SEMI-BLIND TIME-VARYING CHANNEL ESTIMATION USING SUPERIMPOSED TRAINING

Xiaohong Meng *Jitendra K. Tugnait*

Department of Electrical & Computer Engineering
Auburn University, Auburn, Alabama 36849, USA

ABSTRACT

Channel estimation for single-input multiple-output (SIMO) time-varying channels is considered using superimposed training. The time-varying channel is assumed to be described by a complex exponential basis expansion model (CE-BEM). A periodic (non-random) training sequence is arithmetically added (superimposed) at a low power to the information sequence at the transmitter before modulation and transmission. A two-step approach is adopted where in the first step we estimate the channel using only the first-order statistics of the data. Using the estimated channel from the first step, a Viterbi detector is used to estimate the information sequence. In the second step a deterministic maximum likelihood (DML) approach is used to iteratively estimate the SIMO channel and the information sequences sequentially. An illustrative computer simulation example is presented where a frequency-selective channel is randomly generated with different Doppler spreads via Jakes' model.

1. INTRODUCTION

Consider a time-varying SIMO (single-input multiple-output) FIR (finite impulse response) linear channel with N outputs. Let $\{s(n)\}$ denote a scalar sequence which is input to the SIMO time-varying channel with discrete-time impulse response $\{\mathbf{h}(n;l)\}$ (N-vector channel response at time n to a unit input at time $n-l$). The vector channel may be the result of multiple receive antennas and/or oversampling at the receiver. Then the symbol-rate, channel output vector is given by

$$\mathbf{x}(n) := \sum_{l=0}^{L} \mathbf{h}(n;l)s(n-l). \tag{1}$$

In a complex exponential basis expansion representation [4] it is assumed that

$$\mathbf{h}(n;l) = \sum_{q=1}^{Q} \mathbf{h}_q(l)e^{j\omega_q n} \tag{2}$$

where N-column vectors $\mathbf{h}_q(l)$ (for $q = 1, 2, \cdots, Q$) are time-invariant. Eqn. (2) is a basis expansion of $\mathbf{h}(n;l)$ in the time variable n onto complex exponentials with frequencies $\{\omega_q\}$. The noisy measurements of $\mathbf{x}(n)$ are given by

$$\mathbf{y}(n) = \mathbf{x}(n) + \mathbf{v}(n) \tag{3}$$

A main objective in communications is to recover $s(n)$ given noisy $\{\mathbf{x}(n)\}$. In several approaches this requires knowledge of the channel impulse response [11], [9]. In training-based approach, $s(n) = c(n) =$ training sequence (known to the receiver) for (say) $n = 0, 1, \cdots, M_t - 1$ and $s(n)$ for $n > M_t - 1$ is the information sequence (unknown apriori to the receiver) [11], [9]. Therefore, given $c(n)$ and corresponding noisy $\mathbf{x}(n)$, one estimates the channel via least-squares and related approaches. For time-varying channels, one has to send a training signal frequently and periodically to keep up with the changing channel. This wastes resources. An alternative is to estimate the channel based solely on noisy $\mathbf{x}(n)$ exploiting statistical and other properties of $\{s(n)\}$ [11], [9]. This is the blind channel estimation approach. In semi-blind approaches, there is a training sequence but one uses the non-training based data also to improve the training-based results: it uses a combination of training and blind cost functions. This allows one to shorten the training period. Optimal placement and performance lower bounds for semi-blind approaches are in [1] and [2]. More recently a superimposed training based approach has been explored where one takes

$$s(n) = b(n) + c(n), \tag{4}$$

where $\{b(n)\}$ is the information sequence and $\{c(n)\}$ is a training (pilot) sequence added (superimposed) at a low power to the information sequence at the transmitter before modulation and transmission. There is no loss in information rate. On the other hand, some useful power is wasted in superimposed training which could have otherwise been allocated to the information sequence. Superimposed training-based approaches have been discussed in [5], [6] and [8] for SISO systems. Periodic superimposed training for channel estimation via first-order statistics for SISO systems have been discussed in [15] and [13] for time-invariant channels, and in [12] for both time-invariant and time-varying (CE-BEM based) channels. In [3] performance bounds for training and superimposed training-based semi-blind SISO channel estimation for time-varying flat fading channels have been discussed.

