38)$$

Although it is apparent that the real and imaginary parts of the baseline noise ξ_k are not statistically independent, for the sake of obtaining a simple thresholding rule and for the special case of Digital TV system the correlation can be shown to be small, we will proceed as if the real and the imaginary parts of ξ_k are uncorrelated. With this simplified assumption $\|\xi_k\|^2$ is an exponentially distributed random variable with parameter $2\sigma_{\xi_k}^2$, and the density function is

$$p_{\|\xi_k\|^2}(r) \;=\; \begin{cases} \frac{1}{2\sigma_{\xi_k}^2}e^{-\frac{r}{2\sigma_{\xi_k}^2}}, & r \geq 0 \\ 0, & r < 0. \end{cases} \quad (39)$$

where $\sigma_{\xi_k}^2$ is defined by Equation (36).

There is one crucial detail for Algorithm 1 that has not been discussed until this point. Right after obtaining a channel estimate, prior to using that channel estimate for noise variance, σ_η^2, calculation and prior to building the channel convolution matrix $\boldsymbol{H}$, the baseline noise has to be cleaned from the channel estimate. This cleaning can be achieved via thresholding. Previously we have used a fixed thresholding algorithm [6] to get rid of the baseline noise. We have observed that there can be significant performance loss if a fixed thresholding is applied at every iteration. This performance loss is inevitable due to getting rid of significant amount of pulse tails embedded in the channel impulse response while getting rid of the baseline noise. To overcome this problem we propose constant false alarm[1] rate (CFAR) based thresholding, and it is based on determining a threshold based on the approximate statistical distribution of the baseline noise which is already provided in (37).

Recall that the kth tap of the channel estimate vector can be expressed in the form

$$\hat{h}_k = h_k + \underbrace{\mathbf{1}_k^T \boldsymbol{B}\left(\boldsymbol{H}\boldsymbol{d} + \boldsymbol{Q}\boldsymbol{\eta}_{[-N_a-L_q:N+N_c-1+L_q]}\right)}_{\xi_k}, \quad (40)$$

and ξ_k has a Gaussian distribution with zero mean and variance $\sigma^2_{\xi_k} = \mathbf{1}_k^T \boldsymbol{B}\boldsymbol{K_v}\boldsymbol{B}^H \mathbf{1}_k$ where $\boldsymbol{B}$ takes one of the appropriate forms as displayed in Equation (33), and the random variable $\|\xi_k\|^2$ is assumed to have exponential distribution with parameter $2\sigma^2_{\xi_k}$.

The problem of deciding whether the kth tap estimate $\hat{h}_k$ is a zero tap or not can be formulated as a simple hypothesis testing problem. That is we consider

$$H_0 \; : \; \hat{h}_k = \xi_k, \quad (41)$$
$$H_1 \; : \; \hat{h}_k = h_k + \xi_k; \quad (42)$$

where under H_0 the hypothesis is that the kth channel tap is actually zero and we are observing only baseline noise, and under H_1 the hypothesis is that the channel tap is non-zero, and we are observing (non-zero) channel tap plus the baseline noise. We have shown that the probability distributions of the kth channel tap under each hypothesis is given by

$$H_0 \; : \; \hat{h}_k \sim \mathcal{N}(0, \sigma^2_{\xi_k}), \quad (43)$$
$$H_1 \; : \; \hat{h}_k \sim \mathcal{N}(h_k, \sigma^2_{\xi_k}). \quad (44)$$

After defining (43) and (44) we can come up with different decision rules on how to threshold the channel estimate $\hat{\boldsymbol{h}}$, however we choose to pursue the constant false alarm rate (CFAR) based thresholding. False alarm probability based decision rule is chosen so that the resulting threshold rule does not require any a priori knowledge of the distribution of the hypothesis H_1, it is solely based on H_0. False alarm rate is the probability of choosing H_1 when H_0 is true. Our decision rule will be in the form of

$$\text{set } \hat{h}_k^{(th)} = \begin{cases} 0, & \text{if } \|\hat{h}_k\|^2 < \varepsilon_k \\ \hat{h}_k, & \text{otherwise.} \end{cases} \quad (45)$$

[1] In statistical inference literature false alarm (rate) is referred to as the Type I error (probability).

