

# Bandwidth Efficient Channel Estimation Using Super-Imposed Pilots in OFDM Systems

Kaushik Josiam, *Student Member, IEEE*, and Dinesh Rajan, *Member, IEEE*

**Abstract**— This paper proposes a channel estimation algorithm using super-imposed pilots for OFDM systems. The pilot symbols are added linearly to the modulated data symbol at a fraction of the total transmit power, making the method spectrally efficient. We present a series of iterative and non-iterative receivers in which complexity is traded for performance. In particular, a high complexity non-iterative receiver that achieves the theoretically lowest BER and a low complexity iterative receiver that show near-optimal performance are presented. The proposed iterative receiver offers a 50 - 98% reduction in complexity over a non-iterative receiver having equivalent performance, depending on the modulation used. We observe that the performance of the receiver is independent of the channel's Doppler frequency and delay spread. Finally, we compute an analytic bound for the bit error rate of the proposed non-iterative receiver.

**Index Terms**— Iterative methods, maximum likelihood estimation, multipath channels, orthogonal frequency-division multiplexing (OFDM), wireless LAN.

## I. INTRODUCTION

ORTHOGONAL Frequency Division Multiplexing (OFDM) [1] is a multi-carrier modulation technique that comes with an inherent promise of spectrally efficient high data rate transmission and multi-path effect mitigation in wideband wireless systems. In an OFDM system, the available bandwidth is split between the carriers and the bandwidth allocated to each sub-carrier is narrow enough for the channel to be considered flat fading over it. Consequently, OFDM has been selected as the basis for many different high speed wireless standards, e.g., local area network (W-LAN) systems (for example, IEEE 802.11, HIPERLAN2) and for digital audio and video broadcasting (DAB and DVB) in Europe.

For coherent demodulation in an OFDM receiver, the knowledge of each sub-carrier's gain and phase is critical. Often in practice, training symbols known as pilots that are known to receiver a priori, are initially transmitted for channel estimation. Further, pilot symbols are interspersed with data symbols to track the time-varying channel [2]. For example, in 802.11 networks, along with the pilot symbols in the preamble for channel estimation, a few designated sub-carriers are reserved for transmitting pilots to enable channel tracking. However, the addition of pilots cause bandwidth expansion and

reduce the data rate. For spectrally efficient transmission, blind channel estimation methods that do not use pilot symbols have also been proposed. Existing blind methods primarily use the statistics of the received signal to estimate the channel [4]–[6]. Although blind methods have high spectral efficiency, they converge slower and have high computational complexity.

This paper focuses on a bandwidth efficient pilot arrangement for channel estimation in OFDM systems. The pilots are linearly added on the data and hence called super-imposed pilots. The concept of super-imposed pilots for simultaneous data transmission and channel estimation was first proposed for analog communication systems [7] and extended to digital communication systems in [8]. Subsequently, many flavors of systems with super-imposed pilots have been proposed. In [10]–[13], the super-imposed pilots are periodic and the resulting cyclostationarity of the received symbols is exploited to estimate the channel using first order statistics. The channel estimation algorithm relied on time averaging over many symbols, thus requiring that the channel be slow varying over a large block of symbols. Hence, such statistical algorithms are limited in the range of channel conditions over which they can be used. Pilot assisted transmission (PAT) for channel estimation has been described in [3], a specialization of which is the super-imposed pilots. The proposed channel estimation method uses a Kalman filter to estimate the channel and the quality of the channel estimate was the only performance criterion. In contrast to the statistical methods, [9] proposed an maximum likelihood (ML) based receiver for symbols with super-imposed pilots and any arbitrary pilot sequence. Prediction was used to update the channel estimate at each time instant. The primary drawback of such a receiver is the huge computational complexity needed to extract the ML sequence.

In this paper, we propose a joint channel estimation and symbol detection algorithm for OFDM systems with super-imposed pilots using a soft-output Viterbi detector. At each subcarrier, the soft-output Viterbi detector, accepts inputs with super-imposed pilots, continually estimates the channel, and outputs soft value of the detected symbol, subject to a minimum squared error constraint between the transmitted and received symbols with super-imposed pilots. Thus effectively, the soft-output Viterbi detector replaces the channel estimation, channel equalization, symbol detection and demodulation blocks in a conventional OFDM receiver. Further, taking advantage of the channel coding in OFDM transceivers, we propose iterative receivers with soft data exchange between the Viterbi detector and channel decoder. The proposed non-iterative and iterative modes of receiver operation, allow

Manuscript received October 21, 2005; revised July 11, 2006 and July 19, 2006; accepted September 20, 2006. The associate editor coordinating the review of this paper and approving it for publication was K.-K. Wong. This work was supported in part by the Nokia Corporation. Some parts of this work appeared in [36].

The authors are with the Dept of Electrical Engineering, Southern Methodist University, 6251 Airline Drive, Dallas, TX 75275 USA (e-mail: {kjosiam@engr.smu.edu, rajand}@engr.smu.edu).

Digital Object Identifier 10.1109/TWC.2007.05787.

TABLE I  
 SYMBOLS AND NOTATIONS

Symbol	Used for
$[\cdot]_n$	sub-carrier index
$[\cdot]_t$	time index, stage of the trellis
$\hat{a}$	Hypothesis of symbol $a$
$\hat{a}$	Estimate/decision of symbols $a$
$\mathbf{a}_k$	Vector $\mathbf{a}$ of dimension $k$
$\mathbf{A}^{(m,n)}$	Matrix $\mathbf{A}$ of dimension $m \times n$
$g(\cdot)$	Probability density function
$\sigma_x^2$	Variance of a random variable $x$
$\rho$	pilot power fraction
$T_c$	Coherence Time
$N$	FFT Size
$K$	Length of the channel impulse response
$L$	averaging order
$r$	received symbols
$u$	data symbols (interleaved/input to the detector)
$x$	detected coded symbols (de-interleaved)
$p$	pilot symbols
$z$	AWGN symbols
$f$	flat fading coefficient
$h$	Channel filter coefficient $\mathbf{f}_N \rightarrow \mathcal{F}[\mathbf{h}_K]$
$F_s$	Channel Bandwidth, sampling frequency
$\tau_d$	RMS delay spread
$s, \hat{s}$	current and previous state in the trellis
$\gamma_t(\hat{s}, s)$	Transition probability at time $t$ between states $\hat{s}$ and $s$
$\mathcal{M}(\cdot)$	computed metric information
$\mathcal{L}_a(\cdot)$	a priori log-likelihood ratio of symbols
$\mathcal{L}_e(\cdot)$	extrinsic log-likelihood ratio of symbols

for graceful trade-off between performance (e.g., BER) and computational complexity. Specifically, we present results for

- A non-iterative receiver that achieves theoretically lowest BER but has high complexity.
- A low complexity iterative receiver with near-optimal performance.

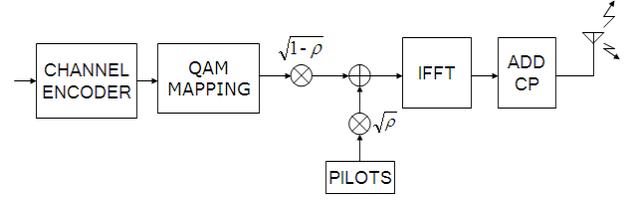
The performance of the proposed receivers is assessed using simulations in an OFDM environment with BPSK and QPSK modulations. It is observed that the proposed receivers' performance is agnostic to the block fading channel's coherence time and the root mean square (RMS) delay spread, as long as there is no inter-symbol interference (ISI). The iterative receiver is shown to perform on par with the near-optimal non-iterative receiver, but with almost 50% and 98% savings in the number of arithmetic operations for BPSK and QPSK modulations respectively. Hence the required computational complexity is substantially reduced. We also derive an analytical bound for the bit error rate (BER) of the proposed receiver with QPSK modulation.

The rest of the paper is organized as follows. The system description along with the channel model used is given in Section II. The channel estimation algorithm along with the receiver design is described in Section III. An iterative receiver that uses the turbo-detection principle is presented in Section IV. Numerical results are given in Section V followed by conclusions in Section VII.

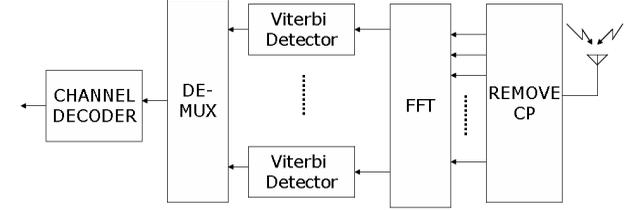
## II. SYSTEM DESCRIPTION

### A. Transmission Strategy

The block diagram detailing the data processing in an OFDM system is shown in Fig. 1(a). At the transmitter,



(a) OFDM Transmitter



(b) OFDM Receiver

Fig. 1. Block diagram of an OFDM system.

information bearing symbols are modulated using a typical regular constellation like QPSK or 16-QAM. A known complex pseudo-noise (PN) pilot sequence,  $p_{n,t}$ , is added to the unknown data symbols,  $u_{n,t}$ , where  $n$  is the subcarrier index and  $t$  is the time index. The data and pilot sequence belong to a  $M$ -ary alphabet depending on the modulation used. However, the pilot sequences are scaled differently from the data symbols. Let  $\rho$  denote the fraction of the transmit energy  $E_s$  allocated to the pilot symbols. Then, the average energy of the pilot symbols is  $E_p = \mathbb{E}[|p_{n,t}|^2] = \rho \cdot E_s$  and the average data energy is  $E_u = \mathbb{E}[|u_{n,t}|^2] = (1 - \rho) \cdot E_s$ . The resulting sums of data and pilot symbols are referred to as data-pilot (DP) symbols, illustrated in Fig. 2. The data and pilot symbols are assumed to have the following statistical properties:

- The data and pilot sequences are each i.i.d and zero mean i.e.,  $\mathbb{E}[u_{n,t}] = \mathbb{E}[p_{n,t}] = 0$  and
 
$$\mathbb{E}[u_{n,t}u_{m,\tau}^*] = \begin{cases} (1 - \rho)E_s & n = m, t = \tau \\ 0 & \text{otherwise} \end{cases} \quad \text{and}$$

$$\mathbb{E}[p_{n,t}p_{m,\tau}^*] = \begin{cases} \rho E_s & n = m, t = \tau \\ 0 & \text{otherwise} \end{cases}$$
- The data and pilot sequences are uncorrelated i.e.,  $\mathbb{E}[u_{n,t}p_{m,\tau}^*] = 0$ .

