SYNCHRONIZATION OF SUPERIMPOSED TRAINING FOR CHANNEL ESTIMATION

Jitendra K. Tugnait *Xiaohong Meng*

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

ABSTRACT

Channel estimation for single-input multiple-output (SIMO) time-invariant or slowly time-varying channels was recently considered in [11] using superimposed training. 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. In [11], the channel is estimated using only the first-order statistics of the data under the assumption that the superimposed training sequence at the receiver is time-synchronized with its transmitted counterpart. In this paper we remove this assumption of synchronization and propose a novel approach to superimposed training synchronization. An illustrative computer simulation example is presented.

1. INTRODUCTION

Consider an 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 channel with discrete-time impulse response $\{\mathbf{h}(l)\}$. 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

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

The noisy measurements of $\mathbf{x}(n)$ are given by ($\{\mathbf{v}(n)\}$ is possible nonzero-mean [11], temporally and spatially white, Gaussian)

$$\mathbf{y}(n) = \mathbf{x}(n) + \mathbf{v}(n). \tag{2}$$

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 [10], [8]. 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) [10], [8]. 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 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 $\{s(n)\}$ [10], [8]. 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

$$s(n) = b(n) + c(n), \tag{3}$$

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

$\{b(n)\}$ 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 [4], [5] and [7] for SISO systems. The UTRA specification for 3G systems [6] allows for a spread pilot (superimposed) sequence in the base station's common pilot channel, suitable for downlinks. Periodic superimposed training for channel estimation via first-order statistics for SISO systems have been discussed in [12] and [11]. In [3] performance bounds for training and superimposed training-based semiblind SISO channel estimation for time-varying flat fading channels have been discussed.

Objectives and Contributions: In [11], the channel is estimated using only the first-order statistics of the data under the assumption that the superimposed training sequence at the receiver is time-synchronized with its transmitted counterpart. In this paper we remove this assumption of synchronization and propose a novel approach to superimposed training synchronization. Synchronization issues have not been considered in [12] (or in [4], [5], [7], and [3]).

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.

2. FIRST-ORDER STATISTICS-BASED SOLUTION OF [11]

Assume the following:

- **(H1)** The information sequence $\{b(n)\}$ is zero-mean, white with $E\{|b(n)|^2\} = 1$.
- **(H2)** 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 $\mathbf{m}$ is unknown.
- **(H3)** The superimposed training sequence $c(n) = c(n+mP)$ $\forall m, n$ is a non-random periodic sequence with period P.

By (**H3**), we have $c_m := \frac{1}{P}\sum_{n=0}^{P-1} c(n)e^{-j\alpha_m n}$,

$$c(n) = \sum_{m=0}^{P-1} c_m e^{j\alpha_m n} \ \ \forall n, \ \ \alpha_m := 2\pi m/P. \tag{4}$$

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

$$E\{\mathbf{y}(n)\} = \sum_{m=0}^{P-1} \underbrace{\left[\sum_{l=0}^{L} c_m \mathbf{h}(l) e^{-j\alpha_m l}\right]}_{=:\mathbf{d}_m} e^{j\alpha_m n} + \mathbf{m}. \tag{5}$$

The sequence $E\{\mathbf{y}(n)\}$ is periodic with cycle frequencies α_m, $0 \le m \le P-1$. A mean-square (m.s.) consistent estimate $\hat{\mathbf{d}}_m$ of $\mathbf{d}_m$, for $\alpha_m \ne 0$, follows as

$$\hat{\mathbf{d}}_m = \frac{1}{T}\sum_{n=1}^{T} \mathbf{y}(n) e^{-j\alpha_m n}. \tag{6}$$

As $T \to \infty$, $\hat{\mathbf{d}}_m \to \mathbf{d}_m$ m.s. if $\alpha_m \ne 0$ and $\hat{\mathbf{d}}_0 \to \mathbf{d}_0 + \mathbf{m}$ m.s. if $\alpha_m = 0$.