Based on the rule of (45) the false alarm rate, denoted by p_{FA} is given by

$$p_{FA} = \Pr\{\|\hat{h}_k\|^2 \geq \varepsilon_k | H_0 \text{ is true }\} = \int_{\varepsilon_k}^{\infty} \frac{1}{2\sigma^2_{\xi_k}} e^{-\frac{r}{2\sigma^2_{\xi_k}}} dr$$
$$= e^{-\frac{\varepsilon_k}{2\sigma^2_{\xi_k}}}. \quad (46)$$

For the given level of false alarm probability p_{FA} the threshold level ε_k is given by

$$\varepsilon_k = -2\sigma^2_{\xi_k} \ln(p_{FA}) \quad (47)$$

where $\sigma^2_{\xi_k}$ is given by (36).

Although we end up with a very simple expression for the threshold of Equation (47), which should be applied to the channel estimate as in (45), we still have the problem of not knowing the true covariance matrix $\boldsymbol{B}(\frac{1}{2}\mathcal{E}_d \boldsymbol{H}\boldsymbol{H}^H + \sigma^2_\eta \boldsymbol{Q}\boldsymbol{Q}^H)\boldsymbol{B}^H$ and the kth diagonal element which we have denoted by $\sigma^2_{\xi_k}$. We can only have an estimate $\widehat{\sigma^2}_{\xi_k}$ available to be used in Equation (47). Thus it is natural to see some performance loss due to using the estimate $\widehat{\sigma^2}_{\xi_k}$ in place of the true variance as will be shown in the simulations. Indeed the thresholding step is going to be incorporated into the iterations of the channel estimation with covariance matrix updated at every iteration. Once the covariance matrix is updated at every iteration we would have a new, and presumably better, threshold ε_k since we will get a better estimate $\widehat{\sigma^2}_{\xi_k}$ at every iteration.

Note that the step **[4-b]** of the Algorithm is the main step to compute the channel estimate, and is repeated here for convenience (the square bracketed index $[n]$ denote the iteration step):

$$\hat{\boldsymbol{h}}_K[n] = \left(\boldsymbol{A}^H \widehat{\boldsymbol{K}}_{\boldsymbol{v}}^{-1}[n]\boldsymbol{A}\right)^{-1} \boldsymbol{A}^H \widehat{\boldsymbol{K}}_{\boldsymbol{v}}^{-1}[n]\boldsymbol{y}_{[-N_a:N+N_c-1]}, \quad (48)$$

where $\widehat{\boldsymbol{K}}_{\boldsymbol{v}}^{-1}[n] = [\frac{\mathcal{E}_d}{2}\boldsymbol{H}(\hat{\boldsymbol{h}}_{(th)}[n-1])\boldsymbol{H}^H(\hat{\boldsymbol{h}}_{(th)}[n-1]) + \widehat{\sigma^2}_\eta[n-1]\boldsymbol{Q}\boldsymbol{Q}^H]^{-1}$ is the inverse of the estimated covariance matrix of $\boldsymbol{v}$, and $\boldsymbol{H}(\hat{\boldsymbol{h}}_{(th)}[n-1])$ is the convolution matrix (with a "hole" inside) constructed as in Equations (12-18) from $\hat{\boldsymbol{h}}_{(th)}[n-1]$ which is the thresholded CIR vector estimated at the previous iteration. The baseline noise for the main channel estimation step of Equation (48) is

$$\boldsymbol{\xi} = \boldsymbol{B}(\boldsymbol{H}\boldsymbol{d} + \boldsymbol{Q}\boldsymbol{\eta}_{[-N_a-L_q:N+N_c-1+L_q]}), \quad (49)$$

where $\boldsymbol{B} = (\boldsymbol{A}^H \widehat{\boldsymbol{K}}_{\boldsymbol{v}}^{-1}[n]\boldsymbol{A})^{-1}\boldsymbol{A}^H \widehat{\boldsymbol{K}}_{\boldsymbol{v}}^{-1}[n]$. Thus the covariance matrix of the baseline noise at the n'th iteration step, denoted by $\boldsymbol{K}_\xi[n]$, is given by

$$\boldsymbol{K}_\xi[n] = \boldsymbol{B}(\frac{\mathcal{E}_d}{2}\boldsymbol{H}\boldsymbol{H}^H + \sigma^2_\eta \boldsymbol{Q}\boldsymbol{Q}^H)\boldsymbol{B}^H. \quad (50)$$

Since we can only use an estimate of the true covariance matrix $\boldsymbol{K_v}$ (in the middle of Equation (50)), after the simplifications we get

$$\widehat{\boldsymbol{K}}_\xi[n] = (\boldsymbol{A}^H \widehat{\boldsymbol{K}}_{\boldsymbol{v}}^{-1}[n]\boldsymbol{A})^{-1}, \quad (51)$$

TABLE I

SIMULATED CHANNEL DELAYS IN SYMBOL PERIODS, RELATIVE GAINS. $L = -1$, $K = 6$, $L_d \approx (1 + 333 + 2N_q)T = 453T \approx 44\mu\text{SEC}$, $N_q = 60$.