The DP symbols are segmented into blocks of length  $N$ . Each block of DP symbols is then transformed using inverse discrete Fourier transform (IDFT) to generate a time domain symbol block. The last  $N_{cp}$  symbols in the block containing time-domain symbols are repeated at its beginning forming the cyclic prefix (CP). The symbols are then sequentially transmitted over a frequency selective channel.

### B. Channel Model

The wireless channel is modeled as a standard block fading channel. In each block, the channel is represented using a linear time-invariant, finite impulse response filter. The length of the channel filter,  $K$ , depends on the channel bandwidth,  $F_s$ , and the root mean square (RMS) delay spread,  $\tau_d$ , of

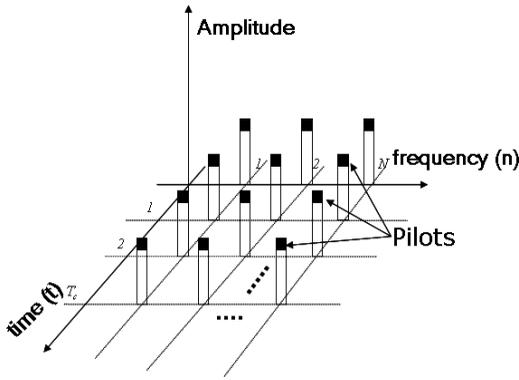


Fig. 2. Pilots super-imposed over data at all temporal and spectral locations.

the channel. The coefficients of the filter,  $h_k$ , have a zero mean normal distribution with variances  $\sigma_k^2$  that follow an exponential power delay profile, *i.e.*,

$$h_k \sim \mathcal{N}(0, \sigma_k^2) \quad k = 0, \dots, K-1 \quad (1)$$

where,

$$\sigma_k^2 = \sigma_0^2 (e^{-\frac{k}{F_s \tau_d}}) \quad k = 0, \dots, K-1 \quad (2)$$

and  $\sigma_0^2 = 1 - e^{-\frac{1}{F_s \tau_d}}$ . Typically,  $K$  is chosen as  $\lceil 10F_s \tau_d \rceil$  [14]. The channel thus generated remains fixed for the duration of the coherence time,  $T_c$  symbols, and changes independently after  $T_c$  symbols.

### III. JOINT CHANNEL ESTIMATION AND DATA DETECTION

The FFT used in OFDM receivers, Fig. 1(b), transforms the frequency selective channel into a multiplicative scaling of the transmitted symbols at each sub-carrier by a single tap channel. Accordingly, an  $N$ -length FFT transforms a  $K$ -tap filter given by coefficients  $\mathbf{h}_K$ , into  $N$  single tap flat fading coefficients given by  $\mathbf{f}_N$ . The orthonormality of the FFT ensures that each sub-carrier can be treated independently, and the block fading nature of the channel results in  $T_c$  symbols in a sub-carrier sharing the same fading co-efficient. Consequently, the received signal  $r_{n,t}$ , after the FFT block in Fig. 1(b), at the  $n^{\text{th}}$  sub-carrier and time instant  $t$ , is given by

$$r_{n,t} = (u_{n,t} + p_{n,t})f_n + z_{n,t} \quad t = 1, 2, \dots, T_c \quad (3)$$

where,  $f_n \in \mathbb{C}$  is the multiplicative fading coefficient of the channel and  $z_{n,t} \in \mathbb{C}$  is the zero mean AWGN with power  $\sigma_z^2$ . Our objective in this paper, is to design a receiver that jointly estimates the channel and detects data assuming perfect time and frequency synchronization.

#### A. Optimal Channel Estimate for Block Fading Channel

We now compute the receiver that maximizes the likelihood of the data symbols. The data symbols can be hypothesized from a set whose elements are defined by the modulation order  $M$ . For a specific  $M$ , it is easy to enumerate all possible data symbol hypotheses for coherence time  $T_c$ . Accordingly, for a sub-carrier  $n$ , there are  $M^{T_c}$  possible,  $T_c$ -dimensional hypothesis vectors,  $[\tilde{u}_{n,T_c}, \tilde{u}_{n,T_c-1}, \dots, \tilde{u}_{n,1}]$ . Since the noise vector  $\mathbf{z}_n$  is complex Gaussian with zero mean and variance  $\sigma_z^2$ , the likelihood function that a given data hypotheses

$\tilde{u}_{n,T_c}, \tilde{u}_{n,T_c-1}, \dots, \tilde{u}_{n,1}$  was transmitted on the  $n^{\text{th}}$  sub-carrier and that the channel estimate is  $\tilde{f}_n$ , can be written as,

$$g_z(\mathbf{r}_n | \tilde{\mathbf{u}}_n, \tilde{f}_n) = \left( \frac{1}{\sqrt{2\pi\sigma_z^2}} \right)^{T_c} \prod_{i=1}^{T_c} e^{-\frac{|r_{n,i} - (\tilde{u}_{n,i} + p_{n,i})\tilde{f}_n|^2}{2\sigma_z^2}} \quad (4)$$

Maximizing the likelihood function is equivalent to minimizing the sum of the exponent terms in (4). Coherent detection of data symbols involves choosing a hypothesis that minimizes the squared norm in the exponent, given by,

$$\Lambda(\tilde{u}_{n,T_c}, \tilde{u}_{n,T_c-1}, \dots, \tilde{u}_{n,1}) = \sum_{i=1}^{T_c} |r_{n,i} - (\tilde{u}_{n,i} + p_{n,i})\tilde{f}_n|^2 \quad (5)$$

Given a set of data hypotheses,  $\tilde{u}_{n,T_c}, \tilde{u}_{n,T_c-1}, \dots, \tilde{u}_{n,1}$ , the squared error minimizing channel hypothesis  $\tilde{f}_n$ , is found by equating  $\frac{\partial \Lambda}{\partial \tilde{f}_n} = 0$ , as,

$$\tilde{f}_n(\tilde{u}_{n,T_c}, \tilde{u}_{n,T_c-1}, \dots, \tilde{u}_{n,1}) = \frac{\sum_{i=1}^{T_c} r_{n,i}(\tilde{u}_{n,i} + p_{n,i})^*}{\sum_{i=1}^{T_c} |\tilde{u}_{n,i} + p_{n,i}|^2} \quad (6)$$

The receiver has to exhaustively search among all  $M^{T_c}$  possible hypotheses to find the maximum likelihood (ML) data sequence that minimizes the squared norm in (5). The computational complexity of the algorithm, measured in terms of the number of arithmetic operations (multiplications and additions) needed for channel estimation and symbol detection is  $O(T_c N M^{T_c})$ . In an OFDM system, the number of sub-carriers  $N$  is fixed. For a fixed  $N$ , with increasing modulation order  $M$ , or coherence time,  $T_c$ , the increase in complexity is exponential. The exponential complexity makes it challenging to implement the algorithm in high data rate environments where higher order modulations are used. We now modify the ML algorithm using the Viterbi detector which is known to efficiently search for the ML sequence.