Objectives and Contributions: In this paper we extend the first-order statistics-based approach of [12] for time-varying (CE-BEM based) channels to semiblind versions using Viterbi detectors. The first-order statistics-based approach views the information sequence as interference whereas in semiblind versions it is exploited to enhance channel estimation and information sequence detection.

Notation: Superscripts H, T and $\dagger$ denote the complex conjugate transpose, the transpose and the Moore-Penrose pseudo-inverse operations, respectively. $\delta(\tau)$ is the Kronecker delta and I_N is the $N \times N$ identity matrix. The symbol $\otimes$ denotes the Kronecker product.

1.1. On CE-BEM Representation

We now briefly discuss the CE-BEM representation of time-varying communications channels, following [4] and particularly [7], to consider practical situations where the basis frequencies ω_q's would be known apriori. Consider a time-varying (e.g. mobile wireless) channel with complex baseband, continuous-time, received signal $x(t)$ and transmitted complex baseband, continuous-time information signal $s(t)$ (with symbol interval T_s sec.) related by $h(t;\tau)$ which is the time-varying impulse response of the channel (response at time t to a unit impulse at time $t-\tau$). Let τ_d denote the (multipath) delay-spread of the channel and let f_d denote the Doppler spread of the channel. If $x(t)$ is sampled once every T_s sec. (symbol rate), then

This work was supported by the US Army Research Office under Grant DAAD19-01-1-0539.

by [7], for $t = nT_s \in [t_0, t_0 + TT_s)$, the sampled signal $x(n) := x(t)|_{t=nT_s}$ has the representation

$$x(n) = \sum_{l=0}^{L} h(n;l)s(n-l) \tag{5}$$

where

$$h(n;l) = \sum_{q=1}^{Q} h_q(l)e^{j\omega_q n}, \tag{6}$$

$$w_q = \frac{2\pi}{T}\left(q - \frac{1}{2} - \frac{Q}{2}\right), \quad L := \lfloor \tau_d/T_s \rfloor, \quad Q := 2\lceil f_d T T_s \rceil + 1. \tag{7}$$

This is a scenario where the CE-BEM representation is appropriate. The above representation is valid over a duration of TT_s seconds (T samples). Eqn. (1) arises if we follow (5) and consider a SIMO model arising due to multiple antennas at the receiver.

2. FIRST-ORDER STATISTICS-BASED SOLUTION OF [12]

Assume the following:

- (**H1**) The time-varying channel $\{\mathbf{h}(n;l)\}$ satisfies (2) where the frequencies ω_q ($q = 1, 2, \cdots, Q$) are distinct and known (cf. Sec. 1.1) with $\omega_q \in [0, 2\pi)$. Also $N \geq 1$.
- (**H2**) The information sequence $\{b(n)\}$ is zero-mean, white with $E\{|b(n)|^2\} = 1$.
- (**H3**) The measurement noise $\{\mathbf{v}(n)\}$ is **nonzero-mean** ($E\{\mathbf{v}(n)\} = \mathbf{m}$), white, uncorrelated with $\{b(n)\}$, with $E\{[\mathbf{v}(n+\tau) - \mathbf{m}][\mathbf{v}(n) - \mathbf{m}]^H\} = \sigma_v^2 I_N \delta(\tau)$. The mean vector $\mathbf{m}$ may be unknown.
- (**H4**) The superimposed training sequence $c(n) = c(n+mP)$ $\forall m, n$ is a non-random periodic sequence with period P.

