Iterative Joint Decoding and Blind Channel Estimation for Coded OFDM systems

Ghazi Al-Rawi, Tareq Al-Nafouri, Ahmad Bahai, and John Cioffi

Abstract— **OFDM systems over frequency selective channels typically use coding across subchannels to exploit frequency diversity. This paper presents a new iterative technique for blind channel estimation in OFDM systems by combining soft decoding and maximum likelihood channel estimation. The proposed technique converges within a single OFDM symbol starting from an arbitrary initial estimate of the channel. A simple technique is suggested to detect convergence and stop iterations. The system performance using the proposed technique is compared to that of coded and uncoded systems that use traditional channel estimation techniques using enough number of pilots.**

Keywords— **Iterative detection, Blind channel estimation, Joint decoding/channel identification.**

I. Introduction

Recently there has been increasing interest in multicarrier modulation (MCM) schemes [1] for digital broadcasting (DAB, DVB) and mobile wireless broadband systems (ETSI BRAN, IEEE802.11a, and MMAC) that use orthogonal frequency division multiplexing (OFDM), and for high speed data transmission over twisted pairs (xDSL) that uses digital multitone (DMT) modulation.

In wireless systems, the use of differential phase-shift keying (DPSK) in OFDM systems avoids the tracking of a time varying channel. However, this limits the number of bits per symbol and results in a 3dB loss in signal-to-noise ratio (SNR) [2]. If channel estimation is done at the receiver, coherent detection, and hence multi-amplitude signaling schemes can be used. If the channel is changing slowly and there is significant correlation between the channel states across multiple OFDM symbols, reference pilot symbols or decision directed channel tracking techniques can be used [3], [4]. Blind channel estimation techniques allow higher data rates, since they avoid training overhead, but they typically require a large number of OFDM symbols to converge to an estimate of the channel that is accurate enough [5], [6], [7]. This not only introduces significant latency in the system, but also is limited to slowly varying channels. If the channel state can change significantly from one symbol to the next due to high Doppler frequency, channel estimation within a single OFDM symbol will be required.

Assuming the impulse response of channel has L non-zero taps and that it is fixed within a single OFDM symbols, L pilot subchannels equally spaced across the N subchannels in the frequency domain can be used to accurately estimate the channel within a single OFDM symbol [4]. A blind channel estimation technique that requires a small number of OFDM symbols to converge by taking advantage of the finite-alphabet property is proposed in [8].

OFDM systems usually use coding across subchannels to exploit frequency diversity in frequency selective channels [9], [4]. Coding gain can also be used to lower the transmit power while operating at the same bit error rate.

Iterative processing is a well-known complexity reduction technique that usually results in close to optimal performance. Iterative techniques have recently captured significant attention and have been proposed as practical solutions to many problems in communications systems [10], [11], [12], [13], [14].

This paper presents an iterative technique for joint soft decoding and blind channel estimation that converges within a single OFDM frame. Therefore, it has a minimum latency and is more appropriate for fast changing channels. The proposed technique does not require any initial estimate of the channel, and hence does not suffer from error propagation and run-away effects that is inherent in decision directed channel tracking techniques [3], [4]. It avoids the overhead of transmitting L pilots, which can be significant if L is a large fraction of N, at the expense of more sophisticated processing at the receiver side.

The system performance using the proposed technique is compared to that of coded systems that use the traditional way of first estimating the channel using L pilots and then doing soft decoding. The performance is also compared to that of uncoded system that uses L pilots for channel estimation. For consistency, all coded systems use the same 4-state rate 1/2 CC.

Section II of this paper presents the assumptions and notations used. Section III presents the proposed Joint iterative decoding and channel estimation technique. SectionIV presents the simulation and comparison results. Concluding remarks are presented in section V.

II. Assumptions and Notations

OFDM is a method for mitigating inter-symbol interference (ISI) in frequency selective channels. It divides the frequency band into N subchannels. Ideally these subchannels are orthogonal, and each one individually has a flat response and can be treated as a single tap no-ISI channel. We assume that each subchannel carries one bit per dimension (i.e., $\bar{b} = 1$). For simplification we assume that each subchannel uses binary phase shift keying (BPSK) modulation.

It is assumed that each OFDM system is coded separately. No coding across OFDM symbols is used. Coded systems referenced in this paper use a recursive systematic encoder with the generator matrix $G(D) = [1 \ \frac{1+D^2}{1+D+D^2}]$, where D is a delay operator. The systematic realization of the encoder allows the pilots to be added before encoding.

By doing this, pilots can not only be used for channel estimation but also to help in the soft decoding process. To increase the transmission efficiency, no trellis termination is assumed. For each coded OFDM symbol, the encoder is assumed to be initialized to state 0. The subchannels are assumed to be orthogonal, and hence no frequency interleaving is assumed.

