

## Research Article

# Optimal Superimposed Training Sequences for Channel Estimation in MIMO-OFDM Systems

Jinesh P. Nair<sup>1</sup> and Ratnam V. Raja Kumar<sup>2</sup>

<sup>1</sup>G.S. Sanyal School of Telecommunications, IIT, Kharagpur 721302, India

<sup>2</sup>Department of Electronics and Electrical Communication, IIT, Kharagpur 721302, India

Correspondence should be addressed to Jinesh P. Nair, mrjineshpn@yahoo.com

Received 7 June 2009; Revised 21 December 2009; Accepted 10 February 2010

Academic Editor: Ricardo Merched

Copyright © 2010 J. P. Nair and R. V. R. Kumar. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this work an iterative time domain Least Squares (LS) based channel estimation method using superimposed training (ST) for a Multiple Input Multiple Output Orthogonal Frequency Division Multiplexing (MIMO-OFDM) system over time varying frequency selective fading channels is proposed. The performance of the channel estimator is analyzed in terms of the Mean Square Estimation Error (MSEE) and its impact on the uncoded Bit Error Rate (BER) of the MIMO-OFDM system is studied. A new selection criterion for the training sequences that jointly optimizes the MSEE and the BER of the OFDM system is proposed. Chirp based sequences are proposed and shown to satisfy the same. These are compared with the other sequences proposed in the literature and are found to yield a superior performance. The sequences, one for each transmitting antenna, offers fairness through providing equal interference in all the data carriers unlike earlier proposals. The effectiveness of the mathematical analysis presented is demonstrated through a comparison with the simulation studies. Experimental studies are carried out to study and validate the improved performance of the proposed scheme. The scheme is applied to the IEEE 802.16e OFDM standard and a case is made with the required design of the sequence.

## 1. Introduction

Channel estimation in MIMO-OFDM is a challenging task. Conventionally this is done by making use of the comb type pilots, block type pilots, or orthogonal pilots as discussed in [1–3]. These methods give good performance at the cost of the bandwidth efficiency of the system which may be significant in a time-varying frequency selective fading channel and also as the number of transmit antennas increases.

Recently, superimposed training- (ST-) based techniques have been proposed for channel estimation. In these schemes, training symbols known to the receiver are algebraically added on to the data at a low power, thus improving the bandwidth efficiency. A drawback of ST-based methods is the mutual interference between the training sequences and data symbols on each other. Superimposed training is considered for channel estimation and equalization in single carrier systems in [4–13], for OFDM based systems in [14–23] and MIMO systems in [24–29]. However, not much

work has been done on ST-based schemes in MIMO-OFDM systems. In view of this and also the bandwidth efficiency of ST-based methods, there is a need for such a method in MIMO-OFDM systems that yields good performance.

Very importantly, the issues that govern the choice of the ST sequences for OFDM systems over frequency selective fading channels are not clearly addressed in the existing literature. Also, the extension to the MIMO-OFDM case is not obvious. The criterion used for selecting the sequences is the minimization of the MSEE or the Cramer-Rao Lower Bound (CRLB) of the channel estimate [14, 17–20]. In [7–9, 11–13, 20, 25], periodic training sequences are used, where the period is determined by the order of the channel estimator. Chirp sequences, pseudorandom binary sequences (PRBS-) and a periodic sequence with a single nonzero component at a specific location in each period were proposed. These approaches are consistent with the general approach of designing optimum and sufficient excitation for open loop informative experiments done in

system identification. However in the superimposed training case, unlike the time multiplexed training case, the channel estimation and symbol detection processes are coupled to each other.

As a result, these ST sequences may lead to a strong interference in a few subcarriers leading to a serious BER performance degradation. The sequences proposed are not obtained by minimizing the BER and may also be nonfeasible in OFDM systems that employ guard bands. This is particularly true in OFDM systems, where the performance is dominated by the higher bit error probabilities of the subcarriers. In view of this major drawback, there is a strong need for the optimization criterion of the training sequence to minimize the error in symbol detection in addition to minimizing the channel MSEE.

Motivated by the above mentioned lacunae in the existing ST-based methods, in this work an iterative time domain LS based channel estimation method using ST for a MIMO-OFDM system over time-varying frequency selective fading channels is proposed. The estimator is generalized to provide scope for exploiting the coherence time and the coherence bandwidth of the channel. The most important contribution of this work is the proposition of a criterion and then a set of optimal training sequences satisfying the same for ST-based MIMO-OFDM systems. Towards this end, the performance of the channel estimate of the proposed scheme is mathematically analyzed in terms of the MSEE and its impact on the BER of the MIMO-OFDM system is studied. By examining the MSEE and the BER, it is found and shown that choosing sequences based on the MSEE or the CRLB alone is not enough and there is scope for further minimizing the total subcarrier BER. In the quest for a solution, a new selection criterion is proposed for the training sequences that jointly optimizes the MSEE and the BER by ensuring a fair distribution of the residual interference due to the training sequence on all the data subcarriers. Such an approach has not been considered in the existing literature.

It is shown that chirp based training sequences can yield the optimum performance and such sequences are proposed for ST-based MIMO-OFDM systems. The sequences are generalized to incorporate the use of guard subcarriers in the OFDM based system. The energy of the proposed sequences is equally distributed among all the data subcarriers. As a result the proposed sequences, one for each transmitting antenna, offers fairness through providing equal interference in all the data carriers unlike the earlier proposals. The proposed training sequences are compared with the other sequences proposed in the literature like the chirp based sequences [11, 20, 25] and also the maximum length PN sequences. Overall, the proposed sequence is shown to significantly improve the BER performance when compared to the existing schemes.

Experimental studies are carried out to study the performance of the proposed scheme and also its comparison with existing schemes with and without guard subcarriers presented. The effectiveness of the mathematical analysis presented in this work is demonstrated through a comparison with the simulation studies. The proposed scheme is applied to the IEEE 802.16e standard, its suitability is

examined, and a case is made with the required design of the sequence.

The organization of this paper is as follows. In Section 2, the system model for the ST-based MIMO-OFDM system is presented and incorporated into a system employing two-branch transmit diversity using Alamouti space-time coding. Then, the proposed iterative time domain based channel estimator is explained. Section 3 gives the performance analysis in terms of the MSEE and BER, which are jointly optimized to propose the selection criterion for the optimal training sequences. The optimal chirp based training sequences are proposed and compared with the other sequences in the literature in Section 4. Section 5 presents the experimental results in terms of the MSEE and the BER and conclusions are drawn in Section 6.

*Notations.* Time-domain matrices are written as  $\mathbf{H}$  and vectors as  $\mathbf{h}$ . Their frequency domain counterparts are  $\bar{\mathbf{H}}$  and  $\bar{\mathbf{h}}$ , respectively. Signals in the time domain are written as  $h(k)$  in frequency domain  $H(k)$ . Superscripts  $(\cdot)^T$ ,  $(\cdot)^H$ , and  $(\cdot)^*$  denote the operations matrix transpose, Hermitian, and complex conjugation respectively, and  $\|\cdot\|$  denotes 2 norms.

## 2. Superimposed Training in MIMO-OFDM Systems

*2.1. Channel Model.* The complex baseband impulse response of a time-varying frequency selective fading channel between the  $m$ th transmit antenna and the  $r$ th receive antenna  $h_{m,r}^c(t, \tau)$  can be given as

$$h_{m,r}^c(t, \tau) = \sum_{i=0}^{L-1} h_{m,r}^l(t) \delta(\tau - \tau_i), \quad (1)$$

where  $h_{m,r}^l$  is the time-varying gain of the  $i$ th path,  $\tau_i$  is the delay of the  $i$ th, path and  $L$  is the number of propagation paths. The channel impulse response  $h_{m,r}^c(t, \tau)$  is modeled as a zero-mean complex Gaussian process with paths being uncorrelated. Each path fades independently according to the Jakes' power spectrum with the time-correlation function  $E[h_{m,r}^k(t)h_{m,r}^{*k}(t')] = \sigma_k^2 J_0[2\pi f_d(t - t')]$ , where  $\sigma_k^2$  is the variance of the  $k$ th path,  $J_0(\cdot)$  is the zeroth order Bessel function of the first kind, and  $f_d$  is the Doppler frequency in Hz.

*2.2. System Model.* Consider a MIMO-OFDM system with  $N_t$  transmitting antennas and  $N_r$  receiving antennas having  $N$  subcarriers and an OFDM symbol duration  $T_s$ . Let  $\bar{\mathbf{s}}_t(k)$  be a vector of data symbols associated with the  $k$ th OFDM symbol of the  $t$ th transmitting antenna  $t \in [1, N_t]$  given by

$$\bar{\mathbf{s}}_t(k) = [S_{t,0}(k) \cdots S_{t,N-1}(k)]^T. \quad (2)$$

These are then passed to the  $N$  point IDFT block whose output may be expressed as

$$\mathbf{s}_t(k) = \mathbf{F}^{-1} \bar{\mathbf{s}}_t(k), \quad (3)$$

where  $\mathbf{F}$  is the  $N \times N$  DFT matrix defined as

$$\mathbf{F}(m, n) = e^{-j2\pi mn/N}. \quad (4)$$

Here  $(m, n)$  is used to denote the  $m$ th row and  $n$ th column of the matrix.

Superimposed training (ST) sequences are added to the data symbols for the purpose of channel estimation. Let the training sequence  $\mathbf{c}_t(k)$  be algebraically added to this IDFT output with a specific data to pilot power ratio  $\alpha = \sigma_{st}^2/\sigma_{ct}^2$ , to get

$$\mathbf{x}_t(k) = \mathbf{s}_t(k) + \mathbf{c}_t(k). \quad (5)$$

The architecture of the transmit part is illustrated in Figure 1. After inserting a CP in the guard intervals between adjacent IDFT blocks, the OFDM symbols are transmitted over a time-varying frequency selective fading channel.