By (**H4**), we have $c_m := \frac{1}{P}\sum_{n=0}^{P-1} c(n)e^{-j\alpha_m n}$,

$$c(n) = \sum_{m=0}^{P-1} c_m e^{j\alpha_m n} \;\; \forall n, \;\; \alpha_m := 2\pi m/P. \tag{8}$$

The coefficients c_m are known at the receiver since $\{c(n)\}$ is known. We have

$$E\{\mathbf{y}(n)\} = \sum_{q=1}^{Q}\sum_{m=0}^{P-1} \underbrace{\left[\sum_{l=0}^{L} c_m \mathbf{h}_q(l)e^{-j\alpha_m l}\right]}_{=:\mathbf{d}_{mq}} e^{j(\omega_q+\alpha_m)n} + \mathbf{m}. \tag{9}$$

Suppose that we pick P to be such that $(\omega_q + \alpha_m)$'s are all distinct for any choice of m and q. Then the sequence $E\{\mathbf{y}(n)\}$ is (almost) periodic with cycle frequencies $(\omega_q + \alpha_m)$, $1 \leq q \leq Q$, $0 \leq m \leq P-1$. A mean-square (m.s.) consistent estimate $\hat{\mathbf{d}}_{mq}$ of $\mathbf{d}_{mq}$, for $\omega_q + \alpha_m \neq 0$, follows as

$$\hat{\mathbf{d}}_{mq} = \frac{1}{T}\sum_{n=1}^{T} \mathbf{y}(n)e^{-j(\omega_q+\alpha_m)n}. \tag{10}$$

As $T \to \infty$, $\hat{\mathbf{d}}_{mq} \to \mathbf{d}_{mq}$ m.s. if $\omega_q + \alpha_m \neq 0$ and $\hat{\mathbf{d}}_{0q} \to \mathbf{d}_{0q} + \mathbf{m}$ m.s. if $\omega_q + \alpha_m = 0$.

It is established in [12] that given $\mathbf{d}_{mq}$ for $1 \leq q \leq Q$ and $1 \leq m \leq P-1$, we can (uniquely) estimate $\mathbf{h}_q(l)$'s if $P \geq L+2$, $\alpha_m \neq 0$, and $c_m \neq 0$ $\forall m \neq 0$. Since $\mathbf{m}$ is unknown, we will omit the term $m = 0$ for further discussion. Define

$$\mathbf{V} := \begin{bmatrix} 1 & e^{-j\alpha_1} & \cdots & e^{-j\alpha_1 L} \\ 1 & e^{-j\alpha_2} & \cdots & e^{-j\alpha_2 L} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & e^{-j\alpha_{P-1}} & \cdots & e^{-j\alpha_{P-1} L} \end{bmatrix}_{(P-1)\times(L+1)}, \tag{11}$$

$$\mathbf{D}_m := [\mathbf{d}_{m1}^T, \; \mathbf{d}_{m2}^T, \; \cdots, \; \mathbf{d}_{mQ}^T]^T, \; [NQ] \times 1, \tag{12}$$

$$\mathbf{H}_l := [\mathbf{h}_1^T(l), \; \mathbf{h}_2^T(l), \; \cdots, \; \mathbf{h}_Q^T(l)]^T, \; [NQ] \times 1, \tag{13}$$

$$\mathcal{H} := \begin{bmatrix} \mathbf{H}_0^H & \mathbf{H}_1^H & \cdots & \mathbf{H}_L^H \end{bmatrix}^H, \; [(L+1)NQ] \times 1, \tag{14}$$

$$\mathcal{D} := \begin{bmatrix} \mathbf{D}_1^H & \mathbf{D}_2^H & \cdots & \mathbf{D}_{P-1}^H \end{bmatrix}^H, \; [(P-1)NQ] \times 1, \tag{15}$$

$$\mathcal{C} := \underbrace{(\mathrm{diag}\{c_1, c_2, \cdots, c_{P-1}\}\mathbf{V})}_{=:\mathcal{V}} \otimes I_{NQ}. \tag{16}$$