Channel taps	Delay $\{\tau_k\}$	Gain $\{\lvert c_k \rvert\}$
$k = -1$	-0.957	0.7263
Main $k = 0$	0	1
$k = 1$	3.551	0.6457
$k = 2$	15.250	0.9848
$k = 3$	24.032	0.7456
$k = 4$	29.165	0.8616
$k = 5$	221.2345	0.6150
$k = 6$	332.9810	0.4900

which is an estimate of the true covariance matrix of $\boldsymbol{\xi}$ of Equation (49) if we could have used the true covariance matrix $\boldsymbol{K_v}$ in Equation (48) to begin with. Then the CFAR based threshold is given

$$\varepsilon_k = -2\widehat{\sigma^2}_{\xi_k} \ln(p_{FA}) \qquad (52)$$

where $\widehat{\sigma^2}_{\xi_k} = \mathbf{1}_k^T \widehat{\boldsymbol{K}}_{\xi}[n]\mathbf{1}_k = \mathbf{1}_k^T(\boldsymbol{A}^H \widehat{\boldsymbol{K}_{\boldsymbol{v}}}^{-1}[n]\boldsymbol{A})^{-1}\mathbf{1}_k$.

IV. SIMULATIONS

We considered an 8-VSB [2] receiver with a single antenna. 8-VSB system has a complex raised cosine pulse shape [2]. The CIR we considered is given in Table I. The phase angles of individual paths for all the channels are taken to be $\arg\{c_k\} = \exp(-j2\pi f_c \tau_k)$, $k = -1, \cdots, 6$ where $f_c = \frac{50}{T}$ and $T = 92.9$nsec. The simulations were run at 28dB Signal-to-Noise-Ratio (SNR) measured at the input to the receive pulse matched filter, and it is calculated by $\text{SNR} = \frac{\mathcal{E}_d \|\{c(t)*q(t)\}_{t=kT}\|^2}{\sigma_\eta^2}$. Figure 1 shows the simulation results for the test channel provided in Table I. Part (a) shows the actual CIR; part (b) shows the correlation based CIR estimate, of Equation (19) $\hat{\boldsymbol{h}}_u$; part (c) shows the ordinary LS based CIR estimate of Equation (20) $\hat{\boldsymbol{h}}_c$; part (d) shows the approximate BLUE CIR estimate of Equation (30) with an assumed SNR of 20dB; part (e) shows the BLUE based CIR estimate of Algorithm 1, after the first iteration, $\hat{\boldsymbol{h}}_K[1]$, where we used CFAR based thresholding with $p_{FA} = 10^{-5}$; part (f) shows the ideal BLUE case for which the true covariance matrix $\boldsymbol{K_v}$ is known. Part (f) provides a bound for the rest of the BLUE algorithm. We note superior performance of the BLUE algorithm even after the first iteration, as compared to the correlation based and ordinary least squares based CIR estimation schemes. However iterative BLUE CIR estimation algorithm is computationally very demanding, thus in many applications the approximate BLUE, as shown in part (d), might be sufficiently acceptable as an initial estimate. The performance measure is the normalized least-squares error which is defined by $\mathcal{E}_{NLS} = \frac{\|\boldsymbol{h} - \hat{\boldsymbol{h}}\|^2}{N_a + N_c + 1}$. Approximate BLUE significantly outperforms the ordinary least squares CIR estimation, but it has virtually identical computational complexity and storage requirement.