It is established in [11] that given $\mathbf{d}_m$ for $1 \le m \le P-1$, we can (uniquely) estimate $\mathbf{h}(l)$'s if $P \ge L+2$, $\alpha_m \ne 0$, and $c_m \ne 0\ \forall m \ne 0$. Since $\mathbf{m}$ is unknown, we will omit the term $m = 0$ for further discussion. Define

$$\mathbf{V} := \begin{bmatrix} 1 & e^{-j\alpha_1} & \cdots & e^{-j\alpha_1 L} \\ 1 & e^{-j\alpha_2} & \cdots & e^{-j\alpha_2 L} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & e^{-j\alpha_{P-1}} & \cdots & e^{-j\alpha_{P-1} L} \end{bmatrix}_{(P-1)\times(L+1)}, \tag{7}$$

$$\mathcal{H} := \begin{bmatrix} \mathbf{h}^H(0) & \mathbf{h}^H(1) & \cdots & \mathbf{h}^H(L) \end{bmatrix}^H, \ [N(L+1)] \times 1, \tag{8}$$

$$\mathcal{D} := \begin{bmatrix} \mathbf{d}_1^H & \mathbf{d}_2^H & \cdots & \mathbf{d}_{P-1}^H \end{bmatrix}^H, \ [N(P-1)] \times 1, \tag{9}$$

$$\mathcal{C} := \underbrace{(\mathrm{diag}\{c_1, c_2, \cdots, c_{P-1}\}\mathbf{V})}_{=:\mathcal{V}} \otimes I_N. \tag{10}$$

Omitting the term $m = 0$ and using the definition of $\mathbf{d}_m$ from (5), it follows that

$$\mathcal{C}\mathcal{H} = \mathcal{D}. \tag{11}$$

It is shown in [11] that if $P - 1 \ge L + 1$ and α_i's are distinct, $\mathrm{rank}(\mathcal{C}) = N(L+1)$; hence, we can determine $\mathbf{h}(l)$'s uniquely. Define $\hat{\mathcal{D}}$ as in (9) with $\mathbf{d}_m$'s replaced with $\hat{\mathbf{d}}_m$'s. Then we have the channel estimate

$$\hat{\mathcal{H}} = (\mathcal{C}^H\mathcal{C})^{-1}\mathcal{C}^H\hat{\mathcal{D}}. \tag{12}$$

Lemma 1. If $P \ge L+2$ and $c_m \ne 0\ \forall m \ne 0$, then (11) has a unique solution. •

3. SYNCHRONIZATION ISSUES

Implicit in the procedure of Sec. 2 is time-synchronization of $\{c(n)\}$ at the receiver with that at the transmitter. Suppose that (compare with (5))

$$E\{\mathbf{y}(n)\} = \sum_{l=0}^{L} \mathbf{h}(l) c(n - l + n_0) + \mathbf{m}$$

$$= \sum_{m=0}^{P-1} \left[\sum_{l=0}^{L} c_m \mathbf{h}(l) e^{-j\alpha_m l} \right] e^{j\alpha_m (n+n_0)} + \mathbf{m} \tag{13}$$

where we allow an offset n_0 in $c(n)$ at the transmitter ($0 \le n_0 \le P-1$: mod P offset for obvious reasons). At the receiver n_0 is unknown and needs to be accounted for, otherwise the approach of [11] will (likely) fail. Define

$$\hat{\mathbf{d}}_m^{(i)} := \frac{1}{T}\sum_{n=1}^{T} \mathbf{y}(n) e^{-j\alpha_m (n+i)}. \tag{14}$$

With respect to (13), we have $\hat{\mathbf{d}}_m^{(n_0)} \to \mathbf{d}_m$ m.s. Our objective is to devise a method which will pick the correct i in (14), given the results for $0 \le i \le P-1$.

Define

$$\mathcal{D}^{(i)} := \begin{bmatrix} \hat{\mathbf{d}}_1^{(i)H} & \hat{\mathbf{d}}_2^{(i)H} & \cdots & \hat{\mathbf{d}}_{P-1}^{(i)H} \end{bmatrix}^H, \tag{15}$$

and

$$\hat{\mathcal{H}}^{(i)} = (\mathcal{C}^H\mathcal{C})^{-1}\mathcal{C}^H\hat{\mathcal{D}}^{(i)}. \tag{16}$$

With $\hat{\mathbf{h}}^{(i)}(l)$ denoting the channel estimate from (16), define

$$\tilde{\mathbf{y}}^{(i)}(n) := \mathbf{y}(n) - \sum_{l=0}^{L} \hat{\mathbf{h}}^{(i)}(l) c(n + i - l). \tag{17}$$

It then follows that ($\mathbf{h}^{(i)}(l)$ is the "true" value of $\hat{\mathbf{h}}^{(i)}(l)$)

$$E\{\tilde{\mathbf{y}}^{(i)}(n)\} \approx \mathbf{m} + \sum_{l=0}^{L} [\mathbf{h}(l) c(n + n_0 - l)$$

$$-\mathbf{h}^{(i)}(l) c(n + i - l)] = \begin{cases} \mathbf{m} & \text{if } i = n_0 \\ \mathbf{m}+? & \text{if } i \ne n_0. \end{cases} \tag{18}$$

