

Taylor Series Approximation for Low Complexity Semi-Blind Best Linear Unbiased Channel Estimates for the General Linear Model with Applications to DTV

Christopher Pladdy¹, S.M. Nerayanuru¹, Mark Fimoff¹, Serdar Ozen², Michael Zoltowski³
¹Zenith R&D Center; ²Izmir Institute of Technology, Izmir, Turkey; ³Purdue University

Abstract— We present a low complexity approximate method for semi-blind best linear unbiased estimation (BLUE) of a channel impulse response vector (CIR) for a communication system which utilizes a periodically transmitted training sequence, within a continuous stream of information symbols. The algorithm achieves slightly degraded results at a much lower complexity than directly computing the BLUE CIR estimate. In addition, the inverse matrix required to invert the weighted normal equations to solve the general least squares problem may be precomputed and stored at the receiver. The BLUE estimate is obtained by solving the general linear model, $y = Ah + w + n$, for h , where w is correlated noise and the vector n is an AWGN process, which is uncorrelated with w . The solution is given by the Gauss-Markoff Theorem as $h = (A^T C(h)^{-1} A)^{-1} A^T C(h)^{-1} y$. In the present work we propose a Taylor series approximation for the function $F(h) = (A^T C(h)^{-1} A)^{-1} A^T C(h)^{-1} y$ where, $F : R^L \rightarrow R^L$ for each fixed vector of received symbols, y , and each fixed convolution matrix of known transmitted training symbols, A . We describe the full Taylor formula for this function, $F(h) = F(h_{id}) + \sum_{|\alpha| \geq 1} (h - h_{id})^\alpha (\partial/\partial h)^\alpha F(h_{id})$ and describe algorithms using, respectively, first, second and third order approximations. The algorithms give better performance than correlation channel estimates and previous approximations used, [15], at only a slight increase in complexity. The linearization procedure used is similar to that used in the linearization to obtain the extended Kalman filter, and the higher order approximations are similar to those used in obtaining higher order Kalman filter approximations, [7]

Keywords— Channel estimation, BLUE, best linear unbiased estimation, general linear model, Taylor series approximation, linearization, Gauss Markoff Theorem.

I. INTRODUCTION

Reliable communication often requires accurate estimation of the channel impulse response (CIR) to facilitate channel equalization. Semi-blind algorithms exploit information used by blind methods (for example, the statistics of the unknown data symbols) as well as information from known training symbols. For general references on blind and semiblind channel estimation, see [8]. Several recent papers consider different aspects of semi-blind channel estimation. Notably [6], [13], [12], and [4]. In [17] and [18] we devised a semi-blind iterative algorithm to construct the best linear unbiased estimate (BLUE) of the channel, which is given by the Gauss-Markoff Theorem ([10] or [14])

as

$$h_{BLUE} = (A^T C(h)^{-1} A)^{-1} A^T C(h)^{-1} y \quad (1)$$

in the case where we have the general linear model for the received data, i.e. where the noise is not white, since we consider correlated noise due to unknown data, where, in our case, the noise is correlated due to convolution with the CIR. For each fixed matrix of training data, A , and each vector of received values, y , we define the function

$$F(h) = F_{y,A}(h) = (A^T C(h)^{-1} A)^{-1} A^T C(h)^{-1} y \quad (2)$$

In [15] and [16] an approximate version of the iterative algorithm of [17] and [18] is described. In the present work we propose a more general framework, within which the approximation given in [15] and [16], would be the zeroth order Taylor series approximation of the function $F(h)$, i.e. a constant approximation, $F(h_{id})$, to the function $F(h)$. We consider the real part of the received data, denoted as vector y , and the real part of the CIR vector, denoted as h . We use the vector-valued function of a real vector variable version of the Taylor series, [3], to develop a series approximation to the function $F(h)$ which gives the BLUE estimate for h , as given in (1). Note that although our proposed algorithm does not fit within a Kalman filtering framework the process of linearization used in deriving the extended Kalman filter (EKF) is similar to that which we propose here, and the higher order approximations which we use are similar to those used in obtaining higher order Kalman filter approximations, [7]. In fact, the use of the Taylor series is a standard tool of approximation, see, for example, [5], [9], and [19] to see instances of its use. We describe the full Taylor formula for this function, (where $h = [h_1, \dots, h_L]^T$, $h_{id} = [h_1^{id}, \dots, h_L^{id}]^T$ and α is a multi-index, where for $\alpha = [0, \dots, 0, \underbrace{1}_{i^{th}}, 0, \dots, 0]^T$, $(h - h_{id})^\alpha =$