#### B. Viterbi Detector

The received vector at sub-carrier  $n$ ,  $[r_{n,1}, r_{n,2}, \dots, r_{n,T_c}]$ , is processed over a small sliding time window of length  $L$  using the Viterbi algorithm [15] on a suitable trellis. The traditional Viterbi detector, is modified to accept a priori soft information values about the received signal and also to output the soft information values according to the soft output Viterbi algorithm described in [19], while implicitly estimating and equalizing the channel. The length of the sliding window,  $L$ , and modulation order,  $M$ , define the trellis completely as shown in Fig. 3. There are  $M^L$  states corresponding to the number of possible data hypotheses over a window of length  $L$ . There are  $M$  branches coming into and out of every state in the trellis. The channel hypothesis at stage  $t$  in the trellis, is now averaged over  $L$  data hypotheses associated with a particular state  $s$  of the trellis and the data hypothesis  $\tilde{u}_{n,t}$  of the current transition from that state and terminating in state  $s$  of the next stage. The sliding window is used to average the channel hypothesis over  $L+1$  symbols, and hence, is also

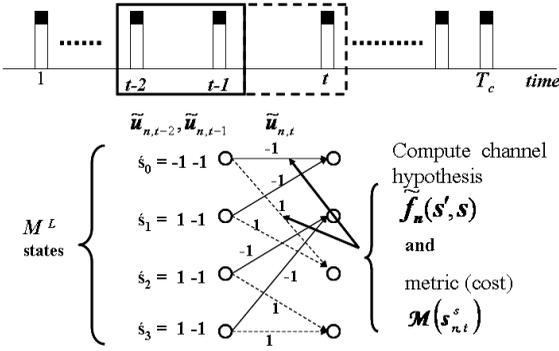


Fig. 3. Trellis showing the states and transitions for a modulation order  $M = 2$  (BPSK) and averaging order  $L = 2$ .  $\tilde{f}_n(\acute{s}, s)$  from (7) and  $\mathcal{M}(\mathbf{s}_{n,t}^s)$  (8) are the computed channel hypothesis and transition metrics for each transition in the trellis.

called *averaging order*. Accordingly the channel hypothesis between states  $\tilde{f}_n(\acute{s}, s)$  is given by,

$$\tilde{f}_n(\acute{s}, s) = \frac{\sum_{l=0}^L r_{n,t-l}(\tilde{u}_{n,t-l} + p_{n,t-l})^*}{\sum_{l=0}^L |\tilde{u}_{n,t-l} + p_{n,t-l}|^2} \quad (7)$$

Let  $\mathbf{s}_t^s$  be the vector of surviving states up to state  $S_t = s$  at stage  $t$  in the trellis. The recursive relation for computing the metric for the transition between state  $S_{t-1} = \acute{s}$  at stage  $t-1$  and state  $S_t = s$  at stage  $t$ , derived in Appendix I, is given by,

$$\begin{aligned} \mathcal{M}(\mathbf{s}_{n,t}^s) &= \mathcal{M}(\mathbf{s}_{n,t-1}^{\acute{s}}) - \frac{1}{2\sigma_z^2} |r_{n,t} - \tilde{f}_n(\acute{s}, s)(\tilde{u}_{n,t} + p_{n,t})|^2 \\ &+ \frac{1}{2} \sum_{i=1}^{\log_2 M} \tilde{b}_{i,n,t} \mathcal{L}_a(b_{i,n,t}) \end{aligned} \quad (8)$$

where  $\mathcal{M}(\mathbf{s}_{n,t-1}^{\acute{s}})$  is the cumulative metric of the surviving path up to state  $S_{n,t-1} = \acute{s}$ ,  $\tilde{b}_{i,n,t}$  is the BPSK symbol value (+1 if bit is 1 and -1 if bit is zero) of the  $i^{\text{th}}$  bit in the symbol  $\tilde{u}_{n,t}$ ,  $\mathcal{L}_a(b_{i,n,t})$  is the a priori soft information about the  $i^{\text{th}}$  bit in the data symbol at stage  $t$  and sub-carrier  $n$  and  $\tilde{f}_n(\acute{s}, s)$  is the channel hypothesis computed using (7).

The soft outputs are calculated by computing the difference between the two highest metrics from among the  $M$  merging paths at any state in the trellis. Let  $\mathbf{s}_{n,t}^s$  and  $\hat{\mathbf{s}}_{n,t}^s$  be the two paths with highest metrics at state  $S_{n,t} = s$  with metrics  $\mathcal{M}(\mathbf{s}_{n,t}^s)$  and  $\mathcal{M}(\hat{\mathbf{s}}_{n,t}^s)$  respectively. The path difference between the metrics is computed as,

$$\Delta_{n,t}^s = \mathcal{M}(\mathbf{s}_{n,t}^s) - \mathcal{M}(\hat{\mathbf{s}}_{n,t}^s) \geq 0 \quad (9)$$

and  $\mathbf{s}_{n,t}^s$  is chosen to be the surviving state since its metric is higher. It has been shown (equation (35) in [18]), that  $\Delta_{n,t}^s$  represents the log likelihood ratio (LLR) for the correct decision at stage  $t$ .

Typically, with grey coded modulation, the LLRs of only those bits whose value would change across the decision boundary are calculated as [20],

$$\mathcal{L}_a(b_{j,n,t}/\mathbf{r}) = b_{j,n,t} \min_{i=k, \dots, k+\delta, b_{j,n,t} \neq b_{j,n,t}^i} \Delta_i^s \quad (10)$$

$$j = 1, \dots, \log_2(M)$$

where  $b_{j,n,t}$  is the BPSK symbol value of the bit whose value changes across the decision boundary,  $b_{j,n,t}^i$  is the value of similar bits for the path which merged with the ML path but was discarded at stage  $t$  and  $\delta$  is a free parameter typically set to five times the value of the averaging order  $L$ . Since the output LLRs of the detector are used as a priori values for the channel decoder, we duplicate the magnitude of the LLR computed in (10) to all the bits in the symbol scaled by their corresponding BPSK symbol value ( $\pm 1$ ). Thus, the magnitude of the LLRs for both bits in a QPSK symbol are identical, with the sign changing according to the detected symbol.

Hard decisions on the transmitted bit are given by a simple sign detection of the output LLRs and an explicit estimate of the channel may be easily computed using the remodulated hard decisions, denoted by,  $[\hat{u}_{n,1}, \hat{u}_{n,2}, \dots, \hat{u}_{n,T_c}]$ , as,

$$\hat{f}_n(\hat{u}_{n,1}, \hat{u}_{n,2}, \dots, \hat{u}_{n,T_c}) = \frac{\sum_{i=1}^{T_c} r_{n,i}(\hat{u}_{n,i} + p_{n,i})^*}{\sum_{i=1}^{T_c} |\hat{u}_{n,i} + p_{n,i}|^2} \quad (11)$$

The Viterbi detector described above, is designed to process the received DP symbols for channel estimation, without any need for explicit pilot symbols to estimate the channel. Such a blind receiver is found to have a flat bit error rate (BER) response with increasing signal to noise ratio (SNR). The problem is attributed to the lack of absolute phase reference for the receiver and has been discussed in [21]. We use a solution, similar to [21], where a single pilot symbol ( $\rho = 1$ ), is initially transmitted to the receiver. The channel estimate from the pilot symbol is used in the initial  $L+1$  stages of the trellis, and the estimator of (7) is used in the stages following the  $(L+1)^{\text{st}}$  stage. The received sequence of  $T_c$  symbols, now consists of one pilot symbol in the beginning, followed by  $T_c - 1$  symbols with super-imposed pilots. In contrast to the single pilot proposed here, conventional training schemes transmit more than one pilot symbol in the preamble of the data packet.

### C. Detector Complexity

The complexity of the Viterbi detector is defined by the trellis size. For an averaging order,  $L$ , there are  $M^L$  states in the trellis. The number of states in the detector trellis grows exponentially with  $L$  and  $M$ . The number of arithmetic operations required for channel estimation and symbol detection across all  $N$  sub-carriers and for a coherence time of  $T_c$  OFDM symbols is  $O(T_c N M^L)$ . The averaging order,  $L$ , is a design constraint, which can be chosen to be less than  $T_c$ . In fact, we show using numerical simulations, that for BPSK modulation, an averaging order,  $L = 5$ , achieves near optimal ML performance for any coherence time  $T_c > 5$  symbols. Clearly, for a fixed  $N$  and  $T_c$ , the complexity of the Viterbi detector ( $M^L$ ) is significantly lower than the ML receiver ( $M^{T_c}$ ). To further reduce complexity, we propose a solution from the realm of iterative detection [23].

Turbo or iterative receivers exchange soft information between the concatenated components as in the OFDM receiver and have been an area of significant research. The constructive

exchange of the soft values between the detector and the channel decoder substantially improves, with every iteration, the overall system performance. Further, we show using simulations, that for BPSK modulation, with just two iterations, an iterative detector with averaging order,  $L = 2$  outperforms a non-iterative detector using  $L = 5$  using substantially lower computations.

#### IV. ITERATIVE RECEIVER FOR CODED - OFDM

Current state of the art non-iterative processing of the received symbols sees the soft outputs from the detector being forwarded to the channel decoder via an interleaver. The output LLRs of the channel decoder is used to make decisions on the data symbols. Iterative receivers loop this processing by forwarding the LLRs from the decoder as  $\hat{a}$  priori information about the received symbol to the detector. Literature in turbo detection [23]–[29] have shown that such looping substantially improved performance over conventional non-iterative receivers *i.e.*, with a degradation of only 0.6dB from the AWGN channel performance [23]. The performance enhancement comes at the cost of increased latency of the output symbols at the receiver. But, with the emergence of improved algorithms for detection and decoding [30]–[34], their usage in practical systems has gained increasing acceptance. Thus, as a logical extension to the proposed receiver in Section III, we propose an iterative receiver using superimposed pilots as shown in Fig. 4.

The proposed OFDM receiver consists of a parallel bank of  $N$  Viterbi detectors, corresponding to each of the  $N$  subcarriers, and one decoder, as illustrated in Fig. 4. The mapping of the values from the detector to the decoder and viceversa is done by simple de-multiplexing and multiplexing operations respectively. The conventional turbo processing of the soft information exchange between the detector and decoder forms Loop I in the system. The initial channel estimate,  $\hat{f}_{n,ini}$  for the first iteration is estimated using the single pilot symbol transmitted, where  $n$  is the sub-carrier index. For every successive iteration, we use the knowledge of the decoded symbols to estimate an updated channel estimate,  $\hat{f}_n$ . To calculate the updated channel estimate, channel estimator block forms Loop II in the system. In the channel estimator, the decoded symbols are re-modulated and multiplexed into their respective sub-carriers. Using the expression in (11), an updated channel estimate is computed, which is then used in the Viterbi detector.