For $i = n_0$, $E\{\tilde{\mathbf{y}}^{(i)}(n)\}$ is a constant, therefore, has no components with nonzero power at cycle frequencies α_m, $m = 1, 2, \cdots, P-1$. What about $i \ne n_0$? What are the $\mathbf{h}^{(i)}(l)$'s for which $E\{\tilde{\mathbf{y}}^{(i)}(n)\} = \mathbf{m}'\ \forall n$, where $\mathbf{m}'$ is some constant?

Suppose that $E\{\tilde{\mathbf{y}}^{(i)}(n)\} = \mathbf{m}'\ \forall n$. From (4) and (18), we then have

$$\sum_{m=0}^{P-1} \left[\sum_{l=0}^{L} c_m \left(\mathbf{h}(l) e^{-j\alpha_m (l-n_0)} - \mathbf{h}^{(i)}(l) e^{-j\alpha_m (l-i)} \right) \right] e^{j\alpha_m n}$$

$$= \mathbf{m}' - \mathbf{m}. \tag{19}$$

If $c_m \ne 0\ \forall m$, then for $m = 1, 2, \cdots, P-1$, (19) implies that

$$\sum_{l=0}^{L} \mathbf{h}(l) e^{-j\alpha_m l} = \sum_{l=0}^{L} \mathbf{h}^{(i)}(l) e^{-j\alpha_m l} \underbrace{e^{j\alpha_m (i-n_0)}}_{e^{j\alpha_m k}} \tag{20}$$

where $k := i - n_0$. Define

$$\mathbf{v}_l := [e^{-j\alpha_1 l},\ e^{-j\alpha_2 l},\ \cdots, e^{-j\alpha_{P-1} l}]^T. \tag{21}$$

Let $\mathcal{H}^{(i)}$ denote $\mathcal{H}$ in (8) but with $\mathbf{h}(l)$ replaced with $\mathbf{h}^{(i)}(l)$. Define

$$\Sigma_k := \mathrm{diag}\left\{ e^{j\alpha_1 k}, e^{j\alpha_2 k}, \cdots, e^{j\alpha_{P-1} k} \right\}, \tag{22}$$

$$\mathbf{V}^{(k)} := [\mathbf{v}_{-k},\ \mathbf{v}_{-k+1},\ \cdots, \mathbf{v}_{-k+L}], \quad \tilde{\mathbf{v}}_l = \mathbf{v}_l \otimes I_N. \tag{23}$$

By (20), we have

$$\left(\mathbf{V}^{(0)} \otimes I_N\right)\mathcal{H} = \left(\mathbf{V}^{(k)} \otimes I_N\right)\mathcal{H}^{(i)},\ \mathbf{V}^{(k)} := \Sigma_k \mathbf{V}^{(0)}. \tag{24}$$

By the nature of α_m's and the definition of $\mathbf{v}_l$, we have

$$\mathbf{v}_l = \mathbf{v}_{l \bmod P}. \tag{25}$$

Also $\mathbf{V}^{(0)}$, hence $\mathbf{V}^{(k)}$, is Vandermonde with rank $L+1$ ($\le P-1$); therefore, $\mathrm{rank}(\mathbf{V}^{(k)} \otimes I_N) = N(L+1)\ \forall k$.

Recall from Sec. 2 that we have $L \le P-2$ for channel identifiability. By (20), only $(i - n_0) \bmod P$ is relevant since $\alpha_m = 2\pi m/P$, $m = 0, 1, \cdots, P-1$. Therefore, we will take $k = (i - n_0) \bmod P$; hence $P - 1 \ge k \ge 0$. For $k = 0$, (19) holds true for $\mathbf{h}^{(i)}(l) = \mathbf{h}(l)$, and this solution is unique (see Sec. 2). What about $k = 1, 2, \cdots, P-1$? **We assume that** $P \ge 2L+3$. First an auxiliary result.

Prop. 1. Rank($[\tilde{\mathbf{v}}_0, \cdots, \tilde{\mathbf{v}}_n, \tilde{\mathbf{v}}_{n+2}, \cdots, \tilde{\mathbf{v}}_{P-1}]$) $= N(P-1)$ for any $0 \le n \le P-3$.
Proof: See the Appendix. •

We now turn to solutions to (19) under the assumption that $P \ge 2L+3$.