$$(h_i - h_i^{id}))$$

$$F(h) = F(h_{id}) + \sum_{|\alpha| \geq 1} (h - h_{id})^\alpha (\partial/\partial h)^\alpha F(h_{id}) \quad (3)$$

and use either a first (denoted $\nabla^1 F(h_{id})$) or second (denoted $\nabla^2 F(h_{id})$) order approximation,

$$F(h) \approx \nabla^1 F(h_{id}) \equiv F(h_{id}) + (h_i - h_i^{id}) (\partial/\partial h_i) F(h_{id}) \quad (4)$$

$$F(h) \approx \nabla^2 F(h_{id}) \equiv F(h_{id}) + \quad (5)$$

$$(h_i - h_i^{id}) (\partial/\partial h_i) F(h_{id}) + \quad (6)$$

$$(h_i - h_i^{id})^2 (\partial^2/\partial h_i^2) F(h_{id}) \quad (7)$$

II. BASEBAND DATA TRANSMISSION MODEL

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

$$y[n] \equiv y(t)|_{t=nT} = \sum_k I_k h[n-k] + \nu[n], \quad (8)$$

$$= \sum_k I_k h[n-k] + \sum_k \eta[k] q^*[-n+k] \quad (9)$$

where $I_k \in \{\alpha_1, \dots, \alpha_M\} \subset \mathbb{C}$ is the M-ary complex valued training sequence; $\{a_k\}$ denotes the first N known (training) symbols within a frame of length N' and $\{d_k\}$ denotes the remaining $N' - N$ random data within the frame; $\nu(t) = \eta(t) * q^*(-t)$ denotes the complex (colored) noise process at the output of the receiver (pulse) matched filter, with $\eta(t)$ being a zero-mean white Gaussian noise process with spectral density σ_η^2 per real and imaginary part; $h(t)$ is the complex valued impulse response of the composite dynamic channel, including the pulse shaping transmit filter $q(t)$, the physical channel impulse response $c(t)$, and the receive filter $q^*(-t)$ and is given by

$$h(t) = p(t) * c(t) = \sum_{k=-K}^L c_k p(t - \tau_k), \quad (10)$$

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 integer multiple of the symbol period, T ; that is $T_q = N_q T = 2L_q$, $N_q \in \mathbb{Z}^+$. $\{c_k\} \subset \mathbb{C}$ denote the complex valued physical channel gains, and $\{\tau_k\}$ denote the multipath delays, or the Time-Of-Arrivals (TOA). We also note that for the 8-VSB system, the transmitter pulse shape is the Hermitian symmetric root raised cosine pulse, which implies that $q(t) = q^*(-t)$. In the sequel, we will denote both the transmit and receive filters by $q[n] \equiv q(t)|_{t=nT}$. Also the sampled matched filter output signal, $y[n]$, will be used extensively in vector form, and so we introduce the notation $y_{[n_1:n_2]} = [y[n_1], \dots, y[n_2]]^T \in \mathbb{R}^{n_2-n_1+1}$. Similarly $\eta_{[n_1:n_2]} = [\eta[n_1], \dots, \eta[n_2]]^T \in \mathbb{R}^{n_2-n_1+1}$ and $\nu_{[n_1:n_2]} = [\nu[n_1], \dots, \nu[n_2]]^T \in \mathbb{R}^{n_2-n_1+1}$.