Let  $\mathbf{h}_{t,r}(k) = [h_{t,r}^0(k) \ h_{t,r}^1(k) \ \dots \ h_{t,r}^{L-1}(k)]^T$  be the  $L \times 1$  channel vector associated with the  $k$ th OFDM symbol between the  $t$ th transmit antenna and the  $r$ th receive antenna. Here  $h_{t,r}^l(k) = h_{t,r}^l(t = kT_s)$  in (1). Hence the channel varies from symbol to symbol depending on the coherence time of the channel. Unrestrictively, all the channel vectors are considered to be of equal length  $L$ . Here  $L$  depends on the delay spread of the channel. Let  $\mathbf{H}_{t,r}(k)$  with dimension  $N \times N$  be the corresponding circulant matrix of the channel formed as a result of cyclic prefixing as discussed in [30].

We assume that the antenna elements at the transmitter, and also at the receiver, are uncorrelated with each other. This enables the consideration of the channel estimate for each transmit-receive antenna pair at each receiver independent of the other receivers. The equivalent baseband vectors at the  $r$ th receive antenna  $r \in [1, N_r]$ , obtained after guard removal from the OFDM symbols may then be expressed as

$$\begin{aligned} \mathbf{y}_r(k) &= \sum_{t=1}^{N_t} [\mathbf{S}_t(k) + \mathbf{C}_t(k)] \mathbf{h}_{t,r}(k) + \mathbf{w}_r(k) \\ &= \sum_{t=1}^{N_t} \mathbf{H}_{t,r}(k) [\mathbf{s}_t(k) + \mathbf{c}_t(k)] + \mathbf{w}_r(k). \end{aligned} \quad (6)$$

Here  $\mathbf{w}_r(k)$  is the AWGN associated with the  $k$ th transmitted vector at the  $r$ th receive antenna.  $\mathbf{S}_t(k)$  and  $\mathbf{C}_t(k)$  are matrices associated with the data and training, respectively, having dimension  $N \times L$ .

We make the following assumptions for the present formulation and study.

- (A1) The channel is wide sense stationary with uncorrelated scattering (WSSUS),  $E(\mathbf{h}_{t,r}(k) \mathbf{h}_{t,r}^H(k)) = \text{diag}(E|h_{t,r}^0(k)|^2, E|h_{t,r}^1(k)|^2, \dots, E|h_{t,r}^{L-1}(k)|^2)$ , with the delay profile normalized to have unit area,  $\sum_{i=0}^{L-1} E(|h_{t,r}^i(k)|^2) = 1$ . Here  $\text{diag}(\cdot)$  stands for the diagonal matrix.
- (A2) The data symbols are uncorrelated with each other and has zero mean. That is,  $E\{\mathbf{s}_t(k) \mathbf{s}_t^H(k)\} = \sigma_{st}^2 \mathbf{I}$ , and  $E\{\mathbf{S}_t(k)\} = \mathbf{0}$ .

(A3) The data is uncorrelated with the superimposed training sequence and also due to (A2), we have  $E\{\mathbf{C}_t^H(k) \mathbf{S}_t(k)\} = E\{\mathbf{S}_t^H(k) \mathbf{C}_t(k)\} = \mathbf{0}$ .

(A4) The noise is additive white Gaussian with variance  $\sigma_{wr}^2$  and mean zero. That is,  $E\{\mathbf{w}_r(k) \mathbf{w}_r^H(k)\} = \sigma_{wr}^2 \mathbf{I}$  and  $E\{\mathbf{w}_r(k)\} = \mathbf{0}$ .

We also assume perfect synchronization and a zero DC offset at the baseband receiver. Methods to deal with this are discussed in [7, 31, 32]. A time-domain channel estimation technique, to be discussed later, is carried out to obtain the estimate  $\hat{\mathbf{h}}_{t,r}(k)$  of the fading channel.

The received vectors are input to the DFT block for demodulation and  $\bar{\mathbf{y}}_r(k) = \mathbf{F} \mathbf{y}_r(k)$  is obtained.  $\hat{\mathbf{h}}_{t,r}(k) = \mathbf{F} \hat{\mathbf{h}}_{t,r}(k)$  is then multiplied with  $\bar{\mathbf{c}}_t(k) = \mathbf{F} \mathbf{c}_t(k)$  which is known to the receiver. These are then subtracted from  $\bar{\mathbf{y}}_r(k)$  to reduce the effect of the interference due to the training sequence on the data symbols as shown in Figure 2. The frequency domain output  $\bar{\mathbf{z}}_r(k)$  is now given by

$$\bar{\mathbf{z}}_r = \sum_{t=1}^{N_t} \left[ \bar{\mathbf{H}}_{t,r}(k) \bar{\mathbf{s}}_t(k) + \left( \bar{\mathbf{H}}_{t,r}(k) - \hat{\bar{\mathbf{H}}}_{t,r}(k) \right) \bar{\mathbf{c}}_t(k) \right] + \bar{\mathbf{w}}_r(k), \quad (7)$$

where  $\bar{\mathbf{H}}_{t,r}(k) = \text{diag}[H_{tr,0}(k) \dots H_{tr,N-1}(k)]$  and  $\hat{\bar{\mathbf{H}}}_{t,r}(k) = \text{diag}[\hat{H}_{tr,0}(k) \dots \hat{H}_{tr,N-1}(k)]$  are diagonal matrices with diagonal entries being the frequency coefficients of the channel and their estimates, respectively,  $\bar{\mathbf{c}}_t(k) = \mathbf{F} \mathbf{c}_t(k) = [C_{t,0}(k) \dots C_{t,N-1}(k)]^T$  and  $\bar{\mathbf{w}}_r(k) = \mathbf{F} \mathbf{w}_r(k) = [W_{r,0}(k) \dots W_{r,N-1}(k)]^T$ .

These are then presented to the other blocks of a typical space-time communication receiver for signal detection.

### 2.3. The Least Squares-Based Channel Estimation Method.

Let composite vectors and matrices be formed for each receive antenna from the corresponding vectors and matrices in (6) to help formulating the proposed system. These are defined as

$$\begin{aligned} [\mathbf{h}_r(k)]_{N_t L \times 1} &= [\mathbf{h}_{1,r}^T(k) \ \mathbf{h}_{2,r}^T(k) \ \dots \ \mathbf{h}_{N_t,r}^T(k)]^T, \\ [\mathbf{C}(k)]_{N \times N_t L} &= [\mathbf{C}_1(k) \ \mathbf{C}_2(k) \ \dots \ \mathbf{C}_{N_t}(k)], \\ [\mathbf{S}(k)]_{N \times N_t L} &= [\mathbf{S}_1(k) \ \mathbf{S}_2(k) \ \dots \ \mathbf{S}_{N_t}(k)] \end{aligned} \quad (8)$$

for  $k = 0, 1, T_p - 1$ . The subscript in  $[\cdot]_{M \times N}$  denotes the matrix dimension. Hence (6) can be rewritten as

$$\mathbf{y}_r(k) = [\mathbf{S}(k) + \mathbf{C}(k)] \mathbf{h}_r(k) + \mathbf{w}_r(k). \quad (9)$$

The estimate of the channel impulse response for each transmit-receive antenna pair at the  $r$ th received antenna can be determined by minimizing the least squares difference between the received signal and the sum of the outputs of the ST sequences passed through the channel estimators.

The fading nature of the channel encountered in practice in high data rate wireless communication systems allows us

to assume that the channels experienced by a plurality of consecutive OFDM symbols are highly correlated, when it amounts to a fraction of the coherence time of the channel  $T_p$ . For a conservative design that ensures highly correlated channel samples

$$T_p \leq \frac{9}{(2 \times 16\pi f_{d\max} T_s)}, \quad (10)$$

where  $f_{d\max}$  is the maximum Doppler frequency and  $T_s$  is the OFDM symbol duration [33]. In order to facilitate averaging over these consecutive OFDM symbols and obtain good channel estimates by exploiting the coherence time of the channel, the least squares difference over  $T_p$  symbols is considered in the criterion. This is under the assumption that the channel estimate is nearly the same for  $T_p$  OFDM symbols

$$\hat{\mathbf{h}}_{t,r}(k) \approx \hat{\mathbf{h}}_{t,r}, \quad \hat{\mathbf{h}}_{t,r}(k) \approx \hat{\mathbf{h}}_{t,r}, \quad k \in [0, T_p - 1]. \quad (11)$$

Therefore from (8) and (9), the cost function  $\mathbf{J}_r$  for the  $r$ th receive antenna to be minimized is given by

$$\mathbf{J}_r = \sum_{k=0}^{T_p-1} (\mathbf{y}_r(k) - \mathbf{C}(k)\hat{\mathbf{h}}_r)^H (\mathbf{y}_r(k) - \mathbf{C}(k)\hat{\mathbf{h}}_r). \quad (12)$$

Here  $\hat{\mathbf{h}}_r = [\hat{\mathbf{h}}_{1,r}^T \cdots \hat{\mathbf{h}}_{N_t,r}^T]^T$  is an  $N_t Q \times 1$  composite channel estimation vector associated with the  $r$ th receive antenna corresponding to  $\mathbf{h}_r$  in (9). These are made up of the channel estimation vectors  $\hat{\mathbf{h}}_{t,r}^T$ ,  $t \in [1, N_t]$ , with dimension  $Q \times 1$  associated with each transmit antenna and the  $r$ th receive antenna. Here  $Q \geq L$  is the upper bound on the order of the channel model. This should be no lesser than the delay spread of the channel to ensure channel identifiability and as close to  $L$  as possible for improved estimation accuracy. Also  $N \geq N_t Q$  in order to ensure unique LS channel estimates.