Loop I in the iterative receiver is the conventional turbo receiver where soft information is exchanged between the decoder and detector takes place. For a sub-carrier  $n$ , the Viterbi detector described in the preceding section, takes the following inputs - the received symbols,  $r_{n,t}$ , the channel estimate vector  $\hat{f}_n$  and the  $\hat{a}$  priori soft values,  $\mathcal{L}_e^D(u_{n,t})$ . The detector outputs the soft values of the detected symbols  $\mathcal{L}_a(u_{n,t})$ . In addition to the extrinsic information from the detector, an estimate of the data symbols computed as,  $\hat{u}_{n,t} = \tanh\left(\frac{\mathcal{L}_a(u_{n,t})}{2}\right)$ , is also computed. The soft value and the estimate of data symbol from all sub-carriers are then de-multiplexed in a single stream whose  $i^{th}$  value is described by  $\mathcal{L}_a(u_i)$  and  $\hat{u}_i$  respectively. The extrinsic information for the  $i^{th}$  symbol,  $\mathcal{L}_e(u_i)$ , can be

calculated from the soft values of the detected symbols, as shown

$$\mathcal{L}_e(u_i) = \mathcal{L}_a(u_i) - \mathcal{L}_e^D(u_i) \quad (12)$$

Let  $\hat{x}_j$  and  $\mathcal{L}_e(x_j)$  denote the de-interleaved version of the detected data estimates and extrinsic LLRs from the detector respectively. The channel decoder is a modified MAP decoder based on the BCJR algorithm [22] that accepts soft values of input bits and outputs soft values of the coded bits [30]. The soft values at the output of the decoder,  $\mathcal{L}_a(x_j)$ , derived in Appendix II, can be shown to be,

$$\mathcal{L}_a(x_j) = \mathcal{L}_e^D(x_j) + \mathcal{L}_e(x_j) + 2\hat{x}_j \quad (13)$$

where,  $\mathcal{L}_e(x_j)$  is the  $\hat{a}$  priori information from the detector and  $\mathcal{L}_e^D(x_j)$  is the extrinsic information of the decoder. The  $\hat{a}$  priori value to the detector  $\mathcal{L}_e^D(u_i)$  is the interleaved version of  $\mathcal{L}_e^D(x_j)$ .

The hard decisions of the decoded data symbols are re-encoded and interleaved to generate a channel estimate as illustrated in loop II of Fig. 4. As stated before, the Viterbi detector uses the estimate of the channel from the pilot to initially start the trellis in the first  $L + 1$  stages of the trellis, and estimates the channel from data hypothesis for stages  $t > L + 1$ . The estimator in (7), though useful when the quality of the channel is unknown, is not effective when improved channel estimates are available. The local averaging in the estimator is subject to local noise variations in the received symbol and is blind to the increased reliability of the channel estimate which is computed using the  $\hat{a}$  posteriori knowledge of the received symbols at the end of the decoder. Thus, the detector is simply set to equalize the received symbols using the computed channel estimate after iteration 1.

##### A. Iterative Detector Complexity

The iterative receiver is proposed to offset the exponential increase in complexity of the detector with increase in averaging order,  $L$ . Recognize that the complexity of the receiver involves two parts, detector and decoder. The design paradigm is to have fewer iterations with smaller averaging order such that, to achieve a set performance level (say, optimal ML performance), the complexity of the iterative scheme with averaging order  $L_1$  is less than the non-iterative receiver with averaging order  $L_2$ , *i.e.*,

$$O(iT_cNM^{L_1}) + iC < O(T_cNM^{L_2}) + C \quad \text{for } L_2 > L_1 \quad (14)$$

where  $i$  is the number of iterations and  $C$  is the complexity of the decoder.

Current implementations of the decoder are computationally optimized and are routinely implemented in current generation turbo receivers. However, the actual computational complexity of the iterative receiver depends on the implementation of the decoder. The decoder, like a standard log-MAP implementation, has complexity,  $C$ , of the order  $O(NT_c(\log_2 M)c2^{\nu-1})$ , where  $c$  is the length of the codeword,  $\nu$  is the constraint length of the channel encoder. For higher order modulation and small  $\nu$ , the detector complexity is much higher than the



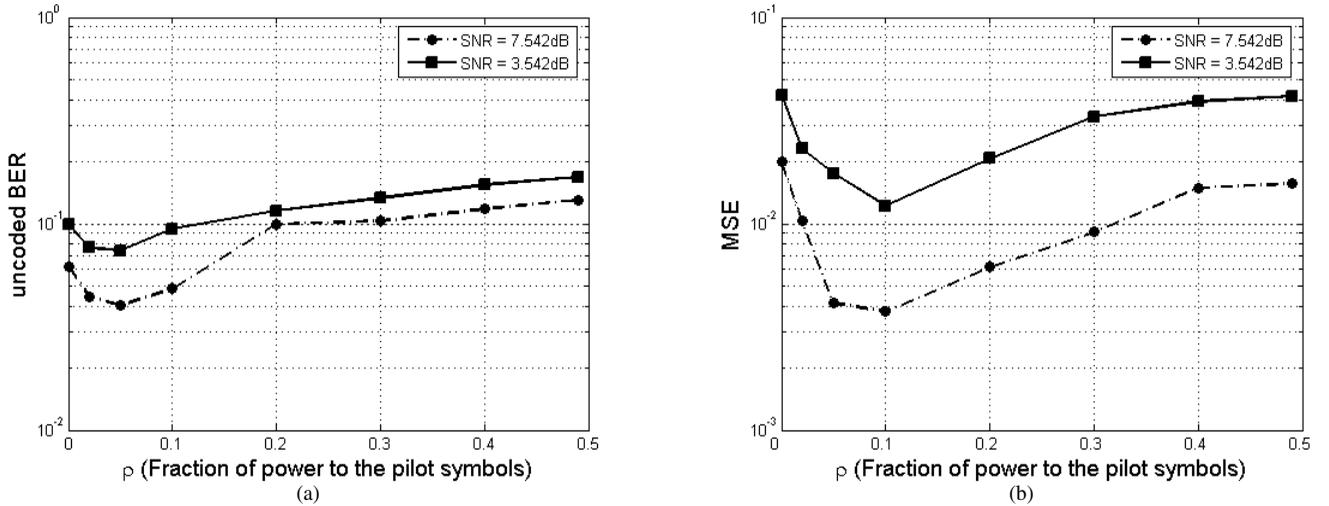


Fig. 5. (a) Optimization of the power  $\rho$  of the superimposed pilot using BER as metric. (b) MSE of the recomputed channel estimate with varying power  $\rho$  of the superimposed pilot.

TABLE III

OPTIMAL VALUES OF THE PILOT SYMBOL POWER FRACTION  $\rho$  THAT MINIMIZE THE BER AND MMSE WITH VARYING SNR AND COHERENCE TIME. FOR BOTH BPSK AND QPSK MODULATIONS, THE VALUES ARE OPTIMAL.

$T_c (\mu s)$	SNR = 3.542dB		SNR = 7.542dB		SNR = 11.542dB	
	$\rho_{BER}$	$\rho_{MMSE}$	$\rho_{BER}$	$\rho_{MMSE}$	$\rho_{BER}$	$\rho_{MMSE}$
48	0.05	0.1	0.05	0.1	0.05	0.1
128	0.1	0.15	0.1	0.15	0.08	0.1
256	0.1	0.2	0.1	0.15	0.08	0.1
512	0.15	0.25	0.15	0.2	0.1	0.15
2048	0.18	0.3	0.15	0.2	0.1	0.18
4096	0.2	0.3	0.15	0.2	0.15	0.2

two different values of SNR. The variation of the MSE with  $\rho$  is similar to the behavior of BER; however the minimum value of MSE occurs at  $\rho = 0.1$ . Further, the value of  $\rho$  that minimizes BER is around 5% for other values of  $L$ . The optimal  $\rho$  concurs with the result in [9] which states that the BER is minimized when the pilot power is approximately 5% of the total transmit power.

It must be noted that the optimal  $\rho$  holds only in the current simulation set-up. For different coherence time  $T_c$  the optimal values of the pilot power fraction that minimizes BER ( $\rho_{BER}$ ) and MMSE ( $\rho_{MMSE}$ ) are summarized in Table III. It is seen that with increasing block length for a given SNR, the optimal  $\rho$  increases. Where as with increase in SNR, for a given coherence time, the required  $\rho$  is smaller. For a coherence time of  $T_c = 48 \mu s$ , the optimal  $\rho = 0.05$  minimizes the BER for all SNRs. Therefore, we super-imposed the pilots at 5% of the total transmit power for all our simulations. However, a theoretical justification for  $\rho$  that minimizes BER or MMSE still remains an open problem.

### B. BER Performance

The BER of the proposed receiver are evaluated over 1000 independent realizations of the channel for different values

of the averaging order,  $L$ , and modulation order,  $M$ . In particular, we present results for BPSK and QPSK modulations corresponding to  $M = 2, 4$  respectively. The BER of the proposed receivers are compared to the *genie-aided* receiver which detects the data from the received signal assuming perfect knowledge of the channel.