Without loss of generality, the symbol rate sampled, complex valued composite CIR, $h[n]$, can be written as a finite dimensional vector

$$h = [h[-N_a], \dots, h[-1], h[0], h[1], \dots, h[N_c]]^T$$

where N_a and N_c denote the number of anticausal and causal taps respectively. Based on (8), and assuming that $N \geq (N_a + N_c + 1)$, we can write the pulse matched filter

output corresponding *only to the known training symbols* compactly, in vector notation, as,

$$\begin{aligned} y_{[N_c:N-N_a-1]} &= \tilde{A}h + \nu_{[N_c:N-N_a-1]} \\ &= \tilde{A}h + \tilde{Q}\eta_{[N_c-L_q:N-N_a-1+L_q]} \end{aligned}$$

where $\tilde{A} = T \left\{ [a_{N_n+N_a}, \dots, a_{N-1}]^T, [a_{N_n+N_a}, \dots, a_0] \right\}$ is the $(N - N_a - N_c) \times (N_a + N_c + 1)$ Toeplitz convolution matrix with first column $[a_{N_n+N_a}, \dots, a_{N-1}]^T$ and first row $[a_{N_n+N_a}, \dots, a_0]$ and $\nu_{[N_c:N-N_a-1]} = \tilde{Q}\eta_{[N_c-L_q:N-N_a-1+L_q]}$ is the colored noise at the receiver matched filter output, where $\tilde{Q} = T \left\{ [q_0, 0, \dots, 0]^T, [q^T, 0, \dots, 0] \right\} \in \mathbb{R}^{(N-N_a-N_c) \times (N-N_a-N_c+N_q)}$ and

$$q^T = [q[+L_q], \dots, q[0], \dots, q[-L_q]] \in \mathbb{R}^{2L_q+1}$$

where q denotes the symbol rate sampled receiver pulse matched filter.

Similarly, we can write the pulse matched filter output which includes *all the contributions from the known training symbols* (including output which includes contributions from adjacent unknown random data) as

$$\begin{aligned} y_{[-N_a:N+N_c-1]} &= (A + D)h + \nu_{[-N_a:N+N_c-1]} \quad (11) \\ &= Ah + Dh + Q\eta_{[-N_a-L_q:N+N_c-1+L_q]} \end{aligned}$$

where $A = T \left\{ [a_0, \dots, a_{N-1}, \underbrace{0, \dots, 0}_{N_a+N_c}]^T, [a_0, 0, \dots, 0] \right\}$ is the $(N + N_a + N_c) \times (N_a + N_c + 1)$ Toeplitz matrix with first column $[a_0, \dots, a_{N-1}, 0, \dots, 0]^T$ and first row $[a_0, 0, \dots, 0]$ and $D = T \left\{ \underbrace{[0, \dots, 0, d_N, \dots, d_{N_a+N_c+N-1}]^T}_N, \underbrace{[0, d_{-1}, \dots, d_{-N_a-N_c}]}_{\text{previous frame data}} \right\}$

is a Toeplitz matrix which includes adjacent random information symbols only, both prior to the training sequence and after the training sequence. We shall only use the statistics of this random data, (since the actual values are unknown) and use these statistics to compute the covariance matrix of the correlated noise, to be used in the solution for the BLUE estimate of h . The data symbols $d_{-1}, \dots, d_{-N_a-N_c}$ denote the unknown information symbols transmitted at the end of the previous frame. $\nu_{[-N_a:N+N_c-1]} = Q\eta_{[-N_a-L_q:N+N_c-1+L_q]}$ is the colored noise at the receiver matched filter output, where $Q \in \mathbb{R}^{(N+N_a+N_c) \times (N+N_a+N_c+N_q)}$ is defined similarly to \tilde{Q} . To compute the covariance matrix for the noise contribution, $Dh + Q\eta_{[-N_a-L_q:N+N_c-1+L_q]}$, to the received vector $y_{[-N_a:N+N_c-1]}$, it is advantageous to rewrite the term Dh .