Omitting the term $m = 0$ and using the definition of $\mathbf{d}_{mq}$ from (9), it follows that

$$\mathcal{C}\mathcal{H} = \mathcal{D}. \tag{17}$$

It is shown in [12] that if $P - 1 \geq L + 1$ and $(\omega_q + \alpha_m)$'s are distinct $\forall q$ and $\forall m$, $\mathrm{rank}(\mathcal{C}) = NQ(L+1)$; hence, we can determine the $\mathbf{h}_q(l)$'s uniquely. Define

$$\hat{\mathbf{D}}_m := [\hat{\mathbf{d}}_{m1}^T, \; \hat{\mathbf{d}}_{m2}^T, \; \cdots, \; \hat{\mathbf{d}}_{mQ}^T]^T \tag{18}$$

and define $\hat{\mathcal{D}}$ as in (15) with $\mathbf{D}_m$'s replaced with $\hat{\mathbf{D}}_m$'s. Then we have the channel estimate

$$\hat{\mathcal{H}} = (\mathcal{C}^H\mathcal{C})^{-1}\mathcal{C}^H\hat{\mathcal{D}}. \tag{19}$$

.

3. DETERMINISTIC MAXIMUM LIKELIHOOD (DML) APPROACH

The first-order statistics-based approach of Sec. 2 views the information sequence as interference. Since the training and information sequences of a given user pass through an identical channel, this fact can be exploited to enhance the channel estimation performance via a semiblind approach. We consider joint channel and information sequence estimation via an iterative DML approach. We have guaranteed convergence to a local maximum. Furthermore, if we initialize with our superimposed training-based solution, one is guaranteed the global extremum (minimum error probability sequence estimator) if the superimposed training-based solution is "good."

Suppose that we have collected T samples of the observations. Form the vector

$$Y = [\mathbf{y}^T(T), \mathbf{y}^T(T-1), \cdots, \mathbf{y}^T(L+1)]^T. \tag{20}$$

Similarly, define

$$\mathbf{s} := [s(T), \, s(T-1), \cdots, s(1)]^T \tag{21}$$

We then have the following linear model ($\tilde{\mathbf{v}}(n) := \mathbf{v}(n) - \mathbf{m}$)

$$Y = \mathcal{T}(\mathbf{s})\mathcal{H} + \underbrace{\begin{bmatrix} \tilde{\mathbf{v}}(T) \\ \vdots \\ \tilde{\mathbf{v}}(L) \end{bmatrix}}_{=:\tilde{V}} + \underbrace{\begin{bmatrix} \mathbf{m} \\ \vdots \\ \mathbf{m} \end{bmatrix}}_{=:\mathcal{M}} \tag{22}$$

where $V = \tilde{V} + \mathcal{M}$ is a column-vector consisting of samples of noise $\{\mathbf{v}(n)\}$, $\mathcal{H}$ is defined in (14),

$$\mathcal{T}(\mathbf{s}) := \begin{bmatrix} s(T)\Sigma_T & \cdots & s(T-L)\Sigma_T \\ s(T-1)\Sigma_{T-1} & \cdots & s(T-L-1)\Sigma_{T-1} \\ \vdots & \vdots & \vdots \\ s(L+1)\Sigma_L & \cdots & s(1)\Sigma_L \end{bmatrix} \tag{23}$$

$$\Sigma_n := \begin{bmatrix} e^{j\omega_1 n} I_N & e^{j\omega_2 n} I_N & \cdots & e^{j\omega_Q n} I_N \end{bmatrix}. \tag{24}$$