1) *Non-Iterative Receivers*: We now demonstrate the effectiveness of channel estimation with super-imposed pilots using the proposed algorithm for an uncoded OFDM system. The variation of BER with SNR for BPSK modulation using averaging orders  $L = 3$  and 5 is shown in Fig. 6(a). Also shown is the performance of the *genie-aided* receiver, which provides a lower bound on the BER of any iterative receiver. Increasing averaging order improves the BER at a given SNR due to improved channel estimation and efficient equalization at each stage of the trellis. From the figure, we can see that with averaging order,  $L = 5$ , the BER of proposed receiver nearly equals the *genie-aided* receiver performance.

The proposed receiver's performance with QPSK modulation is given in Fig. 6(b). For comparison, the performance of a conventional training based channel estimation method is also shown. Training based channel estimation in 802.11a use the transmitted long and short training symbols in the preamble of the packet to estimate the channel. If there are  $n$  pilot symbols, since the channel is block fading, the channel estimate from the first  $n$  received symbols equalize the channel in the  $(T_c - n)$  OFDM symbols that follow. It can be clearly seen in Fig. 6(b) that the proposed system using super-imposed pilots gives lower bit error rates than the conventional estimation scheme employed in the 802.11a networks. We also plot the BER of a *genie-aided* receiver. The BER of a *genie-aided* receiver using QPSK modulation is analytically derived in Appendix III and is given by,

$$P_{QPSK}(E) = \frac{1}{2} \left( 1 - \sqrt{\frac{\bar{\gamma}_b}{\bar{\gamma}_b + 2}} \right), \quad (15)$$

where,  $\bar{\gamma}_b = \frac{(1-\rho)\sigma_f^2}{N_0}$ , is the average signal to noise ratio and  $\sigma_f^2$  is the variance in the channel coefficients.

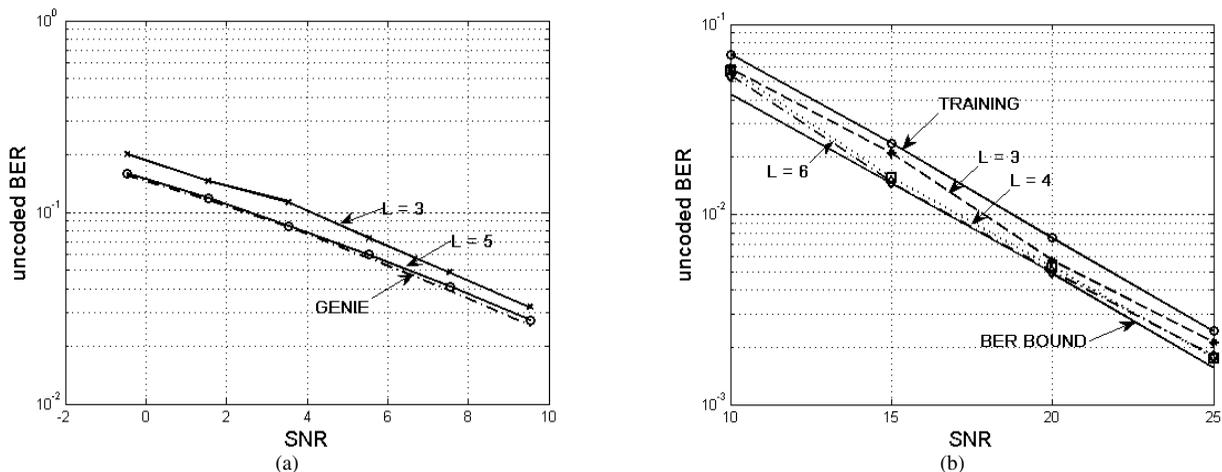


Fig. 6. (a) Uncoded BER of proposed receiver with BPSK modulation for different averaging orders. (b) Comparison of the performance of the proposed receiver using QPSK modulation with the training based method used in 802.11a and the BER bound in (15).

For averaging order,  $L = 6$ , the performance of the proposed receiver nearly equals the bound. Thus, increasing the averaging order beyond 6 results in only a marginal improvement in performance. Further, we recognize that using averaging order to 4, results in a very small loss in performance, but substantial savings in terms of complexity. Hence, the proposed receiver offers graceful trade-off of complexity with performance.

The BER of the proposed receiver with rate  $\frac{1}{2}$  channel coding is shown in Fig. 7(a) for averaging orders  $L = \{2, 4, 5\}$  and BPSK modulation. As in the uncoded case, the performance of the receiver with channel coding at averaging order,  $L = 5$ , almost equals that of the *genie-aided* based receiver. Similarly, with QPSK modulation and the same rate  $\frac{1}{2}$  channel code, near-optimal BER is achieved at  $L = 6$  as shown in Fig. 8(a). We observe that the performance curves plotted do not change with changing RMS delay spread or the coherence time of the channel. The RMS delay spread is adequately compensated by the length of the cyclic prefix in the OFDM symbol, ensuring no ISI. Therefore no-ISI ensures the invariance of the performance to the RMS delay spread. We tested our algorithm for varying packet sizes with the minimum size corresponding to a coherence time of  $12.8\mu s$  and found that the BERs did not vary with changes in the length of the packet. The invariance in performance to the coherence time is due to the fact that changes in coherence time changes the number of OFDM symbols in the packet, but the channel remains fixed.

2) *Low Complexity Iterative Receivers*: Recall that the goal in designing iterative receivers is to achieve superior BER (equalling lower bound if possible) at reduced complexity. The performance of the iterative receiver with BPSK modulation and averaging order,  $L = 2$ , is shown in Fig. 7(b). It can be observed, that the BER after iteration 2 almost equals that of the non-iterative *genie-aided* receiver. Further, most of the gain of the iterative receiver is obtained after the second iteration. Increasing the number of iterations beyond 2 result only in marginal improvement in BER.

Also, from Fig. 7(b), it can be seen that the performance

TABLE IV  
COMPLEXITY COMPARISON BETWEEN VARIOUS DETECTORS WITH QPSK MODULATION AT NEAR-OPTIMAL PERFORMANCE. THE NUMBER OF DATA SUB-CARRIERS IN 802.11A WLAN SYSTEM IS  $N = 48$ .

Detector	Order of Complexity	Typical Value(QPSK)
ML	$O(NT_c M^{T_c})$	$48 \times T_c \times 4^{T_c}$
Viterbi	$O(NT_c M^L)$	$48 \times T_c \times 4^6 = 196608T_c$
iterative Viterbi	$O(iNT_c M^L)$	$3 \times 48 \times T_c \times 4^2 = 2304T_c$

of the *genie-aided* receiver can also be improved with iteration due to the exchange of soft information between the detector and decoder. We also notice a performance gap of 1dB between the second iteration of *genie-aided* receiver and the iterative receiver with  $L = 2$ . We observed using simulations, that increasing the averaging order to  $L = 5$ , the performance of the iterative receiver is nearly that of the iterative *genie-aided* receiver, similar to the non-iterative case. The performance improvement for the iterative receiver with increasing averaging order is similar and is not shown.

The BER of an iterative OFDM system, with QPSK modulation is plotted in Fig. 8(b) for 4 iterations and averaging order  $L = 2$ . The results are similar to the iterative receivers using BPSK modulation, in that, there is very little gain after the third iteration. But unlike in the BPSK case, we see that the we achieve the performance of the *genie-aided* receiver only at the third iteration.

## VI. COMPLEXITY ANALYSIS

Of the detectors proposed in the paper, *i.e.*, ML, Viterbi and iterative Viterbi, ML is optimal. However, we showed via simulations that Viterbi and iterative Viterbi perform near-optimal at different complexities. The overall complexity of each detector to achieve ML equivalent performance with QPSK modulation is summarized in Table IV. While for really fast fading channels (small  $T_c$ ), the complexities of the detectors may be comparable, for slow fading channels (large  $T_c$ ), the reduction in complexity with the iterative receivers is

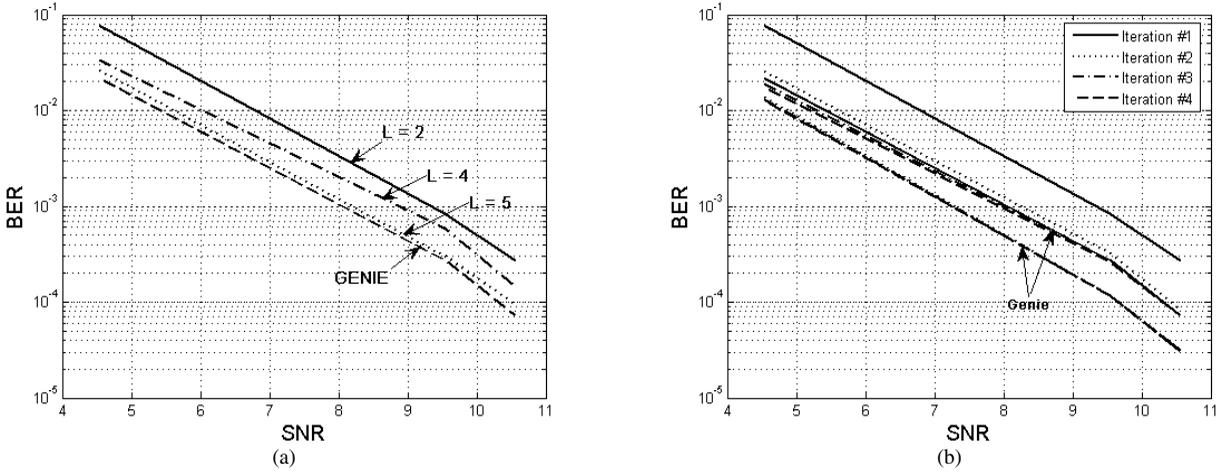


Fig. 7. (a) Coded BER of proposed non-iterative receiver for different averaging orders, with BPSK modulation and channel coding rate  $\frac{1}{2}$ . (b) Coded BER of proposed iterative OFDM receiver for averaging order  $L = 2$ , with BPSK modulation.