We define $d \in \mathbb{R}^{N+2(N_c+N_a)}$ by

$$d = [d_{-N_c-N_a}, \dots, d_{-1}, \mathbf{0}_{1 \times N}, d_N, \dots, d_{N+N_c+N_a-1}]^T$$

and $H = H(h) = T \left\{ [h[-N_a], 0, \dots, 0], [\bar{h}^T, 0, \dots, 0]^T \right\} \in \mathbb{R}^{(N+N_a+N_c) \times (N+2(N_c+N_a))}$ is the Toeplitz channel convolution matrix where

$$\bar{h}^T = [h[N_c], \dots, h[1], h[0], h[-1], \dots, h[-N_a]] \in \mathbb{R}^{N_c+N_a+1}$$

is the time-reversal of the channel vector, h^T . Then, we note that $Dh = Hd$ and $E \left[Hd(Hd)^T \right] = \sigma_d^2 HSH^T$ where $E[dd^T] = \sigma_d^2 S$ and $S \in R^{(N+2(N_c+N_a)) \times (N+2(N_c+N_a))}$ is given by

$$S = \begin{bmatrix} I_{(N_c+N_a)} & 0_{(N_c+N_a) \times N} & 0_{(N_c+N_a)} \\ 0_{N \times (N_c+N_a)} & 0_N & 0_{N \times (N_c+N_a)} \\ 0_{(N_c+N_a)} & 0_{(N_c+N_a) \times N} & I_{(N_c+N_a)} \end{bmatrix} \quad (12)$$

Since $Dh = Hd$, then we may rewrite (11) as

$$y_{[-N_a:N+N_c-1]} = Ah + Hd + Q\eta_{[-N_a-L_q:N+N_c-1+L_q]} \quad (13)$$

Hence,

$$C = C(h) \quad (14)$$

$$= \text{Var}(Hd + Q\eta_{[-N_a-L_q:N+N_c-1+L_q]}) \quad (15)$$

$$= \sigma_d^2 HSH^T + \sigma_\eta^2 QQ^T \quad (16)$$

III. PROBLEM DESCRIPTION AND PREVIOUS WORK

The solution of the general linear model, (13), is given by the Gauss-Markoff Theorem as

$$h = (A^T C(h)^{-1} A)^{-1} A^T C(h)^{-1} y \quad (17)$$

where $C(h)$ is given by (14). We note that in solving (17) we are looking for a fixed point, denoted by h_0 , of the mapping $h \mapsto F(h)$ where, for each fixed vector of received symbols, y , and each fixed convolution matrix of known transmitted training symbols, A , then $F : R^L \rightarrow R^L$ is given by (2).

Our own previous approaches to finding the semi-blind BLUE channel estimate have encompassed:

Case (1): In [17] and [18] an initial thresholded approximation, denoted $h^{(0)}$, to h was obtained via correlation of the training sequence with a stored copy of the training sequence at the receiver. Then the iteration

$$h^{(k+1)} = \left(A^T C(h^{(k)})^{-1} A \right)^{-1} A^T C(h^{(k)})^{-1} y, \quad k = 0, 1, 2, \dots$$

was used to generate a sequence of approximations to the fixed point, h_{BLUE} . Numerical simulations indicated that two or three iterations were sufficient for an error of $\|h^{(k)} - h_{BLUE}\| \sim 10^{-6}$ (here $h \in R^{512}$.) Theoretically convergence of this iteration to the unique fixed point, h_{BLUE} , of the function $F(h) = F_{y,A}(h)$ is guaranteed if $\|J_F(h)\| < 1$ where $J_F(h)$ denotes the Jacobian matrix of F , [18].

Case (2): In [15] and [16], an approximate linear system was derived to give an approximate solution to (17). This approximate linear system was derived by replacing $C(h)$ on the right-hand side of (17) by $C(h_{id})$ where $h_{id} = [0, \dots, 0, 1, 0, \dots, 0]^T \in R^L$ where the 1 appears in the 64^{th} position in our case, to correspond to the position of the cursor in the decision feedback equalizer (DFE) which we use. This approach has the added advantage that the

matrix $(A^T C(h_{id})^{-1} A)^{-1} A^T C(h_{id})^{-1}$ may be computed offline and stored at the receiver.