Case 1: $L-k \ge 0$. Since $1 \le k \le L$, we have $P-k \ge P-L \ge 2L+3-L = L+3$. Eqns. (20)- (25) imply that

$$0 = \mathbf{A} := \sum_{l=0}^{L} \tilde{\mathbf{v}}_l \mathbf{h}(l) - \sum_{l=0}^{L} \tilde{\mathbf{v}}_{l-k} \mathbf{h}^{(i)}(l) \tag{26}$$

$$= \sum_{l=0}^{L} \tilde{\mathbf{v}}_l \mathbf{h}(l) - \sum_{j=0}^{L-k} \tilde{\mathbf{v}}_j \mathbf{h}^{(i)}(j+k) - \sum_{j=-k}^{-1} \tilde{\mathbf{v}}_j \mathbf{h}^{(i)}(j+k). \tag{27}$$

Noting that (see (25))

$$\sum_{j=-k}^{-1} \tilde{\mathbf{v}}_j \mathbf{h}^{(i)}(j+k) = \sum_{j=-k}^{-1} \tilde{\mathbf{v}}_{j+P} \mathbf{h}^{(i)}(j+k),$$

(27) may be rewritten as

$$\mathbf{A} = \sum_{l=0}^{L-k} \tilde{\mathbf{v}}_l [\mathbf{h}(l) - \mathbf{h}^{(i)}(l+k)] + \sum_{l=L-k+1}^{L} \tilde{\mathbf{v}}_l \mathbf{h}(l)$$

$$- \sum_{l=P-k}^{P-1} \tilde{\mathbf{v}}_l \mathbf{h}^{(i)}(l+k-P) = 0. \tag{28}$$

Since $P-k \ge L+3$, we have $P-k-L \ge 3$. Therefore, by Prop. 1, (28) implies that $\mathbf{h}(L) = 0$ which is impossible if the true channel is of length $L+1$ with $\mathbf{h}(L) \ne 0$. Hence, (19) can not be true for $0 < k \le L$ with $P \ge 2L+3$.

Case 2: $L-k < 0$. Now we have

$$\mathbf{A} = \sum_{l=0}^{L} \tilde{\mathbf{v}}_l \mathbf{h}(l) - \sum_{l=P-k}^{P+L-k} \tilde{\mathbf{v}}_l \mathbf{h}^{(i)}(l+k-P) = 0. \tag{29}$$

If $L-k = -1$, then $P-k \ge 2L+3-k = 2L+3-1-L = L+2$. If $L-k \le -2$, then $P+L-k \le P-2$. Therefore, by Prop. 1, (29) implies that $\mathbf{h}(0) = 0$ which is impossible since by assumption $\mathbf{h}(0) \ne 0$ (else true channel is of length L instead of the assumed length $L+1$). Hence, (19) can not be true for $L < k \le P-1$ with $P \ge 2L+3$.

We summarize the above discussion in Lemma 2.

Lemma 2. Given model (1)-(2) with channel length $L+1$ so that $\mathbf{h}(0) \ne 0$ and $\mathbf{h}(L) \ne 0$. If $P \ge 2L+3$ and Lemma 1 holds true, then the only solution to (19) is given by $\mathbf{h}^{(i)}(l) = \mathbf{h}(l)$ for $i = n_0$, $0 \le l \le L$. •

3.1. Proposed Synchronization Method

It is based upon synchronization via exhaustive model fitting via (16) for offsets $i = 0, 1, \cdots, P-1$ and then testing to see if $E\{\tilde{\mathbf{y}}^{(i)}(n)\} = \mathbf{m}'\ \forall n$. Define

$$\hat{\mathbf{e}}_m^{(i)} := \frac{1}{T} \sum_{n=1}^{T} \tilde{\mathbf{y}}^{(i)}(n) e^{-j\alpha_m n} \tag{30}$$

and the cost

$$J_i := \sum_{m=1}^{P-1} \|\hat{\mathbf{e}}_m^{(i)}\|^2. \tag{31}$$

By Lemma 2, for some b,

$$\lim_{T\to\infty} J_i \overset{m.s.}{=} \begin{cases} 0, & i = n_0 \\ b > 0, & i \ne n_0. \end{cases} \tag{32}$$