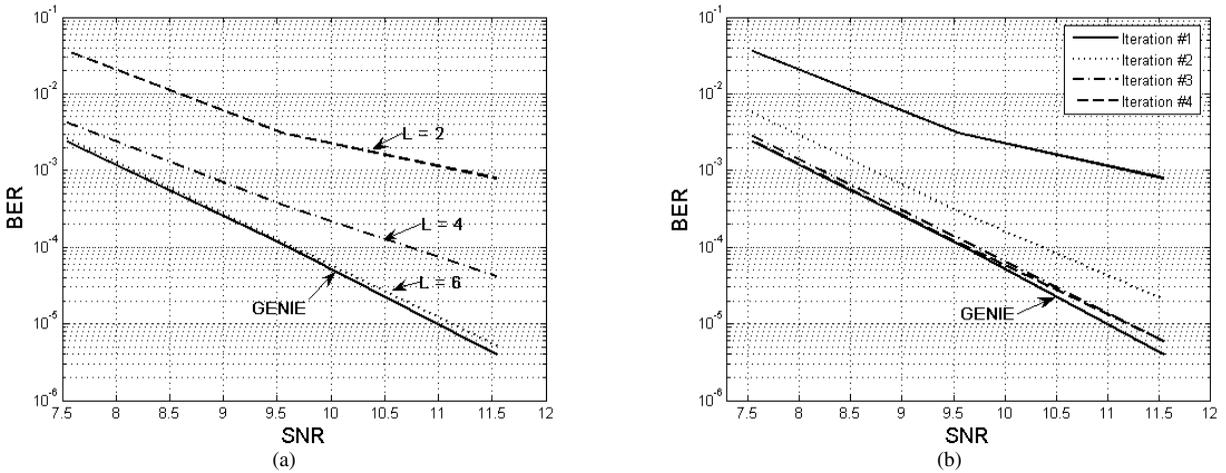


Fig. 8. (a) Coded BER of non-iterative receiver for different averaging orders, with QPSK modulation and channel coding rate  $\frac{1}{2}$ . (b) Coded BER of iterative receiver for averaging order  $L = 2$ , with QPSK modulation.

substantial. Since, the typical usage of WLANs is in stationary and slow moving environments, where the channel is slow fading, the effect of reduced complexity is emphatic.

The complexity of the iterative receiver with BPSK modulation, set to perform only two iterations,  $i = 2$ , with averaging order,  $L = 2$ , is substantially less than that of a non-iterative receiver with averaging order,  $L = 5$ . The performance of both the receivers under consideration, are almost the same and are closest to the genie, as can be seen from Figs. 7(a) and 7(b). In the current simulation set up, the number of arithmetic operations in terms of the coherence time,  $T_c$ , for two iterations of the iterative receiver with  $L = 2$  is of the order of  $O(\max(2T_c \times 48 \times 2^2, 2T_c \times 48 \times 2 \times 2^2)) = 768T_c$ . In the same simulation set up, the non iterative receiver with  $L = 5$  takes about  $O(\max(T_c \times 48 \times 2^5, T_c \times 48 \times 2 \times 2^2)) = 1536T_c$  arithmetic operations in the detector. By using the iterative process, the decoder saves about 50% in arithmetic operations needed to achieve the same result as the non-iterative case for block fading channels.

With QPSK modulation, the savings are more substantial.

The number of iterations required to achieve near optimality is  $i = 3$ , with averaging order  $L = 2$ . Similar performance with an non-iterative receiver is achieved at  $L = 6$ . The total number of computations in the iterative receiver in terms of the coherence time is calculated to be  $O(\max(3T_c \times 48 \times 4^2, 3T_c \times 2 \times 48 \times 2 \times 2^2)) = 2304T_c$  operations. In the same simulation set up, the non-iterative receiver needs  $O(\max(T_c \times 48 \times 4^6, T_c \times 48 \times 2^4)) = 196608T_c$  arithmetic operations, which amounts to about 98% savings in computational power. It should be noted that in commercial applications, where a larger constraint length channel code is used, the reduction in complexity might not be this dramatic.

## VII. CONCLUSION

High spectral efficiency is an important goal in designing high rate, next generation transceiver systems. Pilots that are linearly super-imposed on the data achieve that goal, in that, there is no bandwidth expansion due to pilot symbols. In this paper, we apply the concept of super-imposed training for channel estimation in OFDM systems. Our approach in

this paper was to leverage the presence of super-imposed pilots to obtain better channel estimates while ensuring data detection does not suffer. Such a complementing interplay between channel estimation and data detection in the Viterbi detector results in a near-optimal BER performance as has been established by simulations. Further, we also showed that the Viterbi detector with super-imposed pilots outperforms the conventional training based channel estimation currently used in 802.11a systems. We proposed iterative and non-iterative receivers whose operating modes trade-off performance and computational complexity. The BER performance of the proposed receivers was invariant over a wide range of channel coherence times. Hence, the proposed receivers are applicable equally to fast and slow - varying channels. The receivers' robustness to multipath delays in the channel makes it an attractive option for channel estimation in ultra wide-band (UWB) systems where the number of multi-paths are irresolvable.

A theoretical basis for optimal pilot power fraction that minimizes the BER is still elusive. Future work in this area should focus on optimizing pilot power fraction, sequence and their placement in the data sequence. In this paper, we estimate the channel at each sub-carriers independently without exploiting the correlation of channel impulse response across sub-carriers. It is our belief, that exploiting the correlation among sub-carriers could help in designing simpler receivers. The use of super-imposed pilots in systems with spatial diversity and in multi-user environment is another rich area for exploration.

#### APPENDIX I

##### THE SOFT-OUTPUT VITERBI DETECTOR METRIC

We state in this appendix, relevant modifications to the derivation of the metric for the soft-output Viterbi detector given in [18]. The conditional probability density function,  $g(r_t/u_t, \hat{f})$ , of the received signal at time  $t$ , can be written as:

$$g(r_t/u_t, \hat{f}) = \frac{1}{\sqrt{2\pi\sigma_z^2}} \exp \left[ -\frac{1}{2\sigma_z^2} (r_t - \hat{f}(u_t + p_t))^2 \right] \quad (16)$$

We define the transition probability,  $\gamma_t(\acute{s}, s)$ , at time instant  $t$ , between states  $S_{t-1} = \acute{s}$  and  $S_t = s$  as,

$$\begin{aligned} \gamma_t(\acute{s}, s) &= g(r_t, S_t = s/S_{t-1} = \acute{s}) \\ &= g(r_t/S_t = s, S_{t-1} = \acute{s})g(S_t = s/S_{t-1} = \acute{s}) \\ &= g(r_t/u_t, \hat{f})g(u_t) \end{aligned} \quad (17)$$

For the Viterbi detector to choose the correct ML sequence, the probabilities of surviving paths of all states in the trellis need to be evaluated. Let  $\mathbf{s}_t^s$  be the states along the surviving paths at state  $S_t = s$  at time  $t$  in the trellis. Probability that  $\mathbf{s}_t^s$  is the correct path defined as  $g(\mathbf{s}_t^s/\mathbf{r}_{j \leq t})$  is given by,

$$g(\mathbf{s}_t^s/\mathbf{r}_{j \leq t}) = \frac{g(\mathbf{s}_t^s, \mathbf{r}_{j \leq t})}{g(\mathbf{r}_{j \leq t})} \propto g(\mathbf{s}_t^s, \mathbf{r}_{j \leq t}) \quad (18)$$

The above simplification is valid because the received symbol probability,  $g(\mathbf{r}_{j \leq t})$ , is path independent. The independence of noise values from one sample to another is used to simplify  $g(\mathbf{s}_t^s/\mathbf{r}_{j \leq t})$  as,

$$\begin{aligned} g(\mathbf{s}_t^s, \mathbf{r}_{j \leq t}) &= g(\mathbf{s}_{t-1}^{\acute{s}}, \mathbf{r}_{j \leq t-1})g(r_t, s/\acute{s}) \\ &= g(\mathbf{s}_{t-1}^{\acute{s}}, \mathbf{r}_{j \leq t-1})\gamma_t(s, \acute{s}) \end{aligned} \quad (19)$$

The computation of the probability,  $g(\mathbf{s}_t^s, \mathbf{r}_{j \leq t})$ , is fundamental to ML sequence computation in the Viterbi algorithm. However, computing these probabilities can lead to numerically instability. Hence, for stable implementations, we take the logarithm of the computed densities. We define the metric of the surviving path,  $\mathcal{M}(\mathbf{s}_t^s)$ , as  $\mathcal{M}(\mathbf{s}_t^s)$ , given by

$$\begin{aligned} \mathcal{M}(\mathbf{s}_t^s) &= \ln(g(\mathbf{s}_t^s, \mathbf{r}_{j \leq t})) \\ &= \ln(g(\mathbf{s}_{t-1}^{\acute{s}}, \mathbf{r}_{j \leq t-1})) + \ln(\gamma_t(s, \acute{s})) \end{aligned} \quad (20)$$

Substituting the expression for  $\gamma_t(s, \acute{s})$  from (17) in (20), and expanding the logarithm, we get,

$$\begin{aligned} \mathcal{M}(\mathbf{s}_t^s) &= \mathcal{M}(\mathbf{s}_{t-1}^{\acute{s}}) + \ln \left( \frac{1}{\sqrt{2\pi\sigma_z^2}} \right) \\ &\quad - \frac{1}{2\sigma_z^2} |r_t - \hat{f}(u_t + p_t)|^2 \\ &\quad + \frac{1}{2} \sum_{i=1}^{\log_2 M} \tilde{b}_{i,n,t} \mathcal{L}_a(b_{i,n,t}) \end{aligned} \quad (21)$$

Ignoring the constant term,  $\ln \left( \frac{1}{\sqrt{2\pi\sigma_z^2}} \right)$ , we get the expression for the metric in (8).