Consider (1), (3) and (22). Under the assumption of white Gaussian measurement noise, consider the joint estimators

$$\{\widehat{\mathcal{H}}, \widehat{\mathbf{s}}, \widehat{\mathbf{m}}\} = \arg \left\{ \min_{\mathcal{H},\mathbf{s},\mathbf{m}} ||Y - \mathcal{T}(\mathbf{s})\mathcal{H} - \mathcal{M}||^2 \right\} \tag{25}$$

where $\widehat{\mathbf{s}}$ is the estimate of $\mathbf{s}$. In the above we have followed a DML approach assuming no statistical model for the input sequences $\{s(n)\}$. Under a white Gaussian noise assumption, the DML estimators are obtained by the nonlinear least-squares optimization (25). We have

$$Y = \mathcal{T}(\mathbf{s})\mathcal{H} + \tilde{V} + \mathcal{M} = \mathcal{C}(\mathcal{H})\mathbf{s} + \tilde{V} + \mathcal{M} \tag{26}$$

where

$$\mathcal{C}(\mathcal{H}) = \begin{bmatrix} \mathbf{h}(T;0) & \cdots & \mathbf{h}(T;L) & & \\ & \ddots & & \ddots & \\ & & \mathbf{h}(L+1;0) & \cdots & \mathbf{h}(L+1;L) \end{bmatrix} \tag{27}$$

is a "filtering matrix." We therefore have a separable nonlinear least-squares problem that can be solved sequentially as (joint optimization w.r.t. $\mathcal{H}, \mathbf{m}$ can be further "separated")

$$\begin{aligned} \{\widehat{\mathcal{H}}, \widehat{\mathbf{s}}, \widehat{\mathbf{m}}\} &= \arg \min_{\mathbf{s}} \{ \min_{\mathcal{H},\mathbf{m}} ||Y - \mathcal{T}(\mathbf{s})\mathcal{H} - \mathcal{M}||^2 \} \quad (28) \\ &= \arg \min_{\mathcal{H},\mathbf{m}} \{ \min_{\mathbf{s}} ||Y - \mathcal{C}(\mathcal{H})\mathbf{s} - \mathcal{M}||^2 \}. \quad (29) \end{aligned}$$

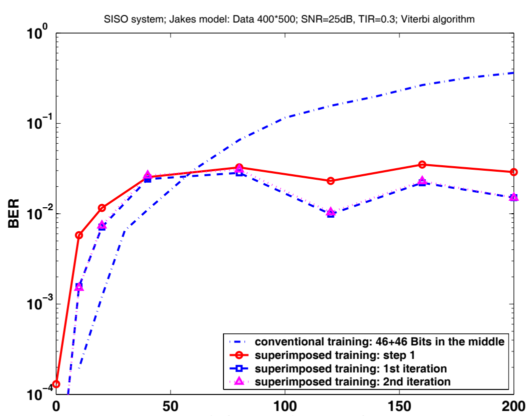

Figure 1. BER: circle: estimate channel using superimposed training and then design a Viterbi detector; square: first iteration specified by Step 2 (Sec. 3); triangle: second iteration specified by Step 2 (Sec. 3); dot-dashed: estimate channel using conventional time-multiplexed training of length 46 bits in the middle of a subblock of length 200 bits and then design a Viterbi detector. Training-to-information symbol power ratio =0.3 (-1.14 dB). Record length = 400 bits. Results based on 500 Monte Carlo runs.

The finite alphabet properties of the information sequences can also be incorporated into the DML methods. These algorithms, first proposed by Seshadri [10] for time-invariant SISO systems, iterate between estimates of the channel and the input sequences. At iteration k, with an initial guess of the channel $\mathcal{H}^{(k)}$ and the mean $\mathbf{m}^{(k)}$, the algorithm estimates the input sequence $\mathbf{s}^{(k)}$ and the channel $\mathcal{H}^{(k+1)}$ and mean $\mathbf{m}^{(k+1)}$ for the next iteration by