Synchronization Algorithm

1) Execute (14)-(16) for $i = 0, 1, \cdots, P-1$.
2) Calculate $\tilde{\mathbf{y}}^{(i)}(n)$ via (17) and $\hat{\mathbf{e}}_m^{(i)}$ via (30) using the results of Step 1) for $i, m = 0, 1, \cdots, P-1$, $m \ne 0$. Determine the cost J_i via (32).
3) The estimate $\hat{n}_0$ of the offset n_0 is give by

$$\hat{n}_0 = \arg\left\{ \min_{0 \le i \le P-1} J_i \right\}. \tag{33}$$

Remark 1. The preceding developments are based on the assumption that $\mathbf{h}(0) \ne 0$ and $\mathbf{h}(L) \ne 0$. If the model order (channel length) is unknown and one overfits with assumed length $\bar{L} > L$, then while $\mathbf{h}(0) \ne 0$, $\mathbf{h}(\bar{L}) = 0$. It is relatively straightforward but tedious to establish that in this case, $\mathbf{h}^{(i)}(l)$ as a time-shifted version of $\mathbf{h}(l)$ will also satisfy (19) for some $i \ne n_0$. Nevertheless, the true channel is captured although with a time-shift. We omit the details.

Remark 2. If $\mathbf{h}(0)$ and/or $\mathbf{h}(L)$ is "small" (in magnitude), the proposed method may regard them as 0. Therefore, comparing estimated channel $\mathbf{h}^{(\hat{n}_0)}(l)$ with the true channel $\mathbf{h}(l)$ for performance comparison may not be an effective measure of the proposed method in the sense that $\mathbf{h}^{(\hat{n}_0)}(l)$ may be a shifted version of $\mathbf{h}(l)$. In Sec. 4, we have chosen to carry out performance comparison on the basis of the bit error rate (BER) of a linear MMSE equalizer designed using the estimated channel. We design a linear minimum mean-square error (LMMSE) equalizer of length L_e and equalization delay d using the estimated channel. Let $\mathbf{v}'(n) := \mathbf{v}(n) - \mathbf{m}$ and define (recall (1) and (2))

$$\mathbf{y}'(n) := \mathbf{y}(n) - \sum_{l=0}^{L} \hat{\mathbf{h}}^{(\hat{n}_0)}(l) c(n + \hat{n}_0 - l) - \hat{\mathbf{m}},$$

$$\approx \sum_{l=0}^{L} \mathbf{h}(l) b(n-l) + \mathbf{v}'(n), \tag{34}$$

where

$$\hat{\mathbf{m}} := (1/T) \sum_{n=1}^{T} [\mathbf{y}(n) - \sum_{l=0}^{L} \hat{\mathbf{h}}^{(\hat{n}_0)}(l) c(n + \hat{n}_0 - l)]. \tag{35}$$

Equalize the channel by applying the LMMSE equalizer to $\{\mathbf{y}'(n)\}$ to estimate $\{b(n)\}$ as $\{\widehat{b}(n)\}$. Quantize $\{\widehat{b}(n)\}$ into $\{\tilde{b}(n)\}$ with the knowledge of the symbol alphabet (hard decisions).

4. SIMULATION EXAMPLE

We consider a random frequency-selective Rayleigh fading channel. We took $N = 1$ and $L = 2$ in (1) with $h(l)$ complex-valued (independent real and imaginary parts), mutually independent for all l, zero-mean unit variance Gaussian. Additive noise was zero-mean complex white Gaussian. The SNR refers to the energy per bit over one-sided noise spectral density with both information and superimposed training sequence counting toward the bit energy. Information sequence as well as superimposed training was binary. We took the superimposed training sequence period $P = 7$ in (**H3**). The average transmitted

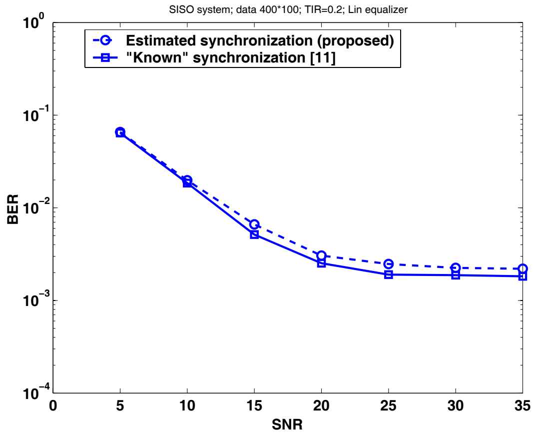

Figure 1. BER: square/solid: estimate channel using superimposed training with known synchronization and then design a linear MMSE equalizer following [11]; circle/dots: estimate channel using superimposed training with estimated synchronization and then design a linear MMSE equalizer (proposed method). Training-to-information symbol power ratio =0.2 (-7 dB). Record length = 400 bits. Results based on 100 Monte Carlo runs.

