

Iterative Channel Estimation, Equalization, and Decoding for Pilot-Symbol Assisted Modulation Over Frequency Selective Fast Fading Channels

Mark F. Flanagan, *Member, IEEE*, and Anthony D. Fagan, *Member, IEEE*

Abstract—A pilot-based channel estimation scheme is proposed for frequency selective Rayleigh fading channels which works in conjunction with the existing paradigm of turbo equalization. The iterative nature of the channel estimation technique provides substantial gain over noniterative methods and makes it a suitable choice for iterative equalization and decoding. The channel estimator has low complexity, is decoupled from the equalizer soft-input soft-output module, and is capable of tracking significant channel variations within a codeword. The scheme is compatible with quadratic-amplitude modulation and with both parallel concatenated convolutional and low-density parity check coding. The proposed scheme provides an attractive low-complexity alternative to iterative receivers based on state-space models for channel parameter evolution. The use of pilot symbols is demonstrated to aid the equalizer both directly through channel estimation and indirectly through pilot message insertion.

Index Terms—Channel estimation, fast fading, frequency selective channels, pilot-symbol assisted modulation, turbo equalization.

I. INTRODUCTION

ITERATIVE joint equalization and decoding via the “turbo principle” was first proposed in [1] and has been shown to give remarkable performance over both static and time-varying channels, particularly when high-performing error correcting codes, such as parallel concatenated convolutional (PCC) and low-density parity check (LDPC) codes, are used. Many schemes, however, [2]–[4] assume perfect channel state information (CSI). This CSI is often difficult to obtain with great precision, particularly at the low SNRs at which PCC and LDPC codes operate. For a frequency selective channel with high Doppler at such SNRs, proper CSI estimation becomes very difficult, and imperfect estimation can significantly degrade bit error rate (BER).

When the fade rate is sufficiently slow, a separate training-aided channel estimator may be used to inform the equalizer soft-input soft-output (SISO) module. The advantage here is that classical signal processing techniques [e.g., LMS, recursive least squares (RLS), and Kalman filtering] may be used within the estimator, with savings in complexity. It has been shown in

[5] and [6] that feedback from the decoder can be successfully incorporated into this type of estimator; however, the limited tracking capability of these estimation algorithms precludes their use on channels with high Doppler such as those described here.

When the fade rate is high enough so that these schemes are no longer sufficient to acquire and track the time-varying frequency selective channel, the predominant design philosophy is that of the incorporation of channel estimation into an expanded trellis-based equalizer SISO module. An adaptive equalizer SISO module is presented in [7] which combines channel estimation into the branch metrics of a Bahl–Cocke–Jelinek–Raviv (BCJR)-type algorithm. Furthermore, this scheme is shown to incorporate, as forward-only cases, many SISO equalizer modules which incorporate channel estimation, e.g., the schemes of [8] and [9]. A shortcoming of this scheme (and its special cases) is the assumption that the unknown and time-varying channel parameters obey a Gauss–Markov model; thus, for a channel with fast fading governed by the Jakes model, a large state space would be required for adequate performance, leading to high complexity. Also, the Jakes model is not actually a finite-state model, nor is it Markov. Davis *et al.* [10] proposed a scheme where no such assumption on the channel parameters is made. In this formulation, each trellis state hypothesizes values for a certain number of previously transmitted symbols; the advantage of this is that, for each state, the minimum mean-square error (MMSE) p th-order linear prediction estimate of the channel filter output may be made on the basis of these hypothesized transmit symbols (this actually circumvents direct channel estimation). However, for spectrally efficient modulations and/or long channels, this method becomes infeasible as it involves expanding an already large state space by a factor exponential in the predictor order.

Valenti and Woerner [11] presented a pilot-aided scheme for BPSK transmission over fast flat-fading channels whereby, during turbo decoding, the channel could be reestimated at each iteration. In this contribution, we will show how their channel estimation scheme may be generalized to perform estimation of the frequency selective channel in a manner useful to joint iterative equalization and decoding. The proposed channel estimation scheme has the following desirable properties: It is of low complexity and decoupled from the equalizer SISO module, it does not rely on a state-space model for channel parameter evolution, it can successfully use feedback from the decoder SISO module, it can estimate frequency selective channels with

Manuscript received July 15, 2005; revised April 18, 2006 and June 27, 2006. This work was supported by Enterprise Ireland. The review of this paper was coordinated by Dr. M. Reed.