Accordingly, the dimension of  $\mathbf{C}_t(k)$ ,  $t \in [1, N_t]$ , in (8) is modified to  $N \times (Q \geq L)$  and used to form the composite matrix  $\mathbf{C}(k) = [\mathbf{C}_1(k) \ \mathbf{C}_2(k) \ \cdots \ \mathbf{C}_{N_t}(k)]$  with dimension  $N \times N_t Q$  in (12). By differentiating (12) with respect to  $\hat{\mathbf{h}}_r$  and equating to zero, the estimate is given by

$$\hat{\mathbf{h}}_r = \left( \sum_{k=0}^{T_p-1} \mathbf{C}^H(k) \mathbf{C}(k) \right)^{-1} \left( \sum_{k=0}^{T_p-1} \mathbf{C}^H(k) \mathbf{y}_r(k) \right). \quad (13)$$

Using (13) with the summation over  $T_p$  number of samples, the estimator can be made to exploit the delay spread of the channel since  $(Q \geq L) \ll N$ . Zeros are padded from  $Q + 1$  to  $N$  and an  $N$  point DFT interpolation on the channel impulse response estimate is performed to obtain the estimate of the frequency coefficients of the channel at all the subcarriers with good performance. In this way the estimator is generalized to exploit the coherence time and the coherence bandwidth of the channel. Hence it is of improved accuracy and is more suitable for practical implementation.

**2.4. Space-Time Coded ST-Based MIMO-OFDM System.** The system model and the channel estimate presented above

are applicable to any MIMO-OFDM system employing superimposed training. Without any loss in generality, in this paper we consider a case that employs a two-branch transmitter diversity using the Alamouti space-time coding scheme with one receiver [34] as shown in Figures 1 and 2. Binary data is grouped into symbols depending on the level of QAM modulation used. The average power of the symbols in each subcarrier is denoted by  $\sigma_s^2$ . Vectors  $\dots \bar{\mathbf{s}}_2, \bar{\mathbf{s}}_1$  are formed with dimension  $N \times 1$  and input to the space-time encoder to get

$$\begin{aligned} \bar{\mathbf{s}}_1(k) &= \bar{\mathbf{s}}_1, & \bar{\mathbf{s}}_2(k) &= \bar{\mathbf{s}}_2, \\ \bar{\mathbf{s}}_1(k+1) &= -\bar{\mathbf{s}}_2^*, & \bar{\mathbf{s}}_2(k+1) &= \bar{\mathbf{s}}_1^*. \end{aligned} \quad (14)$$

Let

$$\mathbf{E}_{t,1} = (\bar{\mathbf{H}}_{t,1} - \hat{\bar{\mathbf{H}}}_{t,1}). \quad (15)$$

Substituting (14) and (15) in (7) the received frequency domain vectors of the two successively space-time encoded vectors can be expressed as

$$\bar{\mathbf{z}}_1(k) = \bar{\mathbf{H}}_{1,1} \bar{\mathbf{s}}_1 + \bar{\mathbf{H}}_{2,1} \bar{\mathbf{s}}_2 + \sum_{t=1}^2 \bar{\mathbf{E}}_{t,1} \bar{\mathbf{c}}_t + \bar{\mathbf{w}}_1(k), \quad (16)$$

$$\bar{\mathbf{z}}_1(k+1) = -\bar{\mathbf{H}}_{1,1} \bar{\mathbf{s}}_2^* + \bar{\mathbf{H}}_{2,1} \bar{\mathbf{s}}_1^* + \sum_{t=1}^2 \bar{\mathbf{E}}_{t,1} \bar{\mathbf{c}}_t + \bar{\mathbf{w}}_1(k+1).$$

These are then presented to the space-time decoder to get

$$\tilde{\mathbf{s}}_1 = \bar{\mathbf{H}}_{1,1}^* \bar{\mathbf{z}}_1(k) + \bar{\mathbf{H}}_{2,1} \bar{\mathbf{z}}_1^*(k+1), \quad (17)$$

$$\tilde{\mathbf{s}}_2 = \bar{\mathbf{H}}_{2,1}^* \bar{\mathbf{z}}_1(k) - \bar{\mathbf{H}}_{1,1} \bar{\mathbf{z}}_1^*(k+1). \quad (18)$$

Hard decisions are taken on  $\tilde{\mathbf{s}}_1$  and  $\tilde{\mathbf{s}}_2$  to obtain  $\hat{\mathbf{s}}_1$  and  $\hat{\mathbf{s}}_2$ , respectively. These may be space time encoded again to obtain  $\hat{\mathbf{s}}_i(k)$  to be used in an iterative fashion for joint channel estimation and data detection as described next.

**2.5. Data-Aided Iterative Channel Estimation.** The detected data vectors  $\tilde{\mathbf{s}}_i(k)$  may be used to further improve the channel estimation accuracy in an iterative fashion as discussed in [22, 23] at the cost of increased complexity and system latency.

The estimator in Section 2.3 is directly incorporated into such a scheme as shown in Figure 2, in which case the channel estimate is of improved accuracy over the existing estimators. In this case, for the  $i$ th iteration, the channel estimate  $\hat{\mathbf{h}}_r$  for  $i \geq 1$  can be obtained as

$$\hat{\mathbf{h}}_r = \left( \sum_{k=0}^{T_p-1} {}^i \mathbf{C}^H(k) {}^i \mathbf{C}(k) \right)^{-1} \left( \sum_{k=0}^{T_p-1} {}^i \mathbf{C}^H(k) \mathbf{y}_r(k) \right), \quad (19)$$

where

$${}^i \mathbf{C}(k) = {}^{i-1} \hat{\mathbf{S}}(k) + \mathbf{C}(k), \quad \forall i \geq 1 \quad (20)$$



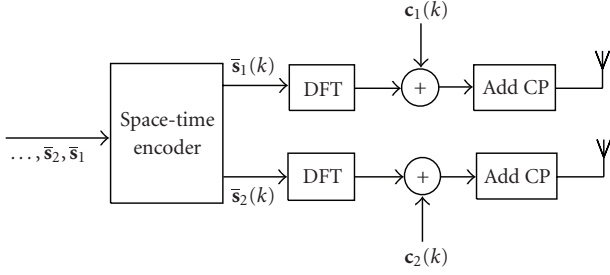


FIGURE 1: Typical space-time MIMO-OFDM transmitter employing superimposed training. The ST sequences are added to each transmit antenna with a specific data to pilot power ratio  $\alpha = \sigma_{st}^2/\sigma_{ct}$ .

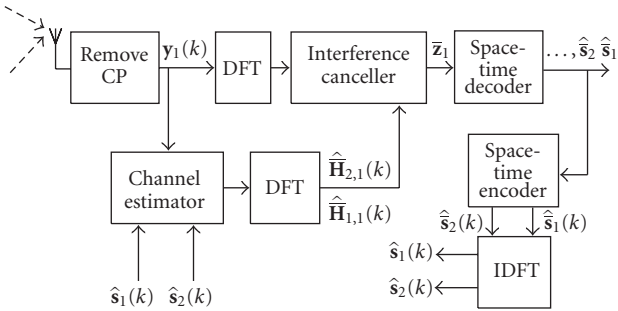


FIGURE 2: Typical space-time MIMO-OFDM receiver employing superimposed training with iterative data-aided channel estimation. The proposed channel estimation method is applied in the time domain and used along with the ST sequences to reduce the training sequence interference on the data.

Here  $[^i\hat{\mathbf{S}}(k)]_{N \times N_t Q}$  is defined as,

$$^i\hat{\mathbf{S}}(k) = \begin{cases} [^i\mathbf{S}_1(k) \ ^i\mathbf{S}_2(k) \ \cdots \ ^i\mathbf{S}_{N_t}(k)], & i \geq 1, \\ 0, & i = 0. \end{cases} \quad (21)$$

The elements  $[^i\hat{\mathbf{S}}_t(k)]_{N \times Q}$  are the matrices obtained from the time-domain vector of the detected symbols  $\hat{\mathbf{s}}_t(k) = \mathbf{F}^{-1}\hat{\mathbf{s}}_t(k)$ , refer to Section 2.4.

Similarly the data output after interference reduction for each iteration is modified in a straightforward manner by replacing the diagonal matrix of the channel frequency response estimate  $\hat{\mathbf{H}}_{t,r}$  by  $^i\hat{\mathbf{H}}_{t,r}$  in (7)

$$^i\mathbf{z}_r(k) = \sum_{t=1}^{N_t} \left[ \bar{\mathbf{H}}_{t,r}(k) \bar{\mathbf{s}}_t(k) + \left( \bar{\mathbf{H}}_{t,r}(k) - ^i\hat{\mathbf{H}}_{t,r} \right) \bar{\mathbf{c}}_t(k) \right] + \bar{\mathbf{w}}_r(k), \quad k \in [0, T_p - 1]. \quad (22)$$

Figure 2 along with (7), (13), (19) and (22) constitutes the iterative time-domain LS-based channel estimation for the ST-based MIMO-OFDM system.

### 3. Performance Analysis

The MSEE in the channel estimate is analyzed, and its impact on the BER of the OFDM system is studied. These are

then used to propose a selection criterion for the training sequences.