In the present work, we take the point of view that we may expand the vector-valued function $F(h) = F_{y,A}(h) = (A^T C(h)^{-1} A)^{-1} A^T C(h)^{-1} y$ of the vector variable, h , using Taylor's Formula, [3], [11] about a fixed "ideal" vector, h_{id} . The approximation to $F_{y,A}(h)$ given by $F_{y,A}(h_{id})$ in *Case (2)*, above, is the zeroeth order Taylor approximation. We then derive explicitly first, second, and third order Taylor approximations to the BLUE CIR estimate. The methods are attractive, since computation of the matrices involved may be done offline and stored at the receiver. The only processing which will be needed at the receiver is that of computing $(h_i - h_i^{id})$, where h_i is the i^{th} entry of the approximate channel vector obtained by correlation, and $h_i^{id} = 1$, if we use $h_{id} = [0, \dots, 0, \underbrace{1}_{i^{th}}, 0, \dots, 0]^T$, and then to form the matrices used in obtaining the vectors given in (4) and (5).

IV. DERIVATION OF THE APPROXIMATIONS $\nabla^1 F(h_{id})$ AND $\nabla^2 F(h_{id})$

We define the matrix

$$A_C(h) = (A^T C(h)^{-1} A)^{-1} A^T C(h)^{-1} \in R^{L \times (N+L-1)}. \quad (18)$$

Then $F : R^L \rightarrow R^L$ defined by $F : h \mapsto F(h)$ in (2) is given by

$$F(h) = A_C(h) y \in R^L.$$

We state the following Propositions, and give brief proofs.

Proposition 1: For any matrix $B = B(h) \in R^{n \times n}$ which depends on a vector parameter $h = [h_1, \dots, h_K] \in R^K$ such that there exists an open set $U \subset R^K$, such that B is nonsingular and differentiable on U , then, we have that for any i , $1 \leq i \leq K$, and for any $h \in U$,

$$\frac{\partial B(h)^{-1}}{\partial h_i} = -B(h)^{-1} \frac{\partial B(h)}{\partial h_i} B(h)^{-1} \quad (19)$$

Proof: Differentiate with respect to h_i both sides of the identity $B(h)^{-1} B(h) = I$, using the product rule. ■

Proposition 2: For $1 \leq i \leq L$, we have

$$\frac{\partial F(h)}{\partial h_i} = A_C(h) \frac{\partial C(h)}{\partial h_i} C(h)^{-1} \{AA_C(h) - I\} y \quad (20)$$

Proof: Differentiate, using the product rule, with respect to h_i , the expression for $A_C(h) y$ and use *Proposition 1*

three times. That is, we have,

$$\begin{aligned}
\frac{\partial F(h)}{\partial h_i} &= \frac{\partial}{\partial h_i} \left((A^T C(h)^{-1} A)^{-1} \right) A^T C(h)^{-1} y \\
&\quad + (A^T C(h)^{-1} A)^{-1} A^T \frac{\partial}{\partial h_i} (C(h)^{-1}) y \\
&= (A^T C(h)^{-1} A)^{-1} A^T C(h)^{-1} \times \\
&\quad \times \left(\frac{\partial}{\partial h_i} C(h) \right) C(h)^{-1} A (A^T C(h)^{-1} A)^{-1} \times \\
&\quad \times A^T C(h)^{-1} y \\
&\quad - (A^T C(h)^{-1} A)^{-1} A^T C(h)^{-1} \times \\
&\quad \times \left(\frac{\partial}{\partial h_i} C(h) \right) C(h)^{-1} y
\end{aligned}$$

which gives the result. ■

Proposition 3: For $1 \leq i, j \leq L$, we have

$$\begin{aligned}
\frac{\partial^2 F(h)}{\partial h_i \partial h_j} &= \left\{ \frac{\partial A_C(h)}{\partial h_j} \frac{\partial C(h)}{\partial h_i} + A_C(h) \left[\frac{\partial^2 C(h)}{\partial h_i \partial h_j} \right. \right. \\
&\quad \left. \left. \frac{\partial C(h)}{\partial h_i} C(h)^{-1} \frac{\partial C(h)}{\partial h_j} \right] \right\} \times \\
&\quad \times C(h)^{-1} \{ AA_C(h) - I \} y + \\
&\quad A_C(h) \frac{\partial C(h)}{\partial h_i} C(h)^{-1} \left\{ A \frac{\partial A_C(h)}{\partial h_j} \right\} y
\end{aligned} \tag{21}$$

(22)

(23)

Proof: Differentiate (20) with respect to h_j , using the product rule and employing *Proposition 1*. We omit the details. ■

We omit the expression for the third derivatives, but it is straightforward to compute from (21) above.