The authors are with the Department of Electronic and Electrical Engineering, University College Dublin, Dublin 6, Ireland (e-mail: mark.flanagan@ee.ucd.ie).

Digital Object Identifier 10.1109/TVT.2007.897215



Fig. 1. Transmitter architecture.

high fade rates, it is capable of tracking significant channel variations within a single codeword, and it is compatible with any 2-D memoryless modulation technique. We note that our system uses joint coding and modulation via bit-interleaved coded modulation (BICM) with iterative detection, which has been shown to outperform trellis-coded modulation on fading channels [12]–[14]; however, the concern of this paper is that of the incorporation of channel estimation into the iterative receiver for BICM rather than the design of joint coding and modulation.

This paper extends the work first presented in [15] to higher order modulations (equivalently to BICM), to channels with arbitrary power profile, and to LDPC and PCC coding. Reference [16] presents an alternative channel estimator for a similar transmission scheme to [15]; their estimator uses feedback from PCC decoding to approximately decouple the L -tap frequency selective channel into L flat fading subchannels. The drawback of this method is that reliable feedback is required for efficient decoupling; also, [16] only considers BPSK modulation. A scheme with a similar design philosophy was presented in [17], where a decoupled MMSE-criterion-based channel estimator for frequency selective channels was shown to be successful in joint multiuser detection and single-user decoding for pilot symbol aided direct-sequence code division multiple access. Accurate channel estimation was achieved even in a highly loaded system, for fade rates significantly below those considered in this paper.

Simulation results are presented for both PCC and LDPC coding, for two reasons. First, it is desirable to illustrate the efficacy of the entire scheme on different types of code which utilize iterative decoding. Second, the practical issue of communication between a mobile station (MS) and a base station (BS) is considered. The message-passing (MP) algorithm in the LDPC section of the factor graph operates in a completely parallel fashion, in stark contrast to the serial BCJR algorithm, which operates in the equalizer section and in the sections corresponding to both constituent recursive systematic convolutional (RSC) codes in the PCC case. Thus, we suggest the use of the PCC encoder and LDPC decoder in the MS and the LDPC encoder and PCC decoder in the BS.

The use of pilot symbols is demonstrated to aid the equalizer SISO module through the periodic injection of “pilot symbol messages” into the equalizer’s trellis, in addition to aiding the channel estimator module through the provision of a known transmit subsequence. This additional aid is marginal; however, it indicates that training/pilot symbols may aid the dual functions of channel estimation and turbo equalization. This is a property not currently exploited in systems which use training/pilots to acquire the channel. This point is explained from a factor graph perspective in [18].

This paper is organized as follows. In Section II, the models for the transmitter and frequency selective channel are pre-

sented, and the function of each block in the receiver is outlined. In Section III, the proposed channel estimator is described in detail, and its position and function in the turbo equalization scheme are explained. Section IV presents simulation results, and Section V gives the conclusion.