**3.1. Mean Square Estimation Error.** In this analysis, since  $T_p$  in (10) ensures that the channels are highly correlated, we assume  $\mathbf{h}_r(k) \approx \mathbf{h}_r, k \in [0, T_p - 1]$ . The error in the initial channel estimate  $\mathbf{e} = \mathbf{h}_r - \hat{\mathbf{h}}_r$  in (13) can be expressed as

$$\mathbf{e} = \left( \sum_{k=0}^{T_p-1} \mathbf{C}^H(k) \mathbf{C}(k) \right)^{-1} \times \sum_{k=0}^{T_p-1} \left( \mathbf{C}^H(k) \mathbf{S}(k) \mathbf{h}_r + \mathbf{C}^H(k) \mathbf{w}_r(k) \right). \quad (23)$$

The MSEE in the channel estimate may then be found using

$$\sigma_e^2 = \text{tr} \{ E(\mathbf{e} \mathbf{e}^H) \}. \quad (24)$$

Here  $\text{tr}(\cdot)$  stands for the trace of the matrix.

Making use of (23) and (24) and the assumptions (A1)–(A4) and also after some matrix manipulations we get

$$\sigma_e^2 = \left( \sum_{t=1}^{N_t} \sigma_{st}^2 + \sigma_{wr}^2 \right) \text{tr} \left\{ \left( \sum_{k=0}^{T_p-1} \mathbf{C}^H(k) \mathbf{C}(k) \right)^{-H} \right\}. \quad (25)$$

We now minimize this MSEE under a power constraint on the training sequence. In (25), the term to be minimized is the trace of the inverse of a positive definite matrix  $\mathbf{R}_{cc} = \sum_{k=0}^{T_p-1} \mathbf{C}_k^H \mathbf{C}_k$ . Let  $c_{ii}$  the  $i$ th diagonal element of  $\mathbf{R}_{cc}$ , where  $c_{ii} = N T_p \sigma_c^2$ , for all  $i \in [0, Q - 1]$ , and  $C_{ii}$  be the  $i$ th diagonal element of  $\mathbf{R}_{cc}^{-1}$ , where  $C_{ii} = \text{Cof}(c_{ii})/|\mathbf{R}_{cc}|$ . Here  $\text{Cof}(c_{ii})$  is the cofactor of the  $i$ th diagonal element.

From the previous discussion and using the property of positive definite matrices,  $0 < |\mathbf{R}_{cc}| \leq c_{ii} \text{Cof}(c_{ii})$ , we see that  $C_{ii} \geq 1/c_{ii}$ . Using this in (25), we see that  $\sigma_e^2 \geq (\sum_{t=1}^{N_t} \sigma_{st}^2 + \sigma_{wr}^2) \sum_{i=0}^{Q-1} 1/c_{ii}$ . The MSEE is minimized when the equality condition is achieved in which case the training sequence that minimizes the MSEE under a power constraint on the training sequences satisfies the following property:

$$\mathbf{R}_{cc} = \sum_{k=0}^{T_p-1} \mathbf{C}(k)^H \mathbf{C}(k) = N T_p \sum_{t=1}^{N_t} \sigma_{ct}^2 \mathbf{I}_{N_t Q \times N_t Q}. \quad (26)$$

This equality condition also minimizes the MSEE given by

$$\sigma_{e \min}^2 = \frac{\left( \sum_{t=1}^{N_t} \sigma_{st}^2 + \sigma_{wr}^2 \right) N_t Q}{N T_p \sum_{t=1}^{N_t} \sigma_{ct}^2}. \quad (27)$$

From (26) and also from the definition of  $\mathbf{C}(k)$  in Section 2.3 we get

$$\sum_{k=0}^{T_p-1} \mathbf{C}_i^H(k) \mathbf{C}_j(k) = \begin{cases} N T_p \sigma_{ct}^2 \mathbf{I}_{Q \times Q}, & i = j, \\ 0, & i \neq j. \end{cases} \quad (28)$$

Hence for MSEE minimization we see that the ST sequences associated with the transmitting antennas are required to

form an orthogonal set. It can also be proved that the autocorrelation matrix of the training sequence for each transmitting antenna satisfying (28) for  $i = j$  will have all their eigenvalues  $\{\lambda_q : 0 \leq q \leq Q - 1\}$  real and identical. The implication of this in the frequency domain is that the training sequence occupies a set of  $Q \geq L$  equally spaced tones in the full band of the OFDM symbol with equipowered components.

However, in most of the currently standardized OFDM-based systems, guard subcarriers are used to enable brick wall shaping of the transmit spectrum. The sequence that is equipowered and equispaced over the full band may still be used, but this puts a restriction on the usable guard subcarriers which may not be permitted by the standard. As a result, in general, in the presence of such guard bands the training sequence may not be able to satisfy (26) and hence the minimum MSEE in (27) cannot be achieved.

In such situations the closest we can get to (26) is by minimizing the off-diagonal decay in  $\mathbf{R}_{cc}$ . Equivalently, the condition number of  $\mathbf{R}_{cc}$ , that is, the ratio of the maximum to the minimum eigenvalues ( $\lambda_{\max}/\lambda_{\min}$ ), should be as close as possible to 1. Towards this end in systems like the IEEE 802.16e, in the OFDM mode, the pilots that are placed are equipowered and equispaced over a subset of the total number of subcarriers.

This approach of minimizing the MSEE for obtaining the optimal training sequences is adequate for pilot-assisted methods because the channel estimation and symbol detection processes are decoupled from each other. In ST-based methods, the training sequence is nonorthogonal to data as seen from (7) and (27). Hence it is important that the effect of the training sequences on the symbol detection process be taken into consideration in the criterion of optimization that the training sequence should satisfy.

**3.2. Bit Error Probability.** Here we study the impact of the ST sequence and channel estimation error on the BER of the OFDM system over a frequency selective channel and use the expression to obtain the training sequence selection criterion.

For simplicity we consider the Alamouti scheme with 2 transmit antennas and one receive antenna. Following a similar procedure, the analysis may be easily extended to the case with two transmit and  $N_r$  receive antennas [34]. Moreover, since the antenna correlations are assumed to be negligible, the selection criterion for the training sequences is dependant only on the number of transmit antennas and independent of the number of receive antennas.

The overall bit error probability  $P_I(E | \gamma)$  of the OFDM system can be expressed in terms of the subcarrier bit error probabilities as

$$P_I(E | \gamma) = \frac{1}{\sum_{j=\langle N' \rangle} I_j} \sum_{j=\langle N' \rangle} I_j P_I(E | \gamma_j). \quad (29)$$

Here  $N'$  is the number of subcarriers used for data transmission,  $I_j$  is the number of information bits allocated for the  $j$ th subcarrier, and  $P_I(E | \gamma_j)$  is the bit error probability of

the  $j$ th subcarrier. If all the data subcarriers are loaded with the same number of bits, (29) can be simplified as

$$P_I(E | \gamma) = \frac{1}{N'} \sum_{j=\langle N' \rangle} P_I(E | \gamma_j). \quad (30)$$

Substituting (15)-(17) in (17) and also using (7), the space-time decoded symbol  $\tilde{\mathbf{s}}_1$  corresponding to  $\mathbf{s}_1$  for the  $j$ th subcarrier is obtained as

$$\begin{aligned} \tilde{\mathbf{s}}_{1,j} &= [H_{11,j} + E_{11,j}]^* \\ &\times \left[ H_{11,j} S_{1,j}(k) + H_{21,j} S_{2,j}(k) \right. \\ &\quad \left. + \sum_{t=1}^2 E_{t1,j} C_{t,j}(k) + W_{1,j}(k) \right] \\ &+ [H_{21,j} + E_{21,j}] \\ &\times \left[ -H_{11,j} S_{2,j}^*(k) + H_{21,j} S_{1,j}^*(k) \right. \\ &\quad \left. + \sum_{t=1}^2 E_{t1,j} C_{t,j}(k+1) + W_{1,j}(k+1) \right]^*. \end{aligned} \quad (31)$$

From (31) the signal to interference and noise ratio at each subcarrier ( $\gamma_j$ ) may be expressed as

$$\gamma_j = \frac{\sum_{t=1}^2 \|H_{t1,j}\|^2 \sigma_s^2}{\sum_{t=1}^2 \|E_{t1,j}\|^2 (\sigma_s^2 + \sigma_{s_j}^2 + \mathfrak{X}) + \sigma_W^2}, \quad (32)$$

where  $\mathfrak{X}$  denotes  $(\sum_{t=1}^2 \|E_{t1,j}\|^2 \sigma_{s_j}^2 + \sigma_W^2) / \sum_{t=1}^2 \|H_{t1,j}\|^2$ . Here,  $\sigma_s^2$ , ( $0 \leq \sigma_{s_j}^2 \leq N/Q \sum_{t=1}^2 \sigma_{ct}$ ) and  $\sigma_W^2$  are the average powers of the data symbol, training, and noise, respectively, associated with the subcarriers.

We assume that the correlation between the channel estimate and the estimation noise is negligible. Also, we approximate the effective interference and noise in (32) to be Gaussian distributed. These assumptions are not always valid. Nevertheless, the channel estimate with length  $Q$  is obtained by averaging the cross-correlation of the received symbols with the training sequence over large data lengths. Also, there is an additional DFT interpolation involved in obtaining the channel frequency response estimate at each subcarrier. This considerably reduces the interference terms and also induces Gaussianity in the noise terms affecting the symbol in each subcarrier. Hence as seen later in the experimental studies in Section 5.1, the approximate BER obtained using this SINR expression is relatively accurate.