Denote the input binary sequence to the convolutional encoder as $u(D)$, and the output sequence as $\tilde{\boldsymbol{v}}(D)$, where:

$$\tilde{\boldsymbol{v}}^T(D) = [v^1(D)\ v^0(D)] = u(D)\cdot G(D) \tag{1}$$

Let $\boldsymbol{u}$ be the input message vector of length $K = r \cdot N$, where $r = 1/2$ is the rate of the code, and let $\boldsymbol{v}$ be the multiplexed output vector of the encoder of length N:

$$\boldsymbol{v}^T = [\tilde{\boldsymbol{v}}_0^T\ \tilde{\boldsymbol{v}}_1^T\ \cdots \tilde{\boldsymbol{v}}_{K-1}^T] \tag{2}$$

where:

$$\tilde{\boldsymbol{v}}_k = \begin{bmatrix} v_k^1 \\ v_k^0 \end{bmatrix} = \begin{bmatrix} v_{2k} \\ v_{2k+1} \end{bmatrix} \tag{3}$$

where $k = 0, 1, \ldots, K-1$. Let $\boldsymbol{X}$, and $\boldsymbol{x}$ represent the transmitted signal vector in the frequency and time domains, respectively.

$$\tilde{\boldsymbol{X}} = (2\tilde{\boldsymbol{v}} - 1)\cdot\sqrt{E_x} \tag{4}$$

$$\boldsymbol{X} = \begin{bmatrix} \tilde{\boldsymbol{X}}_0 \\ \tilde{\boldsymbol{X}}_1 \\ \vdots \\ \tilde{\boldsymbol{X}}_{K-1} \end{bmatrix} = (2\boldsymbol{v} - 1)\cdot\sqrt{E_x} \tag{5}$$

$$\boldsymbol{x} = \boldsymbol{Q}^H \boldsymbol{X} \tag{6}$$

where $\boldsymbol{Q}^H$ is the conjugate transpose of the $N \times N$ Discrete Fourier Transform (DFT) matrix $\boldsymbol{Q}$. By appending $\boldsymbol{x}$ with cyclic prefix, the output of the channel can be obtained through the following cyclic convolution:

$$\boldsymbol{y} = \boldsymbol{h} \otimes \boldsymbol{x} + \boldsymbol{n} \tag{7}$$

where $\boldsymbol{h}$ is the channel impulse response, and $\boldsymbol{n}$ is complex additive white Gaussian noise (AWGN) vector with covariance matrix $\boldsymbol{R_{nn}} = \sigma^2\boldsymbol{I}$, where $\boldsymbol{I}$ is $N \times N$ identity matrix. Because of the cyclic convulotion, we can write:

$$\boldsymbol{Y} = \boldsymbol{H} \odot \boldsymbol{X} + \boldsymbol{N} \tag{8}$$

where $\odot$ represents element by element multiplication, and:

$$\boldsymbol{X} = \boldsymbol{Qx} \tag{9}$$
$$\boldsymbol{Y} = \boldsymbol{Qy} \tag{10}$$
$$\boldsymbol{H} = \boldsymbol{Qh} \tag{11}$$
$$\boldsymbol{N} = \boldsymbol{Qn} \tag{12}$$

Therefore, each subchannel i, where $i = 0, 1, \ldots, N-1$ can be treated as a single-tap AWGN channel with gain H_i and noise variance σ^2. Let $R_i = Y_i/H_i$ be the zero-forcing equalized output of subchannel i, and let $\tilde{\boldsymbol{R}}_k^T = [R_{2k}\ \ R_{2k+1}]$, where $k = 0, 1, \ldots, K-1$, and $\boldsymbol{R}^T = [\tilde{\boldsymbol{R}}_0^T\ \tilde{\boldsymbol{R}}_1^T \cdots \tilde{\boldsymbol{R}}_{K-1}^T]$.

This paper assumes that channel impulse response can have up to L non-zero taps from 0 to $L-1$ and that it is fixed over a single OFMD symbol. It is assumed, however, that there can not be any coordination between different OFMD symbols at the transmit or receive side. This could be because these OFDM symbols are coming from different users as in certain multiple access cases, or that the channel state changes significantly from one OFDM symbol to the next as in broadcast channels with very high Doppler frequency.

The objective is to jointly find the maximum likelihood (ML) estimate of the channel state $\boldsymbol{h}$, and the maximum likelihood estimate of the transmitted OFDM symbol $\boldsymbol{X}$. This joint ML estimation problem can be very difficult to solve optimally. This papers proposes a relatively low-complexity iterative technique to find a good quality approximate solution to this joint estimation problem.