II. SYSTEM MODEL

A. Transmitter

Fig. 1 shows the discrete-time model of the transmitter. N information bits $\{u_k\}$, $k = 1, 2, \dots, N$ are encoded using a rate r code. The code may be any whose decoding is by MP in a factor graph, and hence, the code shall be referred to as an MP code. The two most important and widely used MP codes are PCC and LDPC codes; both shall be considered in this paper. The N/r encoded bits $\{c_k\}$ are then passed to an $R \times S$ block channel interleaver Π_C . The purpose of the channel interleaver is to combat burst errors caused by the fading channel. This interleaving is not required for the LDPC code, since the random structure of the parity check matrix is equivalent to an “in-built” interleaver [19]; the interleaver can be considered to be the identity map in this case. The resulting bit sequence $\{y_k\}$ is parsed into l -bit blocks $\mathbf{y}_k = [y_{(k-1)l+1} y_{(k-1)l+2} \dots y_{kl}]$, and these are modulated to a symbol sequence $\{m_k\}$ via $m_k = \mathcal{M}(\mathbf{y}_k)$. Memoryless QAM modulation is assumed, with a constellation $S \subset \mathbb{C}$ normalized to have unit energy, i.e., $E_{m \in S} \{|m|^2\} = 1$. Note that if the modulator mapping in the case of BPSK is written $\mathcal{M}(\mathbf{y}_k) = \mathcal{M}_B(y_k)$ with $S = \{-1, +1\}$, then the mapping for Gray-coded QPSK may be written $\mathcal{M}([y_{2k-1} y_{2k}]) = (\mathcal{M}_B(y_{2k-1}) + j\mathcal{M}_B(y_{2k}))/\sqrt{2}$. The modulated symbol sequence $\{m_k\}$ is then parsed into blocks of N_P bits (N_P is called the pilot block spacing), and a contiguous group of $2L - 1$ known symbols $\{p_1, p_2, \dots, p_{2L-1}\}$ is inserted into the center of each block (N_P is assumed even). We refer to these known symbols as “pilot symbols” although they could be regarded as a “distributed training sequence.” The pilots are chosen from the same (QAM) constellation S as the information-bearing symbols.

B. Channel

We consider transmission over an L -tap frequency selective Rayleigh fading channel. The receive signal is given by

$$r_k = \sum_{j=0}^{L-1} h_k^{(j)} x_{k-j} + n_k$$

where each channel tap $h_k^{(j)}$ is a complex Gaussian random variable with zero mean, and the different taps are assumed to be independent. The real and imaginary parts of $h_k^{(j)}$ are

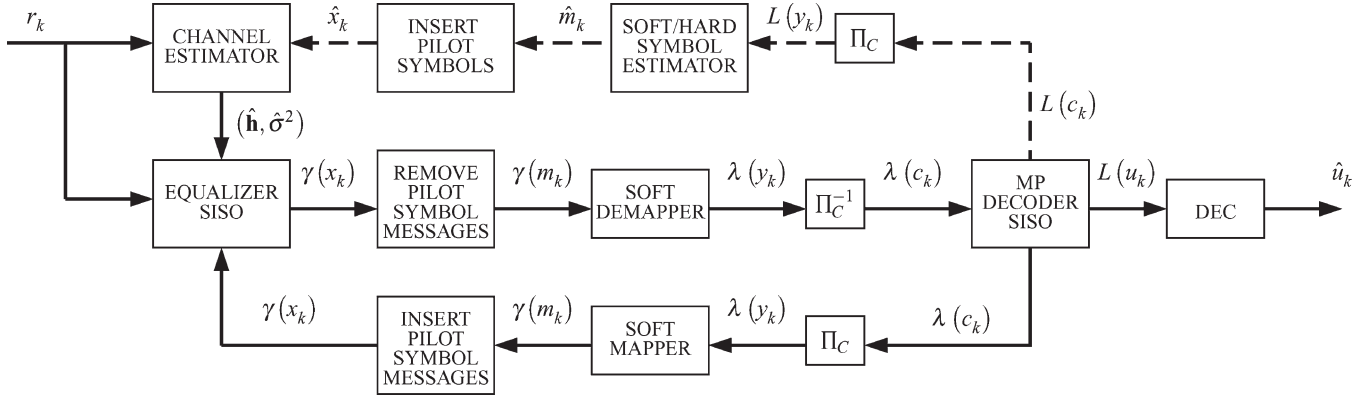


Fig. 2. Receiver architecture. In this figure, γ stands for extrinsic message, λ for extrinsic LLR (LLR corresponding to extrinsic message for a binary variable), and L for actual LLR (LLR corresponding to function summary for a binary variable).