From (30) and (32), the probability of bit error  $P_I(E | \gamma)$  of the OFDM system conditioned on the instantaneous

SINR, when the modulated symbols used are M-QAM, is given by

$$P_I(E | \gamma) \approx \frac{1}{\sum_{j \in \langle N' \rangle} I_j} \times \sum_{j \in \langle N' \rangle} \frac{4}{\log_2 M_j} (\mathfrak{T}) I_j Q \left\{ \sqrt{\frac{3}{M_j - 1}} \gamma_j \right\}. \quad (33)$$

where  $\mathfrak{T}$  denotes  $1 - 1/\sqrt{M_j}$ . Here  $M_j$  denotes the level of QAM modulation used on the  $j$ th subcarrier. It may be noted that the above expression is fairly general to incorporate adaptive modulation and also bit loading scenarios.

For random fading channels the expression for the probability of bit error  $P_I(E)$  averaged over the fading statistics may be obtained by integrating (33) over the probability distribution function of  $\gamma_j$ ,  $p(\gamma_j)$

$$P_I(E) = \frac{1}{\sum_{j \in \langle N' \rangle} I_j} \sum_{j \in \langle N' \rangle} I_j \int_0^\infty P_I(E | \gamma_j) p(\gamma_j) d\gamma_j. \quad (34)$$

**3.3. A Criterion for Minimizing the BER.** In general from (29) it is clearly seen that the BER of the OFDM system is dominated by those subcarriers which have high bit error probabilities. Equations (32) and (33) show that the BER of the subcarriers is affected by the parameters of the training sequence. Hence merely ensuring minimum MSEE yielding sequences in (26) is not enough and there is scope for further minimizing the total subcarrier BER. In order to achieve the same, in this section, a selection criterion for the training sequences minimizing the probability of symbol error expression in (33) is proposed. ST sequences that minimize (33) also minimize (34) as seen later in Sections 5.2.1 and 5.3.1.

Let us consider a frequency selective fading channel with the coherence bandwidth extending over  $N_c$  subcarriers. The fading coefficients of the channel experienced by the transmitted symbols in these subcarriers are nearly the same. This is true for OFDM systems in general. Let the total number of subcarriers  $N$  be divided into  $B_c = N/N_c$  such subbands. Let  $\Phi_b$  represent the set of subcarriers associated with the  $b$ th subband. From the above discussion for the  $b$ th subband we have

$$\sum_{t=1}^2 \|H_{t1,j}\|^2 = \|H_{t1,b}\|^2, \quad j \in \Phi_b, \quad (35)$$

$$\sum_{t=1}^2 \|E_{t1,j}\|^2 = \|E_{t1,b}\|^2, \quad j \in \Phi_b. \quad (36)$$

Since the channel coefficients are the same, in the  $b$ th subband, the channel estimation error is also the same. Also, from (25) and (27), it is readily seen that the magnitude square of the channel estimation error in the average sense

is indeed the same in all the subcarriers. In each of these  $B_c$  subbands the SINR may be expressed as

$$\gamma_{bj} = \frac{\sum_{t=1}^2 \|H_{t1,b}\|^2 \sigma_s^2}{\sum_{t=1}^2 \|E_{t1,b}\|^2 (\sigma_s^2 + \sigma_{sj}^2 + \mathfrak{z}) + \sigma_W^2}, \quad (37)$$

$$b \in [0, B_c - 1], \quad j \in \Phi_b,$$

where  $\mathfrak{z}$  denotes  $(\sum_{t=1}^2 \|E_{t1,b}\|^2 \sigma_{sj}^2 + \sigma_W^2) / \sum_{t=1}^2 \|H_{t1,b}\|^2$ . From (35), (36), and (37), we see that for each of the subbands, the SINR varies only due to the training sequence power  $\sigma_{sj}^2$ . Moreover, for similar reasons the level of modulation and the number of bits allocated in all the subcarriers of the  $b$ th subband are the same. This is true in scenarios where the OFDM system employs adaptive modulation and bit loading depending on subchannel gains, and also when all the subcarriers are modulated with the same number of bits. Hence we have

$$I_j = I_b, \quad M_j = M_b, \quad j \in \Phi_b. \quad (38)$$

Making use of (37) and (38) in (33), and invoking the Jensen's Inequality for convex functions, the expression for the probability of bit error for the  $b$ th subband can be shown to be lower bounded by

$$P_I^b(E | \gamma) \geq \frac{4}{\log_2 M_b} \left(1 - \frac{1}{\sqrt{M_b}}\right) Q \left\{ \frac{1}{N_c} \sum_{j \in \langle N_c \rangle} \sqrt{\frac{3}{M_b - 1}} \gamma_{bj} \right\}, \quad (39)$$

$$j \in \Phi_b.$$

Here we assume that the  $Q(\cdot)$  function is a monotonically decreasing function of the argument. The lower bound can be achieved when all the  $\gamma_{bj}$ s are equal in the  $b$ th subband, in which case

$$\sigma_{sj}^2 = \sigma_b^2, \quad 0 \leq \sigma_b^2 \leq \frac{N}{N_c \sum_{t=1}^2 \sigma_{ct}^2}, \quad j \in \Phi_b. \quad (40)$$

Hence the training sequence should have the same power in all the subcarriers within a subband for the BER to achieve the lower bound within that subband. It may be noted that the above condition does not provide details about the power levels of the subcarriers of different subbands. They may be different from each other in different subbands. However, if the training sequences are allowed to have different power levels in the subcarriers of different subbands, then the channel MSEE is not minimized as the condition in (28) is not satisfied. This in turn adversely affects the BER. Hence for the joint minimization of the MSEE and the BER, we have,

$$\sigma_{sj}^2 = \left(\frac{N}{N'}\right) \sum_{t=1}^2 \sigma_{ct}^2. \quad (41)$$

In view of the above discussion, we see that (41) is very important in the sense that it gives us a criterion function which is useful for selecting the ST sequences

that minimizes the BER. This may be achieved by ensuring a fair distribution of the residual interference due to the training sequences on all the data subcarriers. Under a power constraint on the training sequence, this amounts to having equipowered training sequence components in all the data subcarriers in order to minimize the BER, in which case,  $\sigma_{sj}^2 = (N/N') \sum_{t=1}^{N_t} \sigma_{st}^2$ .

The condition in (41) is specifically applicable for ST-based MIMO-OFDM systems. Violation of the condition in (41) prevents the BER from achieving the lower bound. Hence (41) is a necessary condition for BER minimization over frequency selective fading channels. MIMO-OFDM over flat fading channels can be considered as a special case of the above since in this case,  $N_c = N$  and  $B_c = 1$ , and (41) is directly applicable.

From the above discussion and also from Section 3.1, we arrive at the training sequence selection criterion that jointly optimizes the MSEE and the BER of the ST-based MIMO-OFDM system. Training sequences satisfying this selection criterion requires to have an autocorrelation function that is as close as possible to an impulse for lags extending at least up to the order of the estimator, have negligible cross correlation between each other, and very importantly, ensure a fair distribution of the residual interference due to the training sequence on all the data subcarriers.

It may be noted that we have not followed a joint optimization procedure for minimizing the MSEE and the BER. Such an approach is difficult and may not be mathematically tractable. Instead, we provide an answer to the question, of all the sequences that minimize the MSEE, which sequence also minimizes the BER of the ST MIMO-OFDM system. Thus we arrive at the proposed selection criterion which is in effect jointly optimal. This criterion is useful and significantly different from existing proposals [11, 13, 17] and also results in an improved performance as seen later in the experimental studies in Section 5.

## 4. Chirp-Based Optimal Training Sequences

We propose the following set of chirp-based sequences one for each transmitting antenna as the optimal training sequences in terms of minimizing the MSEE and the BER of the OFDM system as discussed in Section 3.3. We compare the performance of the proposed training sequences with the other training sequences proposed in [11, 20, 25] in view of the proposed selection criterion.

**4.1. Proposed-Chirp Based Sequences.** The proposed set of chirp sequences is a periodic repetition given by

$$c_t(n) = c_t(n + N) = e^{j(2\pi n/N)(Q(t-1)+n/2+1)}, \quad t \in [1, N_t]. \quad (42)$$

Here it is assumed that  $N/Q$  is an integer. The  $N$  point DFT of the sequence is given by

$$c_t(k) = \sum_{n=0}^{N-1} e^{j(2\pi n/N)(Q(t-1)+n/2+1-k)}, \quad t \in [1, N_t], \quad (43)$$

whose magnitude at each subcarrier is given by

$$|c_t(k)| = \sqrt{N}, \quad k \in [0, N-1], \quad t \in [1, N_t]. \quad (44)$$

In view of (42) and (44), the proposed training sequences have the following advantages.

- (i) The energy of the sequences is equally distributed in all the data subcarriers and also they form an orthogonal set thus satisfying the proposed selection criterion in Section 3.3.
- (ii) The sequence has a constant magnitude in the time domain. Such sequences are known to be optimal in terms of the PAPR of the aggregate of the data and training signals as discussed in [20].

First, we compare the performance of the proposed chirp sequences  $c_t(n)$  with another set of chirp sequences given by

$$b_t(k) = e^{j2\pi n[Q(t-1)/N+((n+2)/2Q)]}, \quad t \in [1, N_t]. \quad (45)$$

Essentially, these sequences are similar to those proposed in [11, 20, 25] as the optimal training sequences. In [20], sequences derived from  $m$ -sequences of a similar type are also proposed. We use  $b_t(n)$  as a representative of these sequences for comparison. These are obtained by minimizing the MSEE alone. The magnitude of the  $N$  point DFT of  $b_t(n)$  can be expressed as

$$|b_t(k)| = \begin{cases} \frac{N}{\sqrt{Q}}, & k = iN_Q + t - 1, \quad i \in [0, Q-1], \\ 0, & \text{Otherwise.} \end{cases} \quad t \in [1, N_t], \quad (46)$$

From the above equation, since  $N \gg Q$ , there is a strong interference due to the training sequences in a few subcarriers.