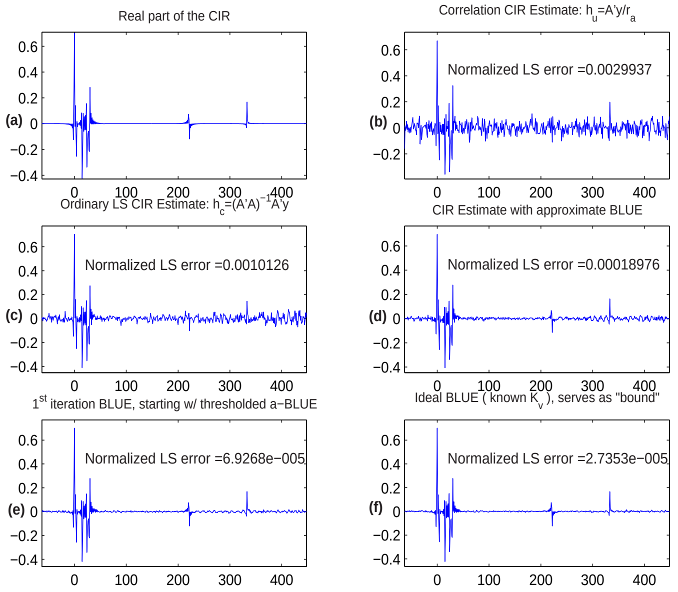

Fig. 1. Part (a) shows the real part of the actual CIR; part (b) shows the correlation based CIR estimate of Equation (19) $\hat{\boldsymbol{h}}_u$; part (c) shows the LS based CIR estimate of Equation (20) $\hat{\boldsymbol{h}}_c$; part (d) shows the approximate BLUE CIR estimate of Equation (30) with an assumed SNR of 20dB; part (e) show the BLUE based CIR estimate of Algorithm 1, after the first iteration, $\hat{\boldsymbol{h}}_K[1]$; part (f) shows the ideal BLUE case for which the true covariance matrix $\boldsymbol{K_v}$ is known, which provides a bound for the rest of the estimators.

ACKNOWLEDGMENTS

This research was funded in part by the NSF under grant number CCR0118842 and by Zenith Electronics Corporation.

REFERENCES

[1] H. Arslan, G. E. Bottomley, “Channel estimation in narrowband wireless communication systems,” Wireless Comunications and Mobile Computing, vol. 1, pp. 201-219, 2001.

[2] ATSC DTV Standard, A/53, September 1995 (http://www.atsc.org).

[3] E. de Carvalho, D. Slock, “Semi-blind methods for FIR multi-channel estimation,” Chapter 7 in *Signal Processing Advances in Wireless and Mobile Communications; Trends in Channel Estimation and Equalization*, G. Giannakis, Y. Hua, P. Stoica, L. Tong, vol. 1, Prentice-Hall, 2001

[4] G. Casella, R. L Berger, “*Statistical Inference*,” Duxbury Press, 1990.

[5] M. Fimoff, S. Özen, S. Nereyanuru, M.D. Zoltowski, W. Hillery, “Using 8-VSB Training Sequence Correlation as a Channel Estimate for DFE Tap Initialization,” Proceedings of the 39th Annual Allerton Conf. in Communications, Control and Computing, Sept. 2001.

[6] S. Özen, M. D. Zoltowski, M. Fimoff, “A Novel Channel Estimation Method: Blending Correlation and Least-Squares Based Approaches,” Proceedings of ICASSP, v. 3, pp. 2281-2284, 2002.

[7] S. Özen, W. Hillery, M. D. Zoltowski, S. M. Nerayanuru, M. Fimoff, “Structured Channel Estimation Based Decision Feedback Equalizers for Sparse Multipath Channels with Applications to Digital TV Receivers,” Conference Record of the 36th Asilomar Conference on Signals, Systems and Computers, v. 1, pp 558-564, Nov. 3-6, 2002.

[8] C. Pladdy, S. Özen, S. M. Nerayanuru, M. Fimoff, M. D. Zoltowski,“Best Linear Unbiased Channel Estimation for Frequency Selective Multipath Channels with Long Delay Spreads” to be published in the Proceedings of 58th Vehicular Technology Conference (VTC), Orlando, Florida 2003.

[9] G. A. F. Seber, *Linear Regresion Analysis*, John Wiley and Sons, 1977.

[10] H-K Song, “A channel estimation using sliding window aproach and tuning algorithm for MLSE,” IEEE Comm. Letters, v. 3, pp. 211-213, 1999.

[11] G. L. Stüber, “*Principles of Mobile Communication*,” Second Edition, Kluwer Academic Publishers, 2001.