independent, and both have autocorrelation function given by the Jakes model [20]

$$R_h^{(j)}(i) = \frac{\sigma_j^2}{2} J_0(2\pi f_d T_s i)$$

where f_d is the maximum Doppler frequency, T_s is the symbol period, and J_0 is the zero-order Bessel function of the first kind. Note that σ_j^2 corresponds to the power of the j th channel tap, and therefore, $\{\sigma_j^2 | j = 0, 1, \dots, L-1\}$ gives the power profile of the channel. The overall channel power gain is normalized to unity, i.e.,

$$\sum_{j=0}^{L-1} E \left\{ |h_k^{(j)}|^2 \right\} = \sum_{j=0}^{L-1} \sigma_j^2 = 1.$$

The additive noise sequence n_k is white, each element being a complex-valued Gaussian random variable with zero mean and variance σ^2 in each of the real and imaginary directions. The SNR per bit is then $E_b/N_0 = (N_P + 2L - 1)/(2N_P l r \sigma^2)$ with pilot-symbol insertion and $E_b/N_0 = 1/(2l r \sigma^2)$ without.

C. Receiver

If it is assumed that the channel tap values and variance of the additive noise are known exactly throughout the entire data frame, then the iterative receiver may be derived in a straightforward manner using factor graph theory [21]. This derivation will not be described in detail here; see [18] for details of the derivation for the case of PCC coding and BPSK modulation. Extension of this derivation to LDPC coding and to higher order modulation is straightforward. The resulting receiver architecture is shown in Fig. 2 (assuming perfect channel estimation and ignoring for the moment in the section in dotted lines). In this figure, γ stands for extrinsic message, λ for extrinsic log-likelihood ratio (LLR) (LLR corresponding to extrinsic message for a binary variable), and L for actual LLR (LLR corresponding to function summary for a binary variable). The equalizer and MP decoder SISO modules exchange extrinsic messages $\{\gamma(x_k)\}$ on the transmit symbols and extrinsic LLRs $\{\lambda(c_k)\}$ on the MP code bits, in the manner of Douillard's original turbo equalization scheme [1]. It is interesting to note

that, in contrast to the scheme of Valenti and Woerner [11], the pilot symbols appear as variables in the factor graph (this is shown in [18]). This results in two effects. First, pilot-symbol messages are removed at the equalizer SISO output; an intuitive explanation for this is that the information corresponding to pilot symbols is of no relevance to the MP decoder. Second, pendant factor nodes attached to the pilot variable nodes in the factor graph result in the insertion of pilot-symbol messages in the feedback path, as shown in Fig. 2. The inserted log-domain message corresponding to the pilot symbol q (with known value p) is ideally

$$\gamma(q) = \begin{cases} 0, & \text{if } q = p \\ -\infty, & \text{otherwise.} \end{cases}$$

Hence, in this scheme, the pilot symbols are capable of aiding the equalizer SISO module directly, as well as indirectly through improved channel estimation.

Between these two main SISO modules are two other SISO modules, here called a soft mapper and soft demapper, which perform bit-to-symbol and symbol-to-bit extrinsic message mappings. These modules are both special cases (where one output is suppressed) of the soft-output maximum *a posteriori* (SOMAP) module described in [22].

Of course, perfect channel estimates are not available in practice, and thus, the equalizer SISO requires estimates of the time-varying complex channel taps and the noise variance in order to form its branch metrics. In factor graph terms, the messages passed from a certain subset F of its pendant factor nodes require knowledge of the channel, for which we substitute channel estimates. These estimates are provided by a channel estimator described in Section III. However, channel estimates derived only from pilot symbol information are generally poor since they ignore the code structure. It is possible to improve the system performance considerably by reestimating the channel after each (equalizer SISO, MP decoder SISO) iteration. This is achieved by including the feedback path shown in dotted lines in Fig. 2. The LLRs $\{L(c_k)\}$ for the MP encoded bits are interleaved and used as priors for a symbol estimator. Hard symbol estimation may be performed by simply making decisions on each bit and modulating the result, i.e., $\hat{m}_k = \mathcal{M}(\hat{y}_k)$, where $\hat{y}_k = [\hat{y}_{(k-1)l+1} \hat{y}_{(k-1)l+2} \dots \hat{y}_{kl}]$, and each $\hat{y}_r = \text{sign}(L(y_r))$.