#### APPENDIX II

##### BLOCK SOFT-IN SOFT-OUT (SISO) DECODER

This appendix describes the SISO algorithm that calculates the soft output for the coded bit stream [30] based on MAP algorithm. Consider a rate  $\frac{1}{2}$  systematic convolutional code. Let  $d_l$  be the input data symbol,  $c_l$  be the output code symbol at stage  $l$  in the trellis. The relation between the code word at stage  $l$  in the trellis and the data estimates at time  $j$  and  $j+1$  respectively is given by,

$$[d_l c_l] = [\hat{x}_j \hat{x}_{j+1}] \quad (22)$$

We use the following identity:

$$a = \log \left[ \sum_{i=1}^L e^{a_i} \right] = \max_i \{a_i\} + \delta(a_1, \dots, a_L) = \max_i^* \{a_i\} \quad (23)$$

where  $\delta(a_1, \dots, a_L)$  is the correction term and that can be computed using a look-up table [30]. The  $\max^*$  operation is defined as maximization plus a correction term.

Define trellis edge,  $e$ , beginning at state  $\acute{s}(e)$  and ending at  $s(e)$ . Let  $[\tilde{d}_l(e) \tilde{c}_l(e)]$  be the codeword associated with the edge  $e$  at stage  $l$  in the trellis. At stage  $l$  in the trellis, the forward and backward recursion can be written as,

$$\begin{aligned} \alpha_l(s) &= \max_{e: \acute{s}(e)=s}^* \{ \alpha_{l-1}[\acute{s}(e)] + d_l(e)\tilde{d}_l(e) + c_l(e)\tilde{c}_l(e) \} + h_{\alpha_l} \\ \beta_l(\acute{s}) &= \max_{e: \acute{s}(e)=s}^* \{ \beta_{l+1}[s(e)] + d_l(e)\tilde{d}_l(e) + c_l(e)\tilde{c}_l(e) \} + h_{\beta_l} \end{aligned}$$

with initial values  $\alpha_0(s) = 0$  if  $s = S_0$ , and  $\alpha_0(s) = -\infty$  otherwise, and  $\beta_E(s) = 0$  if  $s = S_E$  and  $\beta_E(s) = -\infty$  otherwise, where  $E$  is the terminating stage in the trellis,  $h_{\alpha_l}$  and  $h_{\beta_l}$  are normalization constants used to prevent buffer overflow.

The soft value at the output of the decoder can be obtained as,

$$\begin{aligned} \mathcal{L}_a(d_l) &= \max_{e:\tilde{d}(e)=1} \{ \alpha_l[\hat{s}(e)] + \gamma_{l_d}^*(\hat{s}(e), s(e)) + \beta_{l+1}[\hat{s}(e)] \} \\ &\quad - \max_{e:\tilde{d}(e)=-1} \{ \alpha_l[\hat{s}(e)] + \gamma_{l_d}^*(\hat{s}(e), s(e)) + \beta_{l+1}[\hat{s}(e)] \} \\ \mathcal{L}_a(c_l) &= \max_{e:\tilde{c}(e)=1} \{ \alpha_l[\hat{s}(e)] + \gamma_{l_c}^*(\hat{s}(e), s(e)) + \beta_{l+1}[\hat{s}(e)] \} \\ &\quad - \max_{e:\tilde{c}(e)=-1} \{ \alpha_l[\hat{s}(e)] + \gamma_{l_c}^*(\hat{s}(e), s(e)) + \beta_{l+1}[\hat{s}(e)] \} \end{aligned}$$

where  $\gamma_{l_d}^*(\hat{s}(e), s(e)) = d_l \tilde{d}_l + c_l \tilde{c}_l + \frac{d_l \mathcal{L}_e(d_l)}{2}$  and  $\gamma_{l_c}^*(\hat{s}(e), s(e)) = d_l \tilde{d}_l + c_l \tilde{c}_l + \frac{c_l \mathcal{L}_e(c_l)}{2}$ .

Substituting for  $\tilde{d}$  and  $\tilde{c}$  in the expression for soft output values and simplifying, we get,

$$\begin{aligned} \mathcal{L}_a(d_l) &= 2d_l + \mathcal{L}_e(d_l) + \{ \max_{e:\tilde{d}(e)=1} [c_l \tilde{c}_l + \beta_{l+1}(s(e))] \} \\ &\quad - \{ \max_{e:\tilde{d}(e)=-1} [c_l \tilde{c}_l + \beta_{l+1}(s(e))] \} \\ \mathcal{L}_a(c_l) &= 2c_l + \mathcal{L}_e(c_l) + \{ \max_{e:\tilde{c}(e)=1} [d_l \tilde{d}_l + \beta_{l+1}(s(e))] \} \\ &\quad - \{ \max_{e:\tilde{c}(e)=-1} [d_l \tilde{d}_l + \beta_{l+1}(s(e))] \} \end{aligned}$$

which when converted to the correct time index using (22) gives the expression in (13).

### APPENDIX III

#### BER CHARACTERIZATION OF QPSK SIGNALS WITH SUPER-IMPOSED PILOTS

Let  $\hat{f}_n$  denote the estimate of the channel for subcarrier  $n$ . For a given  $\hat{f}_n$ , the transmitted symbols can be recovered from the received symbols in (3) as,

$$\hat{u}_{n,t} = \frac{\hat{f}_n}{f_n} (u_{n,t} + p_{n,t}) + \frac{z_{n,t}}{\hat{f}_n} - p_{n,t} \quad (24)$$

The channel estimate  $\hat{f}_n$  is assumed to have a complex Gaussian distribution with zero mean and variance  $\sigma_f^2$ . Thus, the amplitudes  $|f_n|$  and  $|\hat{f}_n|$  are Rayleigh distributed and their joint distribution is a bivariate Rayleigh distribution,  $g(|f_n|, |\hat{f}_n|)$ . Denote the bit error probability conditioned on  $|f_n|$  and  $|\hat{f}_n|$  as  $g(\hat{u}_{n,t} \neq u_{n,t} / |f_n|, |\hat{f}_n|)$ . Then, the average bit error rate is given as [35],

$$\begin{aligned} P(E) &= \int_0^\infty \int_0^\infty g(\hat{u}_{n,t} \neq u_{n,t} / |f_n|, |\hat{f}_n|) \\ &\quad g(|f_n|, |\hat{f}_n|) d|f_n| d|\hat{f}_n| \end{aligned} \quad (25)$$

The bits are mapped using a gray code to the real and imaginary parts of the modulated symbol. With this mapping, the demodulation is a simple sign detection of the real and imaginary parts independently. Due to symmetry, the BERs of both the MSB and LSB are equal. Hence, the conditional BER can be evaluated as,

$$\begin{aligned} g(\hat{u}_{n,t} \neq u_{n,t} / |f_n|, |\hat{f}_n|) &= \frac{1}{2} [P_{MSB}(E) + P_{LSB}(E)] \\ &= P_{MSB}(E) \end{aligned} \quad (26)$$

Since the noise variance is  $\frac{\sigma_z^2}{|f_n|^2}$ , we can write the bit error probability of the MSB conditioned on knowing  $|f_n|$  and  $|\hat{f}_n|$

as [15],

$$P_{MSB}(E) = Q \left( \frac{\frac{|f_n|}{\sigma_z} d}{\frac{|f_n|}{\sigma_z}} \right) = Q \left( \frac{|f_n| d}{\sigma_z} \right), \quad (27)$$

where  $d$  is the minimum distance between QPSK constellation points. The average energy in the constellation is  $2d^2$  and can be expressed as  $d = \sqrt{\frac{(1-\rho)E_b}{2}}$ . Substituting for  $d^2$  and  $\sigma_z = \sqrt{\frac{N_0}{2}}$  in (27), the conditional bit error rate when the MSB is in error is given by,

$$P_{MSB}(E) = Q \left( \sqrt{\frac{|f_n|^2 (1-\rho) E_b}{N_0}} \right). \quad (28)$$

Since the conditional distribution is independent of  $|\hat{f}_n|$ , the average bit error rate for QPSK modulated symbols is obtained by averaging the conditional BER over the marginal distribution of  $|f_n|$ .