III. Iterative Joint Decoding and Channel Estimation

The proposed technique exploits the following constraints that are typically known to the receiver and iteratively finds estimates of $\boldsymbol{H}$ and $\boldsymbol{X}$ that simultaneously satisfy all of them:
1. X_i can take a value in a finite alphabet.
2. $\boldsymbol{X}$ should satisfy the code constraint.
3. The channel impulse response $\boldsymbol{h}$ has non-zero taps from 0 to $L-1$ only.

The proposed algorithm is considered approximate and not exact, because it separately finds the ML soft estimate of $\boldsymbol{X}$ given $\boldsymbol{H}$, and then finds a ML estimate of $\boldsymbol{H}$ given ML soft estimate of $\boldsymbol{X}$. It then iterates between these two separate estimation processes until convergence occurs.

Given an arbitrary initial estimate of the channel $\boldsymbol{H}^{init} = \boldsymbol{Qh}^{init}$, the algorithm goes through the following three steps:

1. Finding the ML soft estimate of $\boldsymbol{X}$ given $\boldsymbol{H}$:

Max-Log-MAP algorithm gives a soft ML soft estimate of $\boldsymbol{X}$ given $\boldsymbol{H}$ [15]. For each X_i, the Max-Log-MAP algorithm gives a soft output L_i^{app}, which approximates the a posteriori log-likelihood ratio (LLR) of X_i.

$$L_i^{app} \approx \log\left(\frac{P(X_i = \sqrt{E_x}|\boldsymbol{Y}, \boldsymbol{H})}{P(X_i = -\sqrt{E_x}|\boldsymbol{Y}, \boldsymbol{H})}\right) \tag{13}$$

where $i = 0, 1, \ldots, N-1$. The soft outputs can be calculated as follows:

$$\begin{aligned} L_{2k}^{app} &= \max_{(l',l)\in B(v_k^1=1)} [\bar{\alpha}_{k-1}(l') + \bar{\gamma}_k(l',l) + \bar{\beta}_k(l)] \\ &- \max_{(l',l)\in B(v_k^1=0)} [\bar{\alpha}_{k-1}(l') + \bar{\gamma}_k(l',l) + \bar{\beta}_k(l)] \\ L_{2k+1}^{app} &= \max_{(l',l)\in B(v_k^0=1)} [\bar{\alpha}_{k-1}(l') + \bar{\gamma}_k(l',l) + \bar{\beta}_k(l)] \\ &- \max_{(l',l)\in B(v_k^0=0)} [\bar{\alpha}_{k-1}(l') + \bar{\gamma}_k(l',l) + \bar{\beta}_k(l)] \end{aligned}$$

(14)

where $k = 0, 1, \ldots, K-1$, (l', l) is the branch from state l' to state l, and $l', l = 0, 1, \ldots, M_s - 1$, where M_s is the number of states in the code trellis. $B(v_k^1 = 0(1))$ is the set of branches in the kth section of the trellis that have the first output label $v_k^1 = 0(1)$. Similarly, for $B(v_k^0 = 0(1))$. The branch metrics $\bar{\gamma}_k$ and forward and backward parameters $\bar{\alpha}_k$ and $\bar{\beta}_k$ are calculated as follows:

$$\bar{\gamma}_k(l', l) = \log p(u_k = 1) - \frac{d^2(\tilde{\boldsymbol{R}}_k, \tilde{\boldsymbol{X}}_k)}{2\sigma^2} \quad (15)$$

where $k = 0, 1, \ldots, K-1$, and $p(u_k = 1)$ is the a priori probability, which are $1/2$ for all values of k unless if pilots are sent as part of the codeword, in which case $p(u_k = 1) = 0$ or 1 for the pilot bits. $d^2(\tilde{\boldsymbol{R}}_k, \tilde{\boldsymbol{X}}_k)$ is the Euclidean distance between $\tilde{\boldsymbol{R}}_k$ and $\tilde{\boldsymbol{X}}_k$. Recall that $R_i = \frac{Y_i}{H_i^{init}}$ and $\tilde{\boldsymbol{R}}_k^T = [R_{2k}\ R_{2k+1}]$ for $k = 0, 1, \ldots, K-1$.

$$\bar{\alpha}_k(l) = \max_{l'} \{\bar{\alpha}_{k-1}(l') + \bar{\gamma}_k(l', l)\} \quad (16)$$

where $k = 0, 1, \ldots, K-2$, and $\bar{\alpha}_{-1}(l = 0) = 0, \bar{\alpha}_{-1}(l) = -\infty$ for $l \neq 0$.

$$\bar{\beta}_k(l) = \max_{l'} \{\bar{\beta}_{k+1}(l') + \bar{\gamma}_{k+1}(l, l')\} \quad (17)$$

where $k = 0, 1, \ldots, K-2$, and $\bar{\beta}_{K-1}(l) = 0$ for $l = 0, 1, \ldots, M_s - 1$, because the trellis is assumed to be not terminated.