$$\begin{aligned} \mathbf{s}^{(k)} &= \arg \min_{\mathbf{s}\in\mathcal{S}} ||Y - \mathcal{C}(\mathcal{H}^{(k)})\mathbf{s} - \mathcal{M}^{(k)}||^2, \quad (30) \\ \mathcal{H}^{(k+1)} &= \arg \min_{\mathcal{H}} ||Y - \mathcal{T}(\mathbf{s}^{(k)})\mathcal{H} - \mathcal{M}^{(k)}||^2, \quad (31) \\ \mathbf{m}^{(k+1)} &= \arg \min_{\mathbf{m}} ||Y - \mathcal{T}(\mathbf{s}^{(k)})\mathcal{H}^{(k+1)} - \mathcal{M}||^2, (32) \end{aligned}$$

where $\mathcal{S}$ is the (discrete) domain of $\mathbf{s}$. The optimizations in (31) and (32) are linear least squares problems whereas the the optimization in (30) can be achieved by using the Viterbi algorithm [9]. Since the above iterative procedure involving (30), (31) and (32) decreases the cost at every iteration, one achieves a local minimum of the nonlinear least-squares cost (local maximum of DML function).

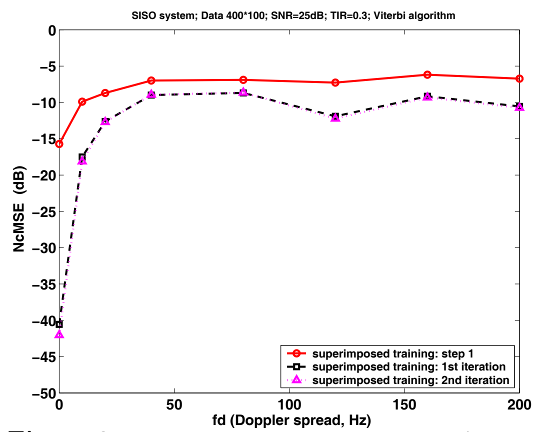

Figure 2. As in Fig. 2 except that NCMSE (normalized channel mean-square error) is shown.

We now summarize our DML approach:

1) a) Use (19) to estimate the channel using the first-order (cyclostationary) statistics of the observations. Denote the channel estimates by $\widehat{\mathcal{H}}^{(1)}$ and $\widehat{\mathbf{h}}_q^{(1)}(l)$. In this method $\{c(n)\}$ is known and $\{b(n)\}$ is regarded as interference.

b) Estimate the mean $\widehat{\mathbf{m}}^{(1)}$ as follows. Define (recall (1)-(3))

$$\hat{\mathbf{m}}^{(1)} := (1/T) \sum_{n=1}^{T} [\mathbf{y}(n) - \sum_{l=0}^{L} \hat{\mathbf{h}}^{(1)}(n;l) c(n-l)] \tag{33}$$

where

$$\hat{\mathbf{h}}^{(1)}(n;l) := \sum_{q=1}^{Q} \hat{\mathbf{h}}_q^{(1)}(l) e^{j\omega_q n} \tag{34}$$

c) Design a Viterbi sequence detector to estimate $\{s(n)\}$ as $\{\tilde{s}(n)\}$ using the estimated channel $\widehat{\mathcal{H}}^{(1)}$, mean $\widehat{\mathbf{m}}^{(1)}$ and cost (30) with $k = 1$. [Note that knowledge of $\{c(n)\}$ is used in $s(n) = b(n) + c(n)$, therefore, we are in essence estimating $b(n)$ in the Viterbi detector.]