power in $c(n)$ (scaled binary) was 0.2 of the power in $b(n)$ – a small penalty in SNR. There was no loss in information rate. Linear MMSE equalizer of length 11 bits and equalization delay of 5 bits was used throughout (see Remark 2). Fig. 1 shows the BER based on 100 Monte Carlo runs, resulting from the linear MMSE equalizer designed on the basis of the estimated channel (see also [11]) using the proposed approach of estimating the synchronization delay as well as the approach of [11] based on the knowledge of true synchronization. [The true channel was different in different Monte Carlo runs.] It is seen that the proposed approach works well. The slight discrepancy between the proposed approach and that of [11] is due the reason cited in Remark 2, namely, at times small $h(0)$ and/or $h(L)$ are regarded as 0 by the proposed method which results in some degradation in the BER performance.

5. CONCLUSIONS

Approach of [11] to SIMO channel estimation using superimposed training sequences (hidden pilots) and first-order statistics is based on the assumption that the superimposed training sequence at the receiver is time-synchronized with its transmitted counterpart. In this paper we have removed this assumption of synchronization and proposed a novel approach to superimposed training synchronization. The results were illustrated via a simulation example involving frequency-selective Rayleigh fading.

The first-order statistics-based approach of [11] views the information sequence as interference. Since the training and information sequences of a given user pass through identical channel, this fact can be exploited to enhance channel estimation performance via a semiblind approach. Such an approach has been considered in a companion paper [13] assuming time-synchronization of the superimposed training sequence.

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] 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.

[5] 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.

[6] H. Holma and A. Toskala, Eds. *WCDMA for UMTS: Radio Access for Third Generation Mobile Communications*, 2nd Ed. Wiley, 2002.

[7] 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.

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

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

[10] 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.

[11] 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.

[12] 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.

[13] X. Meng and J.K. Tugnait, "Semi-Blind channel estimation and detection using superimposed training," in *Proc. ICASSP-2004*, Montreal, Canada, May 17-21, 2004.

6. APPENDIX

Proof of Prop. 1: Define the $(P-1)\times(P-1)$ matrix

$$\mathbf{W} := [\mathbf{v}_0,\ \mathbf{v}_1,\ \cdots,\ \mathbf{v}_n,\ \mathbf{v}_{n+2},\ \cdots,\ \mathbf{v}_{P-1}] \qquad (36)$$

and the $[N(P-1)]\times[N(P-1)]$ matrix

$$\tilde{\mathbf{W}} := \mathbf{W} \otimes I_N. \qquad (37)$$

We have

$$\text{rank}\left(\tilde{\mathbf{W}}\right) = N \times \text{rank}\left(\mathbf{W}\right). \qquad (38)$$

Since $\alpha_m = \alpha_{m-P}$ (recall (4) and (23)), we have

$$\mathbf{W} := [\mathbf{v}_0,\ \mathbf{v}_1,\ \cdots,\ \mathbf{v}_n,\ \mathbf{v}_{n+2-P},\ \cdots,\ \mathbf{v}_{P-1-P}]. \qquad (39)$$

Consider

$$\bar{\mathbf{W}} := [\mathbf{v}_{n+2-P},\ \mathbf{v}_{n+1-P},\ \cdots,\ \mathbf{v}_{-1},\ \mathbf{v}_0,\ \cdots,\ \mathbf{v}_n] \qquad (40)$$

which can be obtained from $\mathbf{W}$ via elementary row operations so that

$$\text{rank}\left(\bar{\mathbf{W}}\right) = \text{rank}\left(\mathbf{W}\right). \qquad (41)$$

Finally

$$\bar{\mathbf{W}} := \text{diag}\left\{e^{-j\alpha_1(n+2-P)}, \cdots, e^{-j\alpha_{P-1}(n+2-P)}\right\}\bar{\mathbf{V}} \qquad (42)$$

where $\bar{\mathbf{V}}$ is a Vandermonde matrix with $\text{rank}\left(\bar{\mathbf{V}}\right) = P-1$ and

$$\bar{\mathbf{V}} := [\mathbf{v}_0,\ \mathbf{v}_1,\ \cdots,\ \mathbf{v}_{P-2}]. \qquad (43)$$

The desired result follows from (36), (37), (38) and (43).