Proposition 4: For $1 \leq i \leq L$, we have

$$\frac{\partial C(h)}{\partial h_i} = \sigma_d^2 H(h) S \left(\frac{\partial H(h)^T}{\partial h_i} \right) + \sigma_d^2 \left(\frac{\partial H(h)}{\partial h_i} \right) S H(h)^T \tag{24}$$

Proof: Differentiating (14) with respect to h_i gives the required result. ■

Proposition 5: For $1 \leq i, j \leq L$, we have

$$\frac{\partial^2 C(h)}{\partial h_i \partial h_j} = \sigma_d^2 \left(\frac{\partial H(h)}{\partial h_j} \right) S \left(\frac{\partial H(h)^T}{\partial h_i} \right) + \tag{25}$$

$$\sigma_d^2 \left(\frac{\partial H(h)}{\partial h_i} \right) S \left(\frac{\partial H(h)^T}{\partial h_j} \right) \tag{26}$$

Proof: Differentiating (24) with respect to h_j , and noting that $\partial^2 H(h)^T / \partial h_i \partial h_j = 0$ for any i, j , gives the required result. ■

Proposition 6: For all derivatives of order 3 and higher, we have $\left(\frac{\partial}{\partial h} \right)^\alpha C(h) \equiv 0_{(N+L-1) \times (N+L-1)} \quad \forall \alpha, |\alpha| \geq 3$

Proof: Differentiating (25) with respect to h_k , and noting that $\partial^2 H(h)^T / \partial h_i \partial h_j = 0$ for any i, j , gives the required result. ■

We use approximations (4) and (5) where we do *not use a full linear approximation*, including all terms in the linear

Taylor series approximation. We only use a linear approximation using the dominant term, which for us is a linear approximation where we have differentiated with respect to the *main tap weight*. This approach is supported by the simulation results where the best results are obtained in this case. This has the added advantage of reducing the complexity since only a single term is used in the approximation. A similar comment holds for both the quadratic and cubic approximations, where *we do not use the full second and third order Taylor series approximations*, but only use a single second order term and a single third order term in the second and third order approximations.

V. ALGORITHM

A. *Input:*

Received vector y ;

Training sequence to form data matrix A

B. *Stored at Receiver:*

Zeroeth order approximation matrix (see (18))

$$A_C(h_{id}) = (A^T C(h_{id})^{-1} A)^{-1} A^T C(h_{id})^{-1} \in R^{L \times (N+L-1)} \tag{27}$$

First order approximation matrix $\in R^{L \times (N+L-1)}$ for first order algorithm (see (20) and note that $\frac{\partial F(h)}{\partial h_i} = \frac{\partial A_C(h_{id})}{\partial h_i} y$)

$$\frac{\partial A_C(h_{id})}{\partial h_i} = A_C(h_{id}) \frac{\partial C(h_{id})}{\partial h_i} C(h_{id})^{-1} \times \tag{28}$$

$$\times \{ AA_C(h_{id}) - I \} \tag{29}$$

Second Order approximation matrix $\in R^{L \times (N+L-1)}$, if second order algorithm is used (see (21) and note that $\frac{\partial^2 F(h)}{\partial h_i \partial h_j} = \frac{\partial^2 A_C(h_{id})}{\partial h_i \partial h_j} y$)

$$\frac{\partial^2 A_C(h_{id})}{\partial h_i^2} = \left\{ \frac{\partial A_C(h_{id})}{\partial h_i} \frac{\partial C(h_{id})}{\partial h_i} \right. \tag{30}$$

$$\left. + A_C(h_{id}) \left[\frac{\partial^2 C(h_{id})}{\partial h_i^2} - \right. \right. \tag{31}$$

$$\left. \left. \frac{\partial C(h_{id})}{\partial h_i} C(h_{id})^{-1} \frac{\partial C(h_{id})}{\partial h_i} \right] \right\} \times \tag{32}$$

$$\times C(h_{id})^{-1} \{ AA_C(h_{id}) - I \} \tag{33}$$

$$+ A_C(h_{id}) \frac{\partial C(h_{id})}{\partial h_i} C(h_{id})^{-1} \times$$

$$\times \left\{ A \frac{\partial A_C(h_{id})}{\partial h_i} \right\} \tag{34}$$

$$\in R^{L \times (N+L-1)} \tag{35}$$

Third order approximation matrix if third order algorithm is used (not given here.)