$$P_{QPSK}(E) = \int_0^\infty Q \left( \sqrt{\frac{|f_n|^2 (1-\rho) E_b}{N_0}} \right) g(|f_n|) d|f_n| \quad (29)$$

where,  $g(|f_n|)$  is the Rayleigh distribution given by  $p(|f_n|) = \frac{2|f_n|}{\bar{\gamma}_b} e^{-\frac{|f_n|^2}{\bar{\gamma}_b}}$  where,  $\bar{\gamma}_b = \frac{(1-\rho)E_b}{N_0} \sigma_f^2$ , is the average signal to noise ratio. The integral in (29) evaluates to the following expression [15],

$$P_{QPSK}(E) = \frac{1}{2} \left( 1 - \sqrt{\frac{\bar{\gamma}_b}{\bar{\gamma}_b + 2}} \right). \quad (30)$$

### REFERENCES

- [1] A. Bahai and B. Saltzberg, *Multicarrier Digital Communications: Theory and Applications of OFDM*. New York: Kluwer Academic, 1999.
- [2] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," *IEEE Trans. Broadcasting*, vol. 27, no. 3, pp. 223-229, Sep. 2002.
- [3] L. Tong, B. M. Sadler, and M. Dong, "Pilot-assisted wireless transmissions - General model, design criteria, and signal processing," *IEEE Signal Processing Mag.*, vol. 21, no. 6, pp. 12-25, Nov. 2004.
- [4] H. Liu, G. Xu, L. Tong, and T. Kailath, "Recent developments in blind channel equalization: From cyclostationarity to subspaces," *IEEE Trans. Signal Processing*, vol. 50, no. 1-2, pp. 83-99, Apr. 1996.
- [5] B. Muquet, M. de Courville, and P. Duhamel, "Subspace based blind and semi-blind channel estimation for OFDM systems," *IEEE Trans. Signal Processing*, vol. 50, no. 7, pp. 1699-1712, July 2002.
- [6] A. Petropulu, R. Zhang, and R. Lin, "Blind OFDM channel estimation through simple linear precoding," *IEEE Trans. Wireless Commun.*, vol. 3, no. 2, pp. 647-655, Mar. 2004.
- [7] C. Kastenholz and W. Birkemeir, "A simultaneous information transfer and channel sounding modulation technique for wide-band channels," *IEEE Trans. Commun. Technol.*, pp. 162-175, June 1965.
- [8] B. Farhang-Boroujeny, "Pilot-based channel identification: Proposal for semi-blind identification of communication channels," *IEE Electron. Lett.*, vol. 31, no. 13, pp. 1044-1046, June 1995.
- [9] P. Hoeher and F. Tufvesson, "Channel estimation with super-imposed pilot sequence," in *Proc. IEEE Global Commun. Conf.*, Dec. 1999, pp. 2162-2166.
- [10] G. Zhou, M. Viberg, and T. Mckelvey, "Super-imposed periodic pilots for blind channel estimation," in *Proc. 35th Asilomar Conf. Signals, Syst., Computers*, Nov. 2001, pp. 653-657.
- [11] N. Chen and G. Zhou, "A superimposed periodic pilot scheme for semi-blind channel estimation of OFDM systems," in *Proc. 10th IEEE DSP Workshop*, Oct. 2002, pp. 362-365.
- [12] J. Tugnait and W. Luo, "On channel estimation using super-imposed training and first order statistics," in *Proc. IEEE International Conf. Acoustics, Speech, Signal Processing*, Apr. 2003, pp. 624-627.

- [13] A. G. Orozco-Lugo, M. M. Lara, and D. C. McLernon, "Channel estimation using implicit training," *IEEE Trans. Signal Processing*, vol. 52, no. 1, pp. 240–254, Jan. 2004.
- [14] J. Terry and J. Heiskala, *OFDM Wireless LANs: A Theoretical and Practical Guide*, 1st ed. Indianapolis, IN: Sams Publishing, 2001.
- [15] J. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 1996.
- [16] *Wireless LAN medium access control (MAC) and physical layer (PHY) specifications*, IEEE Standard 802.11a, pt. 11, Apr. 1999.
- [17] F. Tufvesson, M. Faulkner, P. Hoeher, and O. Edfors, "OFDM time and frequency synchronization by spread spectrum pilot technique," in *Proc. 8th Commun. Theory Mini-Conf./IEEE ICC*, June 1999, pp. 115–119.
- [18] J. P. Woodward and L. Hanzo, "Comparative Study of turbo decoding techniques: An overview," *IEEE Trans. Veh. Technol.*, vol. 49, no. 6, pp. 2208–2233, Nov. 2000.
- [19] J. Hagenauer and P. Hoeher, "A Viterbi algorithm with soft-decision outputs and its applications," in *Proc. IEEE Global Commun. Conf.*, Nov. 1989, vol. 3, pp. 1680–1686.
- [20] J. Hagenauer, "Source-controlled channel decoding," *IEEE Trans. Commun.*, vol. 43, no. 9, pp. 2449–2457, Sep. 1995.
- [21] G. M. Vitetta and D. P. Taylor, "Maximum likelihood sequence estimation of uncoded and coded PSK signals transmitted over Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 43, no. 11, pp. 2750–2758, Nov. 1995.
- [22] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. 20, pp. 284–287, Mar. 1974.
- [23] G. Bauch, H. Khorram, and J. Hagenauer, "Iterative equalization and decoding in mobile communications systems," in *2nd European Pers. Mobile Commun. Conf./ITG-Fachtagung Mobile Kommunikation*, VDE/ITG, Sep./Oct. 1997, pp. 307–312.
- [24] A. Anastasopoulos and K. M. Chugg, "Adaptive soft-input soft-output algorithms for iterative detection with parametric uncertainty," *IEEE Trans. Commun.*, vol. 48, no. 10, pp. 1638–1649, Oct. 2000.
- [25] M. C. Valenti and B. D. Woerner, "Iterative channel estimation and decoding of pilot symbol assisted turbo codes over flat-fading channels," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 1697–1705, Sep. 2001.
- [26] C. Kominakis and R. D. Wesel, "Joint iterative channel estimation and decoding in flat correlated Rayleigh fading," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 1706–1717, Sep. 2001.
- [27] B. Mielczarek and A. Svensson, "Improved iterative channel estimation and turbo decoding over flat-fading channels," in *Proc. IEEE Veh. Technol. Conf.*, Sep. 2002, vol. 2, pp. 975–980.
- [28] H. El Gamal and E. Geraniotis, "Iterative channel estimation and decoding for convolutionally coded anti-jam FH signals," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 321–331, Feb. 2002.
- [29] M. Tüchler, R. Koetter, and A. Singer, "Turbo equalization: Principles and new results," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 754–767, May 2002.
- [30] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial concatenation of interleaved codes: Performance analysis, design and iterative decoding," TDA Progress Report 42-126, Apr.–June 1996, Jet Propulsion Lab., Pasadena, CA, pp. 1–26, Aug 1996. [Online.] (Available: [http://tmo.jpl.nasa.gov/tmo/progress\\_report/42-126/126D.pdf](http://tmo.jpl.nasa.gov/tmo/progress_report/42-126/126D.pdf))
- [31] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "A soft-input soft-output APP module for iterative decoding of concatenated codes," *IEEE Commun. Lett.*, vol. 1, no. 1, pp. 22–24, Jan. 1997.
- [32] P. Robertson, P. Hoeher, and E. Villebrun, "Optimal and sub-optimal maximum a posteriori algorithms suitable for turbo-decoding," *European Trans. Telecommun.*, vol. 8, pp. 118–125, Mar./Apr. 1997.
- [33] S. Tsai, T.-T. Chen, and J. S. Lehnert, "Implementation of fixed-delay soft-output algorithms," in *Proc. IEEE Military Commun. Conf.*, Oct. 1998, vol. 3, pp. 993–997.
- [34] K. M. Chugg and X. Chen, "Efficient architectures for soft-output algorithms," in *Proc. IEEE Int. Conf. Commun.*, June 1998, pp. 126–130.
- [35] X. Tang, M.-S. Alouini, and A. Goldsmith, "Effect of channel estimation error in M-QAM BER performance in Rayleigh fading," *IEEE Trans. Commun.*, vol. 47, no. 12, pp. 83–99, Dec. 1999.
- [36] K. Josiam and D. Rajan, "Bandwidth efficient channel estimation using super-imposed pilots," in *Proc. 39th Conf. Inform. Science Syst.*, Mar. 2005.



**Kaushik Josiam** (S'01) received his B.E. degree in electronics and communications engineering from Bangalore University Visvesvaraya College of Engineering, Bangalore, India in 2001 and a M.S. degree in electrical engineering from Southern Methodist University, Dallas in 2004. He is currently a PhD candidate at SMU. His research interests are in wireless communications, information theory and communication theory.



**Dinesh Rajan** (M'02) is currently an Assistant Professor in the electrical engineering department at SMU. Dinesh received his Ph.D(2002) and MS(1999) in electrical and computer engineering from Rice University and his B.Tech (1997) from IIT, Madras, India. As a senior graduate student, he worked at Nokia Research Centers in Irving and Helsinki during summers of 1998 and 1999. He received an NSF CAREER award in 2006. His research interests are in information theory, wireless communication and networking.