Let us consider the same power for the training sequences (i.e., same  $\alpha$ ). A comparison of (44) and (46) shows that the sequence  $b_t(n)$  puts the entire interference in only  $Q$  number of subcarriers while in the proposed sequence  $c_t(n)$  it is distributed equally into all  $N$  subcarriers. This unequal distribution of power in the existing ST sequences is a major drawback as the interference severely degrades the BER performance. This is in contrast with the proposed training sequences  $c_t(n)$  which offers fairness through providing a uniform low interference in all the data subcarriers as seen in (44). Hence, as reasoned in Section 3.3, in the high SNR regime  $c_t(n)$  perform significantly better in terms of the BER.

When  $c_t(n)$  is compared to the  $m$ -length PN sequences with a period  $N$  denoted by  $p_t(n)$ ,  $t \in [1, N_t]$ , it is seen that  $c_t(n)$  has a more even distribution of the energy in all the subcarriers as compared to  $p_t(n)$ . Also the cross-correlation properties of the proposed chirp sequences  $c_t(n)$  is superior to the  $m$ -sequences. Thus the proposed chirp sequences are better suited according to the optimality conditions in Section 3.3.



**4.2. Impact of Guard Subcarriers.** In systems that employ guard subcarriers at the band edges, the use of  $b_t(n)$  may result in nonzero components in the locations of these guard subcarriers. Hence it would be difficult to incorporate such a sequence in currently standardized systems unlike  $c_t(n)$  which may be shaped appropriately with an associated loss in performance.

In systems that employ guard bands, the proposed training sequence is generalized as

$$c'_t(n) = c'_t(n + N) = \text{IDFT}_N\{\text{DFT}_N[c_t(n)] \times m(K)\}. \quad (47)$$

Here  $t \in [1, N_t]$  and  $m(K)$  is used as the spectral mask. If  $m(K) = 1$ , for all  $K \in [0, N - 1]$ ,  $\Rightarrow c'_t(n) = c_t(n)$ . In the IEEE 802.16e based OFDM system,  $m(K) = 1$  for  $K = (-100, \dots, -1, 1, \dots, 100)$  and  $m(K) = 0$  for  $K = (-128, \dots, -101, 0, 101, \dots, 127)$ .

The proposed training sequence  $c'_t(n)$  is now compared with the ST sequence  $b'_t(n)$ . These are a set of sequences obtained from the frequency shifted versions of the sequence  $b_1(n)$ , which is the superimposed training sequence obtained from the IDFT of the PRBS sequence that occupies the subcarrier indices as proposed in the IEEE 802.16e standard [35]. Hence the DFT of  $b'_1(n) = b'_1(n + N)$  has nonzero components at the indices  $(-88, -63, -38, -13, 13, 38, 88)$  as listed in the standard. This sequence is equipowered and equispaced over a subset of the subcarriers ranging from  $(-100$  to  $100)$ . We assume  $Q = 8$ . As discussed earlier both  $c'_t(n)$  and  $b'_t(n)$  do not satisfy (26).

Nevertheless, the MSEE due to the proposed training sequence  $c'_t(n)$  is superior to  $b'_t(n)$ . This is because the autocorrelation function of  $c'_t(n)$  falls more sharply as compared to that of  $b'_t(n)$  as seen in Figure 3. This is because  $c'_t(n)$  spans a larger range of frequencies as compared to  $b'_t(n)$ . As a result the off-diagonal decay of the  $\mathbf{R}_{cc}$  matrix as defined in Section 3.1 of  $c'_t(n)$  is more rapid than that of  $b'_t(n)$ . In addition to this  $c'_t(n)$  is also optimal in terms of minimizing the BER as discussed in Section 3.3. Hence in the presence of guard subcarriers, the improvement in performance obtained by employing the proposed optimal training sequence  $c'_t(n)$  is more significant than the case when the full band of the OFDM symbol is considered.

## 5. Experimental Results

In this work experimental studies are carried out to study the performance of the proposed scheme for an ST-based MIMO-OFDM system, and its comparison with existing schemes is presented. Representative results using the Alamouti space-time coding scheme with  $N_t = 2$  and  $N_r = 1$  are presented. The MSEE and the BER simulation results presented are obtained by averaging over an ensemble of 1000 Monte Carlo iterations. In order to benchmark against the uncoded system performance, channel coding is not considered here. In all the experiments unless otherwise mentioned we used the cyclic prefix of length 16.

**5.1. Performance over Fading Channels.** The impact of the initial channel estimation error on the OFDM system is

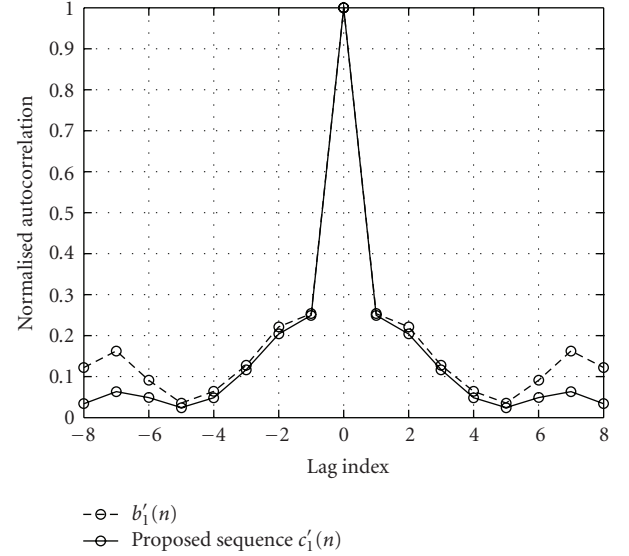


FIGURE 3: Comparison of the normalized autocorrelation of the proposed training sequence  $c'_1(n)$  with  $b'_1(n)$  for lags ranging from  $-8$  to  $8$  as discussed in Section 4.

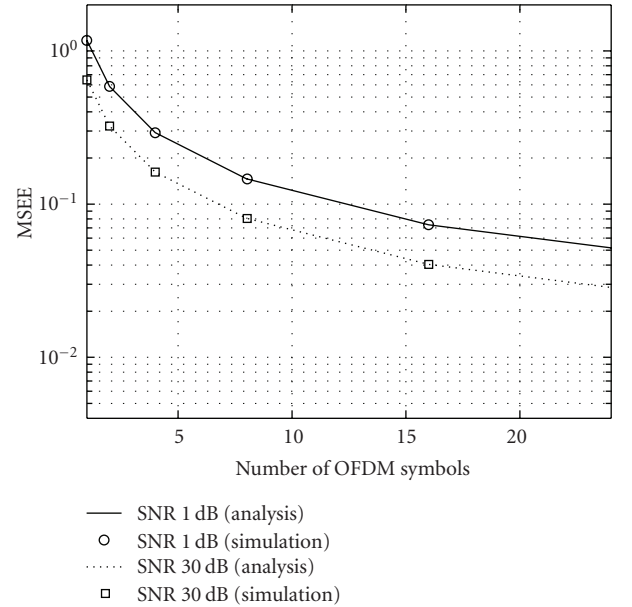


FIGURE 4: Theoretical and simulated MSEs of the initial channel estimate as a function of the number of OFDM symbols for SNR = 30 dB and 1 dB with  $Q = 8$ ,  $\alpha = 10$ ,  $N = 256$ .

studied by employing a time-varying frequency selective fading channel. The proposed chirp-based training sequence  $c_1(n)$  that jointly optimizes the MSEE and the BER as discussed in Section 4 is used. It can be seen that the MSEE reduces by a factor of  $1/(NT_p)$  as seen in Figure 4 and also in (27). The reduction in the MSEE is not that significant for a large number of OFDM symbols. This floor in the MSEE in the high SNR regime is seen as a result of the factor  $(QN_t\alpha)/(NT_p)$  in (27) which reduces only due to  $T_p$ . The data interference in the channel estimate affects the initial

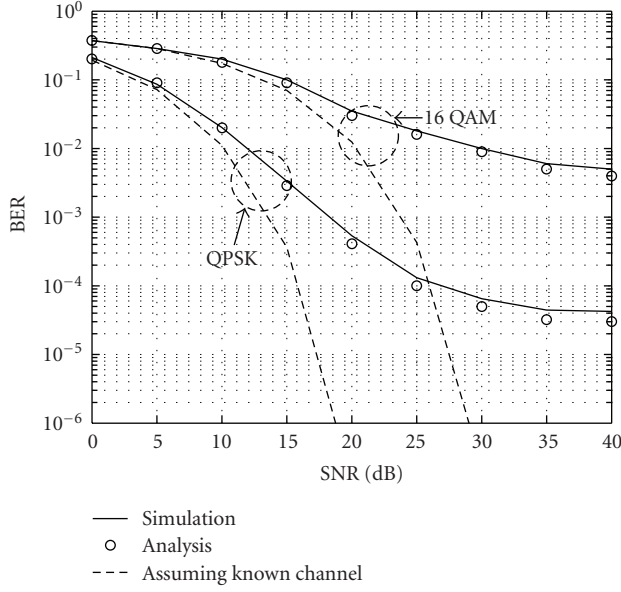


FIGURE 5: Theoretical and simulated BERs as a function of SNR for QPSK and 16-QAM with ST-based channel estimation over a frequency selective fading channel when compared with the corresponding bound by assuming that the channel is known at the receiver. Here  $N = 256$ ,  $T_p = 16$ ,  $Q = 8$ , and  $\alpha = 10$ .

channel MSEE. With the use of iterations as discussed in Section 2.5, the MSEE improves significantly for the SNRfs of practical interest.