C. *Real-Time Processing*

Let $h_{corr} = [h_1^{corr}, h_2^{corr}, \dots, h_L^{corr}]^T \in R^L$ denote the correlation approximation to the channel which is available from timing acquisition. Then compute the scalar value

$h_i^{corr} - h_i^{id} = h_i^{corr} - 1$ when the ideal channel used is $h_{id} = [0, \dots, 0, \underbrace{1}_{i^{th} \text{ position}}, 0, \dots, 0]^T \in R^L$. Use this, together with (27) and (28), to form the matrix

$$(\nabla^1 A_C)(h_{id}) = A_C(h_{id}) + (h_i^{corr} - h_i^{id}) \frac{\partial A_C(h_{id})}{\partial h_i} \quad (36)$$

to be used in the first order algorithm. (Complexity is $L \times (N + L - 1)$ multiplications and $L \times (N + L - 1)$ additions.) Similarly, using (27), (28) and (30), form the matrix

$$(\nabla^2 A_C)(h_{id}) = A_C(h_{id}) + (h_i^{corr} - h_i^{id}) \frac{\partial A_C(h_{id})}{\partial h_i} + (h_i^{corr} - h_i^{id})^2 \frac{\partial^2 A_C(h_{id})}{\partial h_i^2} \quad (37)$$

(Complexity is $2 \times L \times (N + L - 1)$ multiplications and $2 \times L \times (N + L - 1)$ additions.) We omit the third order approximation matrix in the interests of space.

C.1 Compute the approximations to h_{BLUE}

$\nabla^0 F(h_{id}) = (\nabla^0 A_C(h_{id}))y$ zeroeth order approximation [8] identical to that in [15] and [16]

$\nabla^1 F(h_{id}) = (\nabla^1 A_C(h_{id}))y$ first order approximation using (36)

$\nabla^2 F(h_{id}) = (\nabla^2 A_C(h_{id}))y$ second order approximation using (37) [10]

$\nabla^3 F(h_{id}) = (\nabla^3 A_C(h_{id}))y$ third order approximation, not explicitly given here [11]

(Each takes $(N + N_a + N_c) \times (N_a + N_c + 1)$ multiplications.)

VI. SIMULATION RESULTS

Simulation results are shown in Figure 1 (where *Normalized Least Squares Error* = $\|h - \hat{h}\|^2 / (N_a + N_c + 1) \|h\|^2$ where \hat{h} is the channel estimate) for Brazilian channel D which appears in the HDTV literature, [1], at 18,20,22,24,26,28 dB input SNR for the zeroeth, first, second and third order approximations.

REFERENCES

[1] Report of the SET/ABERT DTV Committee, January 2000
[2] S. Benedetto, E. Biglieri, *Principles of Digital Transmission with Wireless Applications*, Kluwer Academic/Plenum Publishers, New York, 1999.
[3] A. Browder, *Mathematical Analysis: An Introduction*, Springer Verlag, New York, 1996.
[4] V. Buchoux, O. Cappé, E. Moulines, and A. Gorokhov, "On the performance of semi-blind subspace based channel estimation," *IEEE Trans. Signal Process.*, vol. 48, no. 6, pp. 1750-59, June, 2004.
[5] H.-Y. Chung, and Y.-Y. Sun, "Analysis and parameter estimation of nonlinear systems with Hammerstein model using Taylor series approach," *IEEE Trans. on Circuits and Systems*, vol. 35, no. 2, pp. 1539-41, Dec. 1988.
[6] E. de Carvalho, D. T. M. Slock, "Blind and semi-blind FIR multichannel estimation: (global) identifiability conditions," *IEEE Trans. Signal Process.*, vol. 52, No. 4, pp 1053-64, April 2004.