The extrinsic LLR of X_i is obtained as follows:

$$L_i^{ext} = L_i^{app} - \frac{2\sqrt{E_x}}{\sigma^2} R_i \quad (18)$$

Notice that the a priori LLR information is intentionally not removed from the extrinsic LLR, because in this case it corresponds to the pilot bits which are known with absolute certainty if they exist.

The Max-Log-MAP is considered as an approximation to the Log-MAP algorithm [15] when the objective is to find soft or hard maximum a posteriori (MAP) estimates of X_i for $i = 0 \ldots N-1$. We chose to use the Max-Log-MAP instead of the Log-MAP for the following reasons:
1. We are interested in ML estimate of the sequence $\boldsymbol{X}$, in which case the Max-Log-MAP offers an optimal not an approximate solution [14].
2. Max-Log-MAP has less complexity, and is more convenient to implement in practice.
3. Max-Log-MAP is less sensitive to accuracy in σ^2 estimation, and hence subchannel gains estimation in our case, compared to Log-MAP [14].

2. Finding the ML estimate of $\boldsymbol{H}$ given soft estimate of $\boldsymbol{X}$:

The next step is to use the extrinsic ML soft estimate of $\boldsymbol{X}$, $\boldsymbol{L^{ext}}$, to find the ML estimate of $\boldsymbol{H}$, $\boldsymbol{H}^{ML}$. Strictly speaking the finite channel spread constraint should be incorporated in finding $\boldsymbol{H}^{ML}$. This could, however, be very difficult to do given the soft estimate of $\boldsymbol{X}$, $\boldsymbol{L^{ext}}$. For simplification here, we ignore the finite channel spread constraint in step 2, and then we incorporate it separately in step 3 using a least square approach. Here, $\boldsymbol{H}^{ML}$ is found as follows:

$$\boldsymbol{H}^{ML} = arg \max_{\boldsymbol{H}} \sum_{\boldsymbol{X}} p(\boldsymbol{X}|\boldsymbol{Y}) \quad (19)$$

Because of the orthogonality of subchannels, the above maximization can be evaluated on each subchannel individually.

$$\begin{aligned} H_i^{ML} &= arg \max_{H_i} \sum_{X_i} p(X_i|Y_i) & (20)\\ &= arg \max_{H_i} \sum_{X_i} p(X_i) p(Y_i|X_i) & (21)\\ &= arg \max_{H_i} \Big[p(X_i = -\sqrt{E_x}) \cdot e^{\frac{(Y_i+H_i)^2}{2\sigma^2}} \\ &+ p(X_i = +\sqrt{E_x}) \cdot e^{\frac{(Y_i-H_i)^2}{2\sigma^2}} \Big] & (22) \end{aligned}$$

where $p(X_i = -\sqrt{E_x}) = \frac{1}{1+e^{L_i^{ext}}}$, and $p(X_i = +\sqrt{E_x}) = \frac{e^{L_i^{ext}}}{1+e^{L_i^{ext}}}$.

In practice, one way to obtain an approximate solution to the maximization problem of the non-linear cost function in eq. 22 is to use numerical techniques. A simple technique is to use a one-dimensional grid and numerically evaluate the cost function at the grid points. Simulations show that a grid size of only 20 points is enough for low and medium SNR ranges, and this grid size was used to obtain the simulation results presented in this paper. By observing the fact that H_i^{ML} will always occur between 0 and Y_i if $p(X_i = \sqrt{E_x}) > p(X_i = -\sqrt{E_x})$ and between 0 and $-Y_i$ otherwise, the cost function then needs to be evaluated over half the range only by simply starting at the right end. In practice, this operation can be implemented using a lookup table with 10 entries.

3. Incorporating the finite channel spread constraint:

The obtained ML estimate of $\boldsymbol{H}$ can be enhanced by utilizing the constraint that impulse response of the channel has a known number $= L$ of non-zero taps. Since $\boldsymbol{H}$ has N points, and we want to use them to find the best estimate of L points of $\boldsymbol{h}$ in the time domain, this is an over-determined estimation problem. It can be solved by finding $\boldsymbol{h}^{LS}$ that satisfies the constraint $h_i^{LS} = 0$ for $i = L+1, \ldots, N-1$ and results in the least norm square distance $||\boldsymbol{H}^{ML} - \boldsymbol{Q}\boldsymbol{h}^{LS}||^2$. This is equivalent to finding the Best Linear Unbiased Estimator (BLUE) of $\boldsymbol{H}$ with finite channel spread constraint given $\boldsymbol{H}^{ML}$ as an initial estimate. Let $\underline{\boldsymbol{h}}^{LS}$ be a vector of length L representing the first L taps of $\boldsymbol{h}^{LS}$. $\underline{\boldsymbol{h}}^{LS}$ that minimizes $||\boldsymbol{H}^{ML} - \boldsymbol{V}\underline{\boldsymbol{h}}^{LS}||^2$ is given by:

$$\underline{\boldsymbol{h}}^{LS} = (\boldsymbol{V}^T\boldsymbol{V})^{-1}\boldsymbol{H}^{ML} \quad (23)$$

where $\boldsymbol{V}$ is $N \times L$ Vandermonde matrix that is given by:

$$V_{n,l} = e^{i\frac{2\pi}{N}nl} \quad (24)$$

where $n = 0, 1, \ldots, N-1$ and $l = 0, 1, \ldots, L-1$.

Therefore,

$$\boldsymbol{h}^{LS} = \left[\left[\begin{array}{c} \underline{\boldsymbol{h}}^{LS} \\ \boldsymbol{0} \end{array} \right]\right] \quad (25)$$

$$\boldsymbol{H}^{LS} = \boldsymbol{Q}\boldsymbol{h}^{LS} \quad (26)$$

where $\boldsymbol{0}$ is a zero vector of length $N - L$.

This completes a single iteration of the algorithm. The algorithm then reiterates by going back to step 1 and uses $\boldsymbol{H}^{LS}$ obtained in the current iteration as $\boldsymbol{H}^{init}$ for the next iteration. When the algorithm converges, $\boldsymbol{h}^{LS}$ of the last iteration is declared as the detected channel state $\boldsymbol{h}^{est}$. Hard estimate of $\boldsymbol{X}$, and hence hard decoded bits are obtained by thresholding $\boldsymbol{L^{app}}$ after convergence.

In practice, step 3 can be implemented, perhaps with some little loss in performance, using other simpler interpolation techniques like frequency domain filtering that would also effectively reduce the noise power by a factor of L/N [4], [16].

IV. Simulation and Comparison Results

The proposed iterative algorithm was simulated using a 4-state rate 1/2 CC with $G(D) = [1 \ \frac{1+D^2}{1+D+D^2}]$. In all of the following simulations we used $N = 128$, $L = 16$, and BPSK modulation in each of the subchannels. The actual channel taps were arbitrarily chosen as: [0.50.70.90.10.50.10.90.30.20.80.70.20.10.50.30.2]. The maximum number of iterations is set to 60 iterations.

It was found that simple codes not only result in less complexity, which grows exponentially with the number of states in the code, but also lead to better convergence behavior. This could be interpreted by the observation that a simple code offers a balance between providing some dependence between the subchannels enough to exploit frequency diversity and allow the system to bootstrap itself, but not too much to result in significant error propagation in the initial iterations when the channel estimate is very inaccurate, which can preclude convergence.

A typical convergence behavior of the system is shown in Figures 1, 2, 3, and 4 for the case of $SNR = E_x \cdot ||\boldsymbol{h}_{actual}||^2/\sigma^2 = 11dB$. The initial channel estimate was randomly selected. In practice, to conserve power, it is possible to stop iterations once convergence occurs. Convergence can be detected by monitoring the average a posteriori log-likelihood ratio $\bar{L}^{app}$ at the output of the soft decoder.

$$\bar{L}^{app} = \frac{1}{N}\sum_{i=0}^{N} L_i^{app} \quad (27)$$

$$\bar{L}^{ext} = \frac{1}{N}\sum_{i=0}^{N} L_i^{ext} \quad (28)$$

Fig. 1 shows plots of $\bar{L}^{app}$ and $\bar{L}^{ext}$ versus number of iterations. Fig. 2 shows a plot of the mean square error (MSE) in channel estimate $MSE = ||\boldsymbol{h}^{est} - \boldsymbol{h}^{actual}||^2$ versus number of iterations. Fig. 3 shows the number of decoded bit errors versus number of iterations. Examining Figures 1, 2, and 3 shows that there is a close correlation between the behavior of $\bar{L}^{app}$ and that of MSE and number or decoded bit errors. At convergence $\bar{L}^{app}$ becomes constant, or repeats in a periodic manner. Moreover, local peaks in $\bar{L}^{app}$ curve usually correspond to local valleys in MSE and number of decoded bit errors curves. In our simulations, convergence is detected by recording the latest four local peaks of $\bar{L}^{app}$, and the estimated channel states at the corresponding iterations. $\bar{L}^{app}(iter_p)$ is considered a local peak if $\bar{L}^{app}(iter_p - 1) \leq \bar{L}^{app}(iter_p) \leq \bar{L}^{app}(iter_p + 1)$. Convergence is declared once one of recoded local peaks recurs. At convergence the channel state corresponding to the maximum local peak among those between and including the recuring peaks is chosen as the detected channel state. Fig. 1 shows that in this case convergence has occurred at iteration 15. In the cases where convergence is not detected before the maximum number of iterations is reached, the channel state corresponding to the maximum local peak among the latest four is chosen as the detected channel state.