Figure 5 illustrates the loss in BER performance due to the channel estimation error. The BER results were obtained by averaging over channels drawn from a uniform power delay profile with 4 taps. The influence of the channel estimation error becomes significant in the high SNR regime as compared to the known channel case, primarily due to the influence of the residual errors due to the training sequences on the data as seen in (32) and (33). The floor in the MSEE as observed in Figure 4 also contributes to the performance degradation.

Also, we see from Figure 5 that there is an excellent match between the analytical BER expression and the simulation result. This clearly indicates that the assumptions made in Section 3.2 regarding the negligible cross-correlation terms and the Gaussianity of the effective interference on the subcarriers are valid.

## 5.2. Optimal Training Sequences

**5.2.1. Performance Comparison.** In this simulation, a time-varying frequency selective fading channel model is used to compare the performance of the training sequences  $c_t(n)$  in (42) and  $b_t(n)$  in (45). Each channel realization is drawn from a zero-mean complex Gaussian process with the variance in each path normalized to follow an exponentially decaying power delay profile. Also,  $T_p$  is chosen so as to satisfy (10). As discussed in Section 4, the MSEs of these two sequences are the same but the BER performance in the

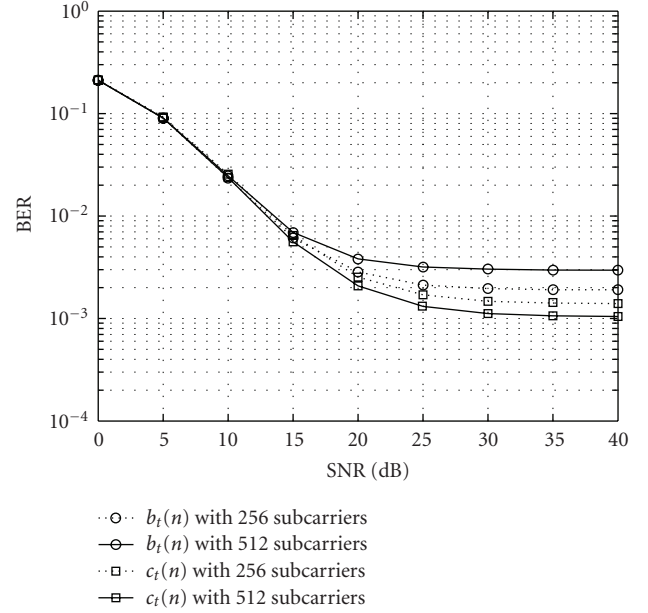


FIGURE 6: Simulated BER versus SNR comparison between the training sequences  $c_t(n)$  and  $b_t(n)$  with  $Q = 8$ ,  $\alpha = 10$ ,  $T_p = 8$  for  $N = 256$  and  $T_p = 4$  for  $N = 512$ , thus maintaining the same MSEE (27). The modulation format used is QPSK. The exponentially decaying power delay profile used was  $e^{-0.8l}$ ,  $l = \{0, 1, \dots, Q-1\}$ .

case of the proposed optimal training sequence is observed to be significantly better as seen in Figure 6. This result clearly strengthens the claim on essentiality of the BER as the optimality criterion and that the proposed training sequence  $c_t(n)$  yields that optimality in terms of both, the MSEE and the BER. In fact the difference in the BER is more pronounced for a larger  $N$ , maintaining the same MSEE as in the earlier case as seen in Figure 6. The same MSEE is maintained by keeping the product  $NT_p$  constant and keeping all the other parameters in the simulation constants as seen in (27). This is because the residual interference in the  $Q$  subcarriers increases by a factor of  $N_Q$  in  $b_t(n)$ , resulting in a higher error floor as compared to the case with a lesser number of subcarriers, thus increasing the gap in the BER between  $c_t(n)$  and  $b_t(n)$  in the high SNR regime. At an SNR of 30 dB, using 256 subcarriers the gap in the BER is approximately 1.5 dB. By using 512 subcarriers the gap increases to 4.5 dB at the same SNR.

**5.3. Application to the IEEE 16e OFDM-Based System.** In this work the proposed scheme is applied to this standard, its suitability is examined by comparing it with the comb type pilot-assisted scheme, and a case is made with the required design of the sequence.

We use the ITU Vehicular Channel A (ITUV-A, rms delay spread 0.364 micro seconds) [36]. For the simulation parameters considered in Table 1, the value of  $T_p = 8$  from (10), ensures that the fading channels experienced by these OFDM symbols are highly correlated with a correlation coefficient in excess of 0.95. Also the maximum mobility leading to  $f_{d\max}$  is 72 Km/Hr.

TABLE 1: Simulation parameters.

Symbol	Simulation parameters	Values
$f_c$	Center frequency	3 GHz
BW	Channel bandwidth	5 MHz
$f_s$	Signal sampling rate	5.76 MHz
$f_{d\max}$	Maximum Doppler frequency	200 Hz
CP	Cyclic prefix	16 samples
$Q$	Order of the estimator	8 samples
$T_p$	Number of OFDM symbols used for channel estimation	8 symbols
$\alpha$	Data to pilot power ratio	12

For the comb type pilot-assisted scheme, the estimator in (13) is used along with  $b'_i(n)$  as the training sequences resulting in 8 out of the 200 subcarriers for each transmitting antenna element being dedicated as pilot tones. In this case data transmission is avoided at these pilot tone subcarriers. This is compared with the proposed ST-based scheme using the proposed sequence  $c'_i(n)$  as shown in Figure 7. For a complexity similar to the pilot-assisted case, there is a considerable improvement in the bandwidth efficiency of the the ST-based scheme at the cost of system performance because of using all the 200 subcarriers for data transmission in both the transmitting antenna elements. However using only one additional iteration of the iterative data-aided channel estimation scheme as discussed in Section 2.5, the performance of the ST-based scheme becomes comparable to the pilot-assisted case as seen from Figure 7. We see that a BER of  $1 \times 10^{-3}$  is achieved at nearly the same SNR of 16 dB in the case of the comb type pilot-assisted scheme and the ST-based scheme with 2 iterations. Only an additional 1 dB is required for the ST-based scheme with 1 iteration for the same BER performance. Moreover forward error correcting (FEC) coding and decoding is a part of the standard. With its use, a BER of  $1 \times 10^{-6}$  can be easily achieved, thus satisfying the BER requirements listed in the standard.

**5.3.1. Performance Comparison with Guard Bands.** The performance of the proposed training sequence in the presence of guard bands  $c'_i(n)$  is now compared with  $b'_i(n)$  which is described in Section 4. The simulation parameters as in Table 1 were used but for a maximum Doppler of 140 Hz enabling the averaging of the estimate over  $T_p = 14$  OFDM symbols as discussed in Section 2.3. Figure 8 clearly demonstrates the improved performance obtained by employing the proposed training sequence  $c'_i(n)$  as compared to  $b'_i(n)$ . In the iterative scheme, the improvement in the initial channel estimate in the case of  $c'_i(n)$  improves the BER, which in turn further improves the MSEE in the next iteration. The BER performance obtained by using  $b'_i(n)$  with 4 iterations is still inferior to the BER performance obtained by employing the proposed sequence  $c'_i(n)$  with only 3 iterations. Hence for nearly the same BER performance, there is a considerable saving in the computational complexity when the proposed sequence set  $c'_i(n)$  that attempts to jointly

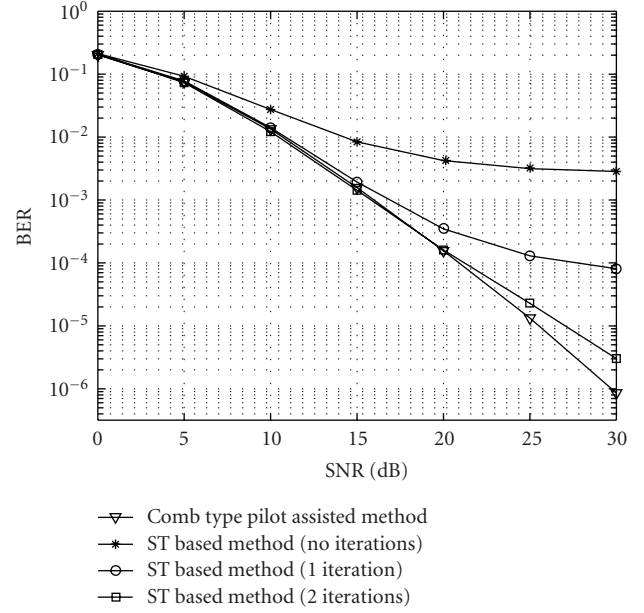


FIGURE 7: Simulated BER versus SNR comparison using the ITU Vehicular Channel Model A, with the simulation parameters of Table 1, between the comb type pilot-assisted scheme and the iterative data aided ST-based scheme as discussed in Section 2.5. Each data subcarrier uses QPSK modulation.