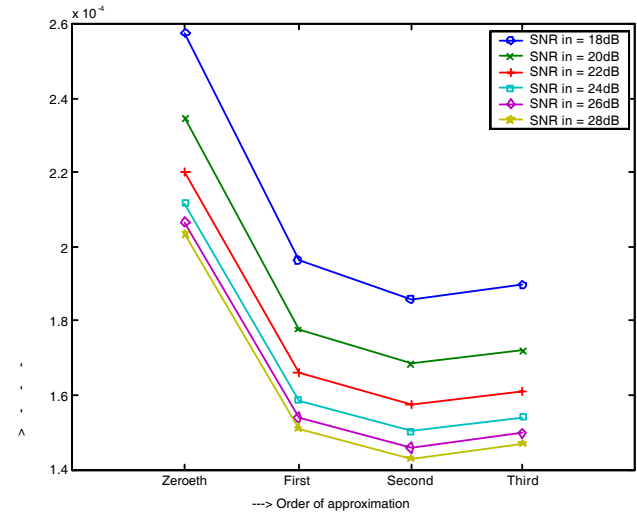


Fig. 1.

[7] A. Gelb, J. F. Kaspar, Jr., R. A. Nash, Jr., C. F. Price, A. A. Sutherland, Jr., *Applied Optimal Estimation*, MIT Press, Cambridge, MA, 1974.
G. B. Giannakis, Y. Hua, P. Stoica, L. Tong, *Signal Processing Advances in Wireless and Mobile Communications: Trends in Channel Estimation and Equalization*, Prentice Hall, Upper Saddle River, NJ, 2001.
[9] G. T. Heydt, M. Bakroun, and A. Inan, "Voltage flicker estimation based on linearization and L_p norms," *IEEE Trans. Power Delivery*, vol. 18, no. 4, pp. 1564-1566, Oct. 2003.
[10] S. L. Kay, *Fundamentals of Statistical Signal Processing: Volume 1, Estimation Theory*, Prentice-Hall Signal Processing Series, Upper Saddle River, NJ, 1993.
[11] S. Lang, *Undergraduate Analysis, 2nd Edition*, Springer Verlag, New York, 1997.
[12] S. Lasaulce, P. Loubaton, and E. Moulines, "A semi-blind channel estimation technique based on second-order blind method for CDMA systems," *IEEE Trans. Signal Process.* vol. 51, no. 7, pp. 1894-1904, July, 2003.
[13] A. Medles, D. T. M. Slock, and E. De Carvalho, "Linear prediction based semi-blind estimation of MIMO FIR channels," in *Proc. Third IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications*, Taoyuan, Taiwan, March 20-23, 2001.
[14] J. M. Mendel, *Lessons in Estimation Theory for Signal Processing, Communications and Control*, Prentice-Hall Signal Processing Series, Upper Saddle River, NJ, 1995.
[15] S. Özen, M. Fimoff, C. Pladdy, S. M. Nerayanuru, M. D. Zoltowski, "Approximate best linear unbiased channel estimation for frequency selective multipath channels with long delay spreads," *Asilomar Conference on Signals and Systems*, Monterey, CA, Nov. 2003.
[16] S. Özen, M. Fimoff, C. Pladdy, S. M. Nerayanuru, M. D. Zoltowski, "Approximate best linear unbiased channel estimation for frequency selective multipath channels with long delay spreads," unpublished.
[17] C. Pladdy, S. Özen, S. M. Nerayanuru, M. D. Zoltowski, M. Fimoff, "Best linear unbiased channel estimation for frequency selective multipath channels with long delay spreads," *Vehicular Technology Conference*, Orlando, FL, Fall 2003.
[18] C. Pladdy, S. Özen, S. M. Nerayanuru, M. D. Zoltowski, M. Fimoff, "A semi-blind iterative algorithm for best linear unbiased channel estimation", unpublished.
[19] M. K. Tsatsanis, and G. B. Giannakis, "Subspace methods for blind estimation of time-varying FIR channels," *IEEE Trans. Signal Processing*, vol. 45, no. 12, pp. 3084-93, Dec. 1997.