The soft symbol estimator uses as symbol estimate the expected value of the symbol given the symbol priors, i.e.,

$$\hat{m}_k = \sum_{\mathbf{y}_k \in [GF(2)]^l} \mathcal{M}(\mathbf{y}_k) P(\mathbf{y}_k). \quad (1)$$

Here, the symbol priors are $P(\mathbf{y}_k) = P(y_{(k-1)l+1}) \times P(y_{(k-1)l+2}) \dots P(y_{kl})$, where each bit prior is related to the corresponding LLR via $L(y_r) = \log(P(y_r = 1)/P(y_r = 0))$. Equation (1) simplifies in the case of BPSK to

$$\hat{m}_k = \begin{cases} \text{sign}(L(y_k)), & \text{for hard symbol estimation} \\ \tanh\left(\frac{L(y_k)}{2}\right), & \text{for soft symbol estimation} \end{cases} \quad (2)$$

and in the case of Gray-coded QPSK to

$$\hat{m}_k = [\text{sign}(L(y_{2k-1})) + j\text{sign}(L(y_{2k}))] / \sqrt{2} \quad (3)$$

for hard symbol estimation, and

$$\hat{m}_k = \left[\tanh\left(\frac{L(y_{2k-1})}{2}\right) + j \tanh\left(\frac{L(y_{2k})}{2}\right) \right] / \sqrt{2} \quad (4)$$

for soft symbol estimation. The original pilot symbols are inserted into this symbol estimate stream, and the resulting estimated transmit stream $\{\hat{x}_k\}$ is used by the channel estimator to produce a refined channel estimate. This improved estimate aids the equalizer SISO module in the next iteration. In factor graph terms, this amounts to resending the messages from all nodes in F [18].

III. CHANNEL ESTIMATOR

A. Known Transmit Sequence

Here, a generalization of Valenti and Woerner's channel estimator [11] is proposed; the generalization is to higher order modulations and to frequency selective channels. Assuming the transmit sequence $\{x_k\}$ is known at the receiver, then the correlation sequence $q_k^{(j)} = r_k x_{k-j}^*$ may be formed for each $j = 0, 1, \dots, L-1$. The autocorrelation function of the sequence $\{q_k^{(j)}\}$ is

$$\begin{aligned} R_q^{(j)}(i) &= E\{q_k^{(j)} q_{k-i}^{(j)*}\} \\ &= \sum_{r=0}^{L-1} \sum_{s=0}^{L-1} E\{h_k^{(r)} h_{k-i}^{(s)*}\} E\{x_{k-r} x_{k-j}^* x_{k-i-s}^* x_{k-i-j}\} \\ &\quad + E\{n_k n_{k-i}^*\} E\{x_{k-j}^* x_{k-i-j}\} \\ &= \sum_{r=0}^{L-1} 2R_h^{(r)}(i) E\{x_{k-r} x_{k-j}^* x_{k-i-r}^* x_{k-i-j}\} + 2\sigma^2 \delta(i) \\ &= 2R_h^{(j)}(i) + (1 - \sigma_j^2 + 2\sigma^2) \delta(i). \end{aligned}$$

Here, we have assumed that the fourth-order moment

$$E\{x_k x_p^* x_q^* x_r\} = 0 \quad \text{for } k \notin \{p, q, r\}. \quad (5)$$

This condition will approximately hold for modulated MP codewords in general; it is also a property to be taken into account when designing the pilot sequence.

The cross correlation between the sequences $\{q_k^{(j)}\}$ and $\{h_k^{(j)}\}$ is

$$v^{(j)}(i) = E\{q_{k-i}^{(j)} h_k^{(j)*}\} = 2R_h^{(j)}(i).$$

The sequence $\{q_k^{(j)}\}$ is filtered by a length- K finite-impulse response filter to produce the optimum MMSE estimate of the j th channel tap (K is assumed odd). This proceeds according to [23]

$$\hat{h}_k^{(j)} = \sum_{i=-\lfloor \frac{K}{2} \rfloor}^{\lfloor \frac{K}{2} \rfloor} w_i^{(j)} q_{k-i}^{(j)} \quad \text{for } j = 0, 1, \dots, L-1 \quad (6)$$

where $\mathbf{w}^{(j)} = (w_{-\lfloor K/2 \rfloor}^{(j)} \dots w_{\lfloor K/2 \rfloor}^{(j)})^T$ is given by the Wiener solution

$$\left[\mathbf{R}_j + \left(\frac{1 - \sigma_j^2}{2} + \sigma^2 \right) \mathbf{I} \right] \mathbf{w}^{(j)} = \mathbf{c}_j.$$