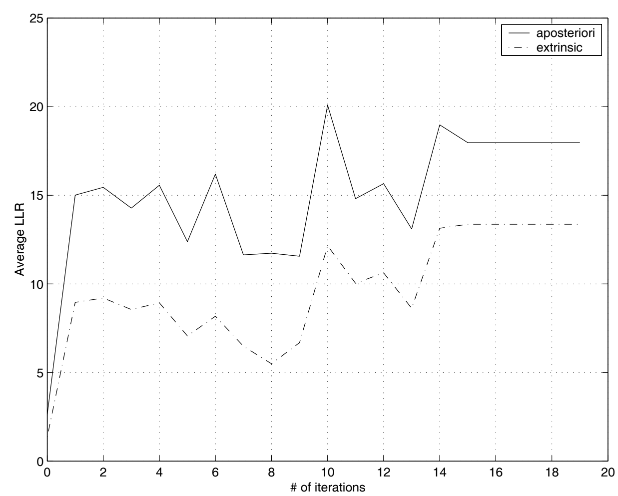

Fig. 1. Average a posteriori and extrinsic LLR's at the output of the soft decoder versus number of iterations

Fig. 4 shows the convergence behavior of selected three taps (2, 6, and 10) among the 16 taps of the channel impulse response.

The system performance using the proposed technique is compared to that of reference coded and uncoded systems that use the traditional technique of using L equally spaced pilots in the frequency domain to estimate the channel, and then, if necessary, perform soft decoding using the Max-Log-MAP algorithm.

To show the benefit of utilizing pilots not only in channel estimation, but also to assist in decoding, for the coded reference systems, we show results for both cases: when pilots are added after encoding where they can only be used for channel estimation, and when they are added before encoding where they could also assist in the soft decoding

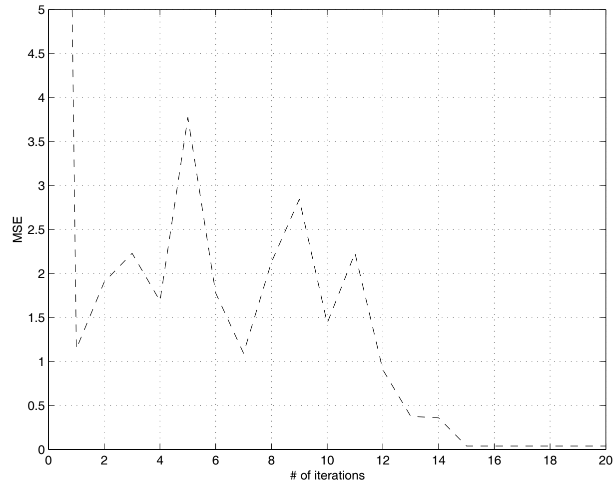

Fig. 2. Mean square error in channel estimation versus number of iterations

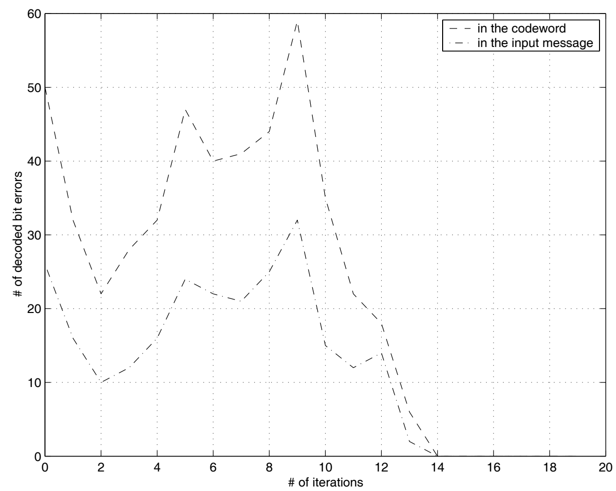

Fig. 3. Number of decoded bit errors versus number of iterations

process. The coded systems use the same CC as that used with the proposed technique.