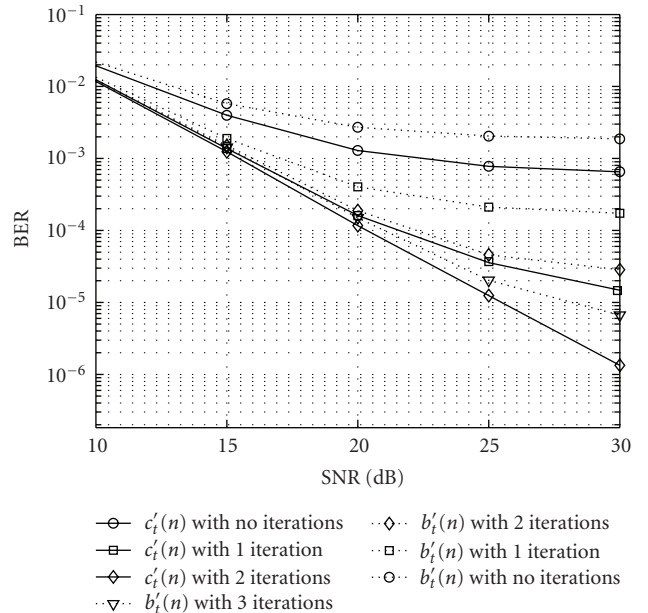


FIGURE 8: Simulated BER versus SNR comparison between the proposed sequence  $c'_i(n)$  and  $b'_i(n)$  for the ITU Vehicular Channel Model A with the simulation parameters of Table 1, except for  $f_{d\max} = 140$  Hz and  $T_p = 14$ . Each data subcarrier uses QPSK modulation.

minimize the MSEE and BER is used, as compared to the sequences that only minimizes the MSEE.

## 6. Conclusions

In this work, an iterative time-domain LS based channel estimation scheme for ST-based MIMO-OFDM systems is proposed. The method includes an interference canceler that reduces the interference due to the training symbols on the data symbols prior to data detection. The method is generalized to provide scope for exploiting the coherence bandwidth and the coherence time of the channel for an improved accuracy. It is shown that a major drawback of the existing schemes is the set of ST sequences arbitrarily chosen on the basis of minimizing the MSEE alone. This severely degrades the BER performance. The MSEE and the BER of the MIMO-OFDM system are analyzed, and a new selection criterion for the training sequences that jointly optimizes the MSEE and the BER is proposed. This yields a consistent selection criterion that requires the training sequences to be used at the transmitting antennas in such a way so as to form an orthogonal set and also ensure a fair distribution of the residual interference due to the training sequence on all the data subcarriers. It is shown that the proposed chirp-based sequences satisfy the selection criterion and the same are compared with the other existing proposals. ST-based methods are inherently bandwidth efficient. With the use of the proposed sequences along with the iterative channel estimation scheme, a BER performance which is comparable to pilot-assisted methods can be obtained in such ST-based systems, even in channels with a high Doppler bandwidth.

## References

- [1] Y. Li, N. Seshadri, and S. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 3, pp. 461–471, 1999.
- [2] Y. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Transactions on Wireless Communications*, vol. 1, no. 1, pp. 67–75, 2002.
- [3] I. Barhum, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels," *IEEE Transactions on Signal Processing*, vol. 51, no. 6, pp. 1615–1624, 2003.
- [4] F. Mazzenga, "Channel estimation and equalization for M-QAM transmission with a hidden pilot sequence," *IEEE Transactions on Broadcasting*, vol. 46, no. 2, pp. 170–176, 2000.
- [5] B. Farhang-Boroujeny, "Pilot-based channel identification: proposal for semi-blind identification of communication channels," *Electronics Letters*, vol. 31, no. 13, pp. 1044–1046, 1995.
- [6] D. C. McLernon, E. Alameda-Hernandez, A. Orozco-Lugo, and M. M. Lara, "New results for channel estimation via superimposed training," in *Proceedings of the 2nd International Symposium on Communications, Control and Signal Processing (ISCCSP '06)*, Marrakech, Morocco, March 2006.
- [7] J. K. Tugnait and X. Meng, "On superimposed training for channel estimation: performance analysis, training power allocation, and frame synchronization," *IEEE Transactions on Signal Processing*, vol. 54, no. 2, pp. 752–765, 2006.
- [8] J. K. Tugnait and W. Luo, "On channel estimation using superimposed training and first-order statistics," *IEEE Communications Letters*, vol. 7, no. 9, pp. 413–415, 2003.
- [9] J. K. Tugnait, X. Meng, and S. He, "Doubly selective channel estimation using superimposed training and exponential bases models," *Eurasip Journal on Applied Signal Processing*, vol. 2006, Article ID 85303, 11 pages, 2006.
- [10] X. Meng, J. K. Tugnait, and S. He, "Iterative joint channel estimation and data detection using superimposed training: algorithms and performance analysis," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 4, pp. 1873–1880, 2007.
- [11] A. G. Orozco-Lugo, M. Mauricio Lara, and D. C. McLernon, "Channel estimation using implicit training," *IEEE Transactions on Signal Processing*, vol. 52, no. 1, pp. 240–254, 2004.
- [12] G. T. Zhou, M. Viberg, and T. McKelvey, "A first-order statistical method for channel estimation," *IEEE Signal Processing Letters*, vol. 10, no. 3, pp. 57–60, 2003.
- [13] D. H. Pham and J. H. Manton, "An optimal pilot embedding scheme for channel identification and source recovery," in *Proceedings of IEEE International Symposium on Spread Spectrum Techniques and Applications (ISSSTA '04)*, pp. 77–81, Sydney, Australia, August–September 2004.
- [14] Q. Yang and K. S. Kwak, "Time-varying multipath channel estimation with superimposed training in CP-OFDM systems," *ETRI Journal*, vol. 28, no. 5, pp. 688–691, 2006.
- [15] C.-T. Lam, D. D. Falconer, and F. Danilo-Lemoine, "Iterative frequency domain channel estimation for DFT-precoded OFDM systems using in-band pilots," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 2, pp. 348–358, 2008.
- [16] K. Josiam and D. Rajan, "Bandwidth efficient channel estimation using super-imposed pilots in OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 6, no. 6, pp. 2234–2245, 2007.
- [17] T. Cui and C. Tellambura, "Superimposed pilot symbols for channel estimation in OFDM systems," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '05)*, vol. 4, pp. 2229–2233, St. Louis, Mo, USA, November–December 2005.
- [18] A. Vosoughi and A. Scaglione, "On the effect of receiver estimation error upon channel mutual information," *IEEE Transactions on Signal Processing*, vol. 54, no. 2, pp. 459–472, 2006.
- [19] A. Vosoughi and A. Scaglione, "Everything you always wanted to know about training: guidelines derived using the affine precoding framework and the CRB," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 940–954, 2006.
- [20] N. Chen and G. T. Zhou, "Superimposed training for OFDM: a peak-to-average power ratio analysis," *IEEE Transactions on Signal Processing*, vol. 54, no. 6, pp. 2277–2287, 2006.
- [21] N. Ohkubo and T. Ohtsuki, "Added pilot semi-blind iterative channel estimation for OFDM packet transmission," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '03)*, vol. 2, pp. 878–882, San Francisco, Calif, USA, December 2003.
- [22] C. K. Ho, B. Farhang-Boroujeny, and F. Chin, "Added pilot semi-blind channel estimation scheme for OFDM in fading channels," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '01)*, vol. 5, pp. 3075–3079, San Antonio, Tex, USA, November 2001.
- [23] S. Balasubramanian, B. Farhang-Boroujeny, and V. J. Mathews, "Pilot embedding for channel estimation and tracking in OFDM systems," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '04)*, vol. 2, pp. 1244–1248, Dallas, Tex, USA, November–December 2004.
- [24] H. Zhu, B. Farhang-Boroujeny, and C. Schlegel, "Pilot embedding for joint channel estimation and data detection in MIMO communication systems," *IEEE Communications Letters*, vol. 7, no. 1, pp. 30–32, 2003.



- [25] S. He, J. K. Tugnait, and X. Meng, "On superimposed training for MIMO channel estimation and symbol detection," *IEEE Transactions on Signal Processing*, vol. 55, no. 6 II, pp. 3007–3021, 2007.
- [26] A. Scaglione and A. Vosoughi, "Turbo estimation of channel and symbols in precoded MIMO systems," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '04)*, vol. 4, pp. 413–416, Montreal, Canada, May 2004.
- [27] N. N. Tran, H. D. Tuan, and H. H. Nguyen, "Superimposed training designs for spatially correlated MIMO-OFDM systems," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '08)*, pp. 3549–3554, New Orleans, La, USA, November-December 2008.
- [28] M.-A. Khalighi and S. Bourennane, "Semi-blind channel estimation based on superimposed pilots for single-carrier MIMO systems," in *Proceedings of the 65th IEEE Vehicular Technology Conference (VTC '07)*, pp. 1480–1484, Dublin, Ireland, April 2007.
- [29] Q. Yang and K. S. Kwak, "Superimposed pilot aided multiuser channel estimation for MIMO-OFDM uplinks," *ETRI Journal*, vol. 28, no. 5, pp. 688–691, 2006.
- [30] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications: where Fourier meets Shannon," *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp. 29–48, 2000.
- [31] E. Alameda-Hernandez, D. C. McLernon, A. G. Orozco-Lugo, M. M. Lara, and M. Ghogho, "Frame/training sequence synchronization and DC-offset removal for (data-dependent) superimposed training based channel estimation," *IEEE Transactions on Signal Processing*, vol. 55, no. 6 I, pp. 2557–2569, 2007.
- [32] C.-P. Li and W.-W. Hu, "Super-imposed training scheme for timing and frequency synchronization in OFDM systems," *IEEE Transactions on Broadcasting*, vol. 53, no. 2, pp. 574–583, 2007.
- [33] Y. Zhang and H.-H. Chen, *Mobile WiMAX—Toward Broadband Wireless Metropolitan Area Networks*, Auerbach Publications, Boca Raton, Fla, USA, 2008.
- [34] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.
- [35] "Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems," IEEE 802.16e-2005, February 2006.
- [36] "Mobile WiMAX—Part I: A Technical Overview and Performance Evaluation," WiMAX Forum, May 2006.