Here, \mathbf{R}_j is a matrix whose (m, n) entry is $R_h^{(j)}(m-n)$, and $\mathbf{c}_j = (R_h^{(j)}(-\lfloor K/2 \rfloor) \dots R_h^{(j)}(\lfloor K/2 \rfloor))^T$. Thus, the channel estimator consists of a bank of Wiener filters, one for each channel tap. Each Wiener filter has real coefficients. The MMSE for Wiener filter $\mathbf{w}^{(j)}$ is given by

$$\text{MMSE}_j = \sigma_j^2 - \mathbf{c}_j^T \left[\mathbf{R}_j + \left(\frac{1 - \sigma_j^2}{2} + \sigma^2 \right) \mathbf{I} \right]^{-1} \mathbf{c}_j. \quad (7)$$

Once the channel taps have been estimated, the noise variance may be estimated simply by taking the sample variance of

$$z_k = r_k - \sum_{j=0}^{L-1} \hat{h}_k^{(j)} x_{k-j}. \quad (8)$$

If the fade rate is slow enough ($f_d T_s \ll 1$) and the channel estimation filter length K is small enough, then the coefficients $w_i^{(j)}$ are all approximately equal and are

$$w_i^{(j)} \approx \frac{\sigma_j^2}{(1 + \sigma^2) + (K-1)\sigma_j^2} \quad (9)$$

for each $i = -\lfloor K/2 \rfloor, \dots, \lfloor K/2 \rfloor$. We may also make the further approximation $w_i^{(j)} \approx 1/(K-1)$ if $\sigma_j^2 \gg (1 + 2\sigma^2)/(K-1)$. This means that for slow fade rates, the Wiener filters may be replaced by moving average filters which do not require knowledge of the autocorrelation function of the channel taps (or knowledge of the noise variance, if the channel tap power is high enough). Note that in the case of BPSK modulation and $L = 1$, this analysis collapses to that of Valenti and Woerner [11].

B. Unknown Transmit Sequence

Initially, the receiver has no knowledge of the transmit symbol sequence other than the pilot symbol values. After the first iteration, estimates of the transmit sequence are available via feedback from the MP decoder. In both cases, it is possible to use a modified version of the scheme described previously to perform channel estimation.

1) *Initial Channel Estimation:* We consider a hypothetical transmit sequence $\{\bar{x}_k\}$ in which the nonpilot symbols are made up of a sequence of symbols in $\{-1, +1\}$ which satisfy the fourth-order moment condition (5). Consider the first pilot-assisted block for ease of exposition. We shall assume that the channel taps are reasonably static over the length of this block, i.e.,

$$h_k^{(j)} \approx h^{(j)} \quad k \in \{1, 2, \dots, N_P + 2L - 1\}.$$

The received symbols corresponding to the pilots are

$$\mathbf{p} = \mathbf{A}\mathbf{h} + \mathbf{n}$$

where

$$\begin{aligned} \mathbf{p} &= (r_{(N_P/2)+L} \quad r_{(N_P/2)+L+1} \quad \dots \quad r_{(N_P/2)+2L-1})^T \\ \mathbf{h} &= (h^{(0)} \quad h^{(1)} \quad \dots \quad h^{(L-1)})^T \\ \mathbf{n} &= (n_{(N_P/2)+L} \quad n_{(N_P/2)+L+1} \quad \dots \quad n_{(N_P/2)+2L-1})^T \end{aligned}$$

and

$$\mathbf{A} = \begin{pmatrix} p_L & p_{L-1} & \dots & p_1 \\ p_{L+1} & p_L & \dots & p_2 \\ \vdots & \vdots & & \vdots \\ p_{2L-1} & p_{2L-2} & \dots & p_L \end{pmatrix}. \quad (10)$$