Fig. 5 shows plots of Bit Error Rate (BER) versus E_b/N_0 for the proposed system and the reference coded and uncoded systems over the E_b/N_0 range of 8 to $15dB$. In all simulations the number of observed block errors was larger than 1000. Energy per information bit E_b instead of energy per symbol E_x is used in this comparison to appropriately account for the different redundancy overheads experienced by the different systems. Notice that the redundancy overhead caused by the cyclic prefix is common to all of the systems under comparison. For a given E_b/N_0, SNR for the different systems is obtained as follows:

$$SNR = 2 \cdot R_{eff} \cdot \frac{E_b}{N_0} \qquad (29)$$

where R_{eff} is equal to $\frac{N-L}{N+L}$ for the uncoded system with L pilots, $\frac{K-L}{N+L}$ for the coded system with L pilots added after encoding, $\frac{K-\frac{K}{N}L}{N+L}$ for the coded system with L pilots added before encoding, and $\frac{K}{N+L}$ for the coded system without pilots.

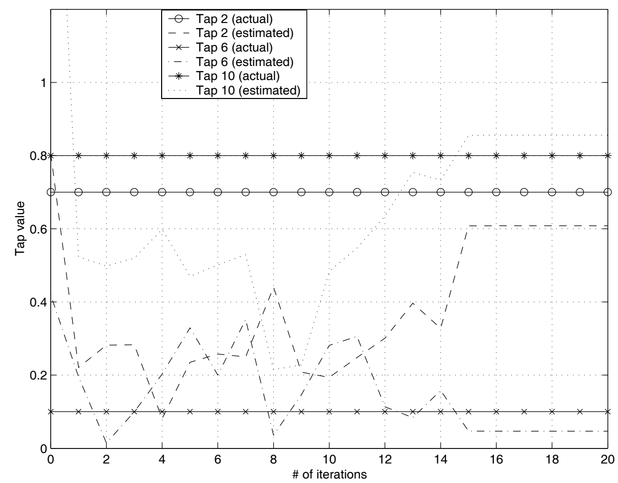

Fig. 4. Learning curves for selected channel taps

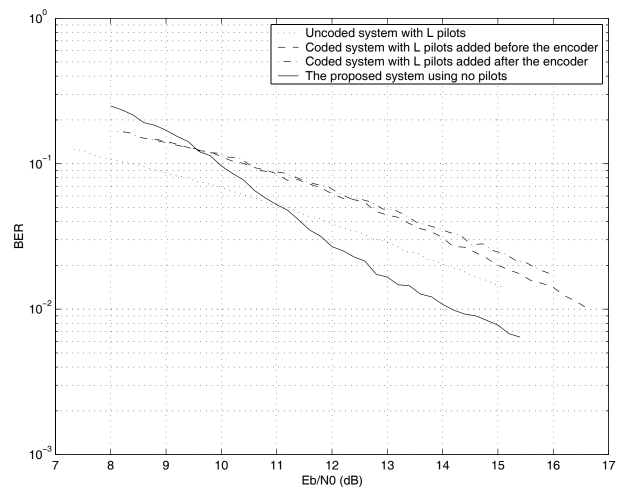

Fig. 5. Comparison of BER performance of the proposed system and reference systems

In the proposed algorithm, maximum number of iterations is set to 60 iterations, and the stopping criterion presented above is used to stop iterations once convergence is detected. The system that uses the proposed algorithm shows a better overall performance compared to the coded systems with L pilots for $E_b/N_0 > 9.7dB$. It shows a better overall performance compared to the uncoded system with L pilots for $E_b/N_0 > 11dB$. On the other hand, the other coded systems, start to show a coding gain relative to the uncoded system at a much higher value of E_b/N_0. The figure also shows that adding the pilots before encoding and using them in soft decoding results in better overall system performance.

Fig. 6 shows the average number of iterations required for convergence using the proposed technique versus E_b/N_0. As expected this number drops significantly as E_b/N_0 increases.

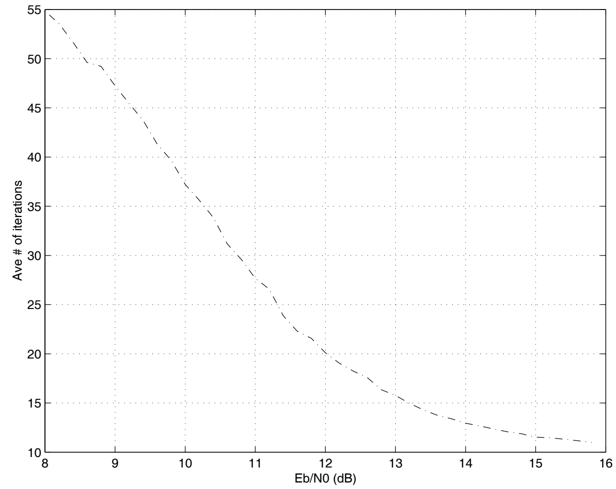

Fig. 6. Average number of iterations required for convergence versus Eb/N0

V. Conclusion

Coding is typically used in OFDM systems to exploit frequency diversity. This paper presented a new relatively low-complexity iterative technique for joint decoding and channel estimation. By using a simple 4-state rate 1/2 CC, it is possible to accurately estimate the channel state in OFDM systems within a single frame without using any pilots, and starting from an arbitrary initial state. The proposed technique was compared to coded and uncoded systems that use the traditional channel estimation technique that uses L pilots uniformly spaced in the frequency domain. It starts to offer a coding gain relative to coded and uncoded reference systems at $E_b/N_0 = 9.7dB$, and $11dB$, respectively. When pilots are used in a coded system it is better to add them before encoding so that they can not only assist in the channel estimation, but also in the soft decoding process.