2) a) Substitute $\tilde{s}(n)$ for $s(n)$ in (1) and use the corresponding formulation in (22) to estimate the channel $\mathcal{H}$ as

$$\widehat{\mathcal{H}}^{(2)} = \mathcal{T}^{\dagger}(\tilde{\mathbf{s}}) \left[Y - \widehat{\mathcal{M}}^{(1)} \right]. \tag{35}$$

The mean $\mathbf{m}$ is estimated as $\widehat{\mathbf{m}}^{(2)}$ using (33) with $\hat{\mathbf{h}}^{(1)}$ replaced with $\hat{\mathbf{h}}^{(2)}$ obtained from $\hat{\mathbf{h}}_q^{(2)}$ as in (34).

b) Design a Viterbi sequence detector using the estimated channel $\widehat{\mathcal{H}}^{(2)}$, mean $\widehat{\mathbf{m}}^{(2)}$ and cost (30) with $k = 2$, as in Step 1c.

3) Step 2 provides one iteration of (30)-(31). Repeat a few times if so desired.

4. SIMULATION EXAMPLE

Consider (1) with $N = 1$ and $L = 2$. We simulate a random time- and frequency-selective Rayleigh fading channel following [14]. For different l's, $h(n; l)$'s are mutually independent and for a given l, we follow the modified Jakes' model [14] to generate $h(n; l)$:

$$h(n; l) = X(t)|_{t=nT_s}, \tag{36}$$

$X(t) = \frac{2}{\sqrt{M}} \sum_{i=1}^{M} \mathrm{e}^{j\psi_i} \cos(2\pi f_d t \cos(\alpha_i) + \phi)$, $\alpha_i = (2\pi i - \pi + \theta)/(4M)$, $i = 1, 2, \cdots, M$, random variables θ, ϕ and ψ_i are mutually independent $(\forall i)$ and uniformly distributed over $[0, 2\pi)$, T_s = symbol interval, f_d= (max.) Doppler spread and $M = 25$. For a fixed l, (36) generates a random process $\{h(n; l)\}_n$ whose power spectrum approximates the Jakes' spectrum as $M \uparrow \infty$. We consider a system with carrier frequency of 2GHz, data rate of 40kB (kB= kilo-Bauds), therefore, $T_s = 25 \times 10^{-6}$ sec., and a varying Doppler spread f_d in the range 0Hz to 200Hz (corresponding to a maximum mobile velocity in the range 0 to 108km/hr). We picked a data record length of 400 symbols (time duration of 10 msec.). For a given Doppler spread, we pick Q as in Sec. 1.1 (T=400, L=2 in (7)). For the chosen parameters it varies within the values $\{1, 3, 5\}$. We emphasize that the CE-BEM was used only for processing at the receiver; the data were generated using (36), not the CE-BEM.

For comparison, we consider conventional training assuming time-invariant channels. The block of data of length 400 symbols was split into two non-overlapping blocks of 200 symbols each, Each subblock had a training sequence length of 46 symbols in the middle of the data subblock. Assuming synchronization, time-invariant channels were estimated using conventional training and used for information detection via a Viterbi algorithm; this was done for each subblock. We take all sequences (information and training) to be binary. For superimposed training, we take a periodic (scaled) binary sequence of period $P = 7$ with the training-to-information sequence power ratio (TIR) of 0.3 where

$$\mathrm{TIR} = \sigma_c^2 / \sigma_b^2, \tag{37}$$

and σ_b^2 and σ_c^2 denote the average power in the information sequence $\{b(n)\}$ and training sequence $\{c(n)\}$, respectively. Complex white zero-mean Gaussian noise was added to the received signal and scaled to achieve a target bit SNR at the receiver (relative to the contribution of $\{s(n)\}$).

Fig. 1 show the BER (bit error rate) based on 500 Monte Carlo runs for conventional training, the method of [12], and the proposed approximate DML approach with two iterations, under varying Doppler spreads f_d and a bit SNR of 25dB. It is seen that as Doppler spread f_d increases beyond about 60Hz (normalized Doppler $T_s f_d$ of 0.0015), superimposed training approach of [12] (Step 1) outperforms the conventional (midamble) training with time-invariant channel approximation, without decreasing the information rate. Furthermore, the proposed DML enhancement can lead to a significant improvement with just one iteration. Fig. 2 shows the normalized channel mean-square error (NCMSE), defined (before averaging over runs) as

$$\mathrm{NCMSE} = \left[\sum_{n=1}^{T} \sum_{l=0}^{2} |\widehat{h}(n; l) - h(n; l)|^2 \right] / \left[\sum_{n=1}^{T} \sum_{l=0}^{2} |h(n; l)|^2 \right]$$