The paper presented a stopping criterion based on monitoring the average a posteriori LLR at the output of the soft decoder to detect convergence of the proposed iterative algorithm and stop iterations. The number of iterations required for convergence depends on the operating SNR, and generally drops as the SNR increases.

References

[1] Z. Wang and G. B. Giannakis, “Wireless multicarrier communications: Where Fourier meets Shannon,” *IEEE Signal Proc. Mag.*, vol. 17, no. 3, pp. 29–48, May 2000.

[2] J. G. Proakis, *Digital Communications*, McGraw-Hill, 2001.

[3] S. K. Wilson, R. Khayata, and J. M. Cioffi, “16 QAM modulation with orthogonal frequency division multiplexing in a Rayleigh-fading environment,” in *Proc. IEEE Vehicular Tech. Conf.*, Stockholm, Sweden, Jun. 1994, vol. 3, pp. 1660–1664.

[4] V. Mignone and A. Morello, “CD3-OFDM: A novel demodulation scheme for fixed and mobile receivers,” *IEEE Trans. Commun.*, vol. 44, no. 9, pp. 1144–1151, Sep. 1994.

[5] B. Muquet, M. de Courville, P. Duhamel, and V. Buzenac, “A subspace based blind and semi-blind channel identification method for OFDM systems,” in *Proc. IEEE Workshop on Signal Proc. Advances in Wireless Comm.*, Annapolis, MD, May 1999, pp. 170–173.

[6] Jr. R. W. Heath and G. B. Giannakis, “Exploiting input cyclostationarity for blind channel indentification in OFDM systems,” *IEEE Trans. Signal Processing*, vol. 47, no. 3, pp. 848–956, Mar. 1999.

[7] X. Wang and K. J. Ray Liu, “Adaptive channel estimation using cyclic prefix in multicarrier modulation system,” *IEEE Commun. Lett.*, vol. 3, no. 10, pp. 291–293, Oct. 1999.

[8] S. Zhou and G. B. Giannakis, “Finite-alphabet based channel estimation for OFDM and related multicarrier systems,” *IEEE Trans. Commun.*, vol. 49, no. 8, pp. 1402–1414, Aug. 2001.

[9] J. F. Helard and B. Le Floch, “Trellis coded orthogonal frequency division multiplexing for digital video transmission,” in *Proc. IEEE GlobeCom*, Phoenix, AZ, Dec. 1991, vol. 2, pp. 785–791.

[10] C. Berrou, A. Glavieux, and P. Thitmajshima, “Near Shannon limit error-correction coding and decoding: turbo-codes,” in *Proc. IEEE Int. Conf. Commun.*, Geneva, Switzerland, May 1993, pp. 1064–1070.

[11] A. Picart, P. Didier, and A. Glavieux, “Turbo-detection: A new approach to combat channel frequency selectivity,” in *Proc. IEEE Int. Conf. Commun.*, Montreal, Canada, Jun. 1997, pp. 1498–1502, IEEE.

[12] D. J. C. MacKay, “Good error correcting codes based on very sparse matrices,” *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 399–431, Mar. 1999.

[13] X. Wang and H. V. Poor, “Iterative (turbo) soft interference cancellation and decoding for coded CDMA,” *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 1046–1061, Jul. 1999.

[14] K. Chugg, A. Anastasopoulos, and X. Chen, *Iterative Detection: Adaptivity, Complexity Reduction, and Applications*, Kluwer Academic Publishers, 2001.

[15] P. Robertson, E. Villebrun, and P. Hoeher, “A comparison of optimal and sub-optimal map decoding algorithms operating in the log domain,” in *Proc. IEEE Int. Conf. Commun.*, Seattle, WA, Jun. 1995, vol. 2, pp. 1009–1013.

[16] J. van de Beek, O. Edfors, M. Sandell, S. Wilson, and P. Borjesson, “On channel estimation in ofdm systems,” in *Proc. IEEE Vehicular Tech. Conf.*, Chicago, IL, Jul. 1995, vol. 2, pp. 815–819.