5. CONCLUSIONS

The approach of [12] to SIMO CE-BEM-based time-varying channel estimation using superimposed training sequences (hidden pilots) and first-order statistics was extended to semiblind versions thereof. The results were illustrated via a simulation example involving time- and frequency-selective Rayleigh fading.

REFERENCES

[1] E. de Carvalho and D.T.M. Slock, "Semi-blind Methods for FIR Multichannel Estimation," Chapter 7 in *Signal Processing Advances in Wireless & Mobile Communications, vol. I*, G.B. Giannakis, Y. Hua, P. Stoica, and L. Tong (Eds.) Upper Saddle River, NJ: Prentice Hall, 2001.

[2] M. Dong and L. Tong, "Optimal design and placement of pilot symbols for channel estimation," *IEEE Trans. Signal Proc.*, vol. 50, pp. 3055-3069, Dec. 2002.

[3] M. Dong L. Tong and B.M. Sadler, "Optimal insertion of pilot symbols for transmissions over time-varying flat fading channels," in *Proc. 2003 IEEE Workshop on Signal Processing Advances in Wireless Commun.*, pp. 472-476, Rome, Italy, June 15-18, 2003.

[4] G.B. Giannakis and C. Tepedelenlioglu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," *Proc. IEEE* , vol. 86, pp. 1969-1986, Oct. 1998.

[5] P. Hoeher and F. Tufvesson, "Channel estimation with superimposed pilot sequence," in *Proc. IEEE GLOBECOM Conf.*, pp. 2162-66, Rio de Janeiro, Brazil, Dec. 5-9, 1999.

[6] T.P. Holden and K. Feher, "A spread spectrum based system technique for synchronization of digital mobile communication systems," *IEEE Trans. Broadcasting*, vol. 36, pp. 185-194, Sept. 1990.

[7] X. Ma and G.B. Giannakis, "Maximum-diversity transmissions over doubly-selective wireless channels" *IEEE Trans. Information Theory*, vol. IT-49, pp. 1832-1840, July 2003.

[8] F. Mazzenga, "Channel estimation and equalization for M-QAM transmission with a hidden pilot sequence," *IEEE Trans. Broadcasting*, vol. 46, pp. 170-176, June 2000.

[9] J.G. Proakis, *Digital Communications, 4th ed.*. NY: McGraw-Hill, 2001.

[10] N. Seshadri, "Joint data and channel estimation using fast blind trellis search techniques," *IEEE Trans. Commun.*, vol. COM-42, pp. 1000-1011, March 1994.

[11] J.K. Tugnait, L. Tong and Z. Ding, "Single-user channel estimation and equalization," *IEEE Signal Proc. Mag.*, vol. 17, pp. 16-28, May 2000.

[12] J.K. Tugnait and Weilin Luo, "On channel estimation using superimposed training and first-order statistics," in *Proc. 2003 ICASSP*, Hong Kong, April 2003.

[13] J.K. Tugnait and Weilin Luo, "On channel estimation using superimposed training and first-order statistics," *IEEE Commun. Letters*, vol. CL-8, pp. 413-415, Sept. 2003.

[14] Y.R. Zheng and C. Xiao, "Simulation models with correct statistical proerties for Rayleigh fading channels," *IEEE Trans. Commun.*, COM-51, pp. 920-928, June 2003.

[15] G.T. Zhou, M. Viberg and T. McKelvey, "A first-order statistical method for channel estimation" *IEEE Signal Proc. Letters*, vol. SPL-10, pp. 57-60, March 2003.