Then, an estimate of the receive sequence for this block is

$$\bar{\mathbf{r}} = \mathbf{V}\mathbf{A}^{-1}\mathbf{p} \quad (11)$$

where

$$\mathbf{V} = \begin{pmatrix} \bar{x}_1 & 0 & 0 & \dots & 0 \\ \bar{x}_2 & \bar{x}_1 & 0 & \dots & 0 \\ \bar{x}_3 & \bar{x}_2 & \bar{x}_1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \bar{x}_{N_P+2L-1} & \bar{x}_{N_P+2L} & \dots & \bar{x}_{N_P+L} \end{pmatrix}$$

and the pilots have been chosen to make \mathbf{A} invertible. Note that in practice, the sequence $\{\bar{x}_k\}$ is generated in the receiver by a maximal-length pseudonoise sequence generator. The receive sequence estimate (11) is computed for each pilot-assisted block. Initial estimation of the channel taps then proceeds according to

$$\hat{h}_k^{(j)} = \sum_{i=-\lfloor \frac{K}{2} \rfloor}^{\lfloor \frac{K}{2} \rfloor} w_i^{(j)} \bar{r}_{k-i} \bar{x}_{k-i-j}, \quad \text{for } j = 0, 1, \dots, L - 1$$

and subsequently, the noise variance is estimated by taking the sample variance of

$$z_k = \bar{r}_k - \sum_{j=0}^{L-1} \hat{h}_k^{(j)} \bar{x}_{k-j}.$$

It is easily seen that this scheme necessitates $2L - 1$ pilots per block in order to work.

2) *Subsequent Channel Estimation:* After the first iteration, channel estimation may proceed using the method outlined in part III-A of this section, the actual transmit sequence $\{x_k\}$ being replaced by the estimate of the transmit sequence $\{\hat{x}_k\}$ provided by the MP decoder.

IV. SIMULATION RESULTS

A. BER Performance of the System

In this section, we present simulation results which illustrate the BER performance of the system. Four different simulation environments are tested, in order to demonstrate the performance of the entire scheme under variation of coding scheme, modulation type, channel length, power profile, and fade rate. Within each simulation environment, five different schemes were tested: 1) perfect knowledge of the channel and no transmission of pilot symbols; 2) perfect knowledge of the channel and transmission of pilot symbols; 3) channel estimation with no feedback from MP decoder to the channel estimator; 4) hard decision feedback from decoder to channel estimator; and 5) soft decision feedback from decoder to channel estimator.

The PCC code used in simulations consisted of two rate 1/2 RSC encoders with constraint length 4 and feedback and feed-forward generators in octal notation 15_o and 17_o , respectively. The trellis of the upper encoder was terminated with $m = 3$ tail bits, while the trellis of the lower encoder was left open. A 1250-b S-random interleaver was used within the PCC code, with $S = 20$ [24]. Half of the parity outputs of the PCC encoder were punctured to give an overall rate of $r = N/[2(N + m)] \approx 1/2$. The channel interleaver used with the PCC code was a 50×50 block interleaver. For all PCC coded system simulations, 12 iterations of channel estimation, equalization, and decoding were performed.

The LDPC code used in simulations was the Margulis code with $p = 11$; this is a (3, 6)-regular Gallager code based on a Cayley graph construction [25]. It is a rate 1/2 code with codeword length 2640. In the case of LDPC coding, simulations using a 55×48 block interleaver Π_C showed little or no performance improvement over those with no interleaver; these results are omitted. For all LDPC coded system simulations, iterative channel estimation, equalization, and decoding continued until either a valid codeword was detected (via syndrome check), or 20 iterations were completed. Before presenting simulation results, we note the interesting result that in all simulations involving LDPC coding, all frame errors were detected errors.

The pilot sequences we used in simulations were in each case multiples of binary sequences from [26]. These were found by exhaustive computer search for binary sequences with good autocorrelation properties. The pilot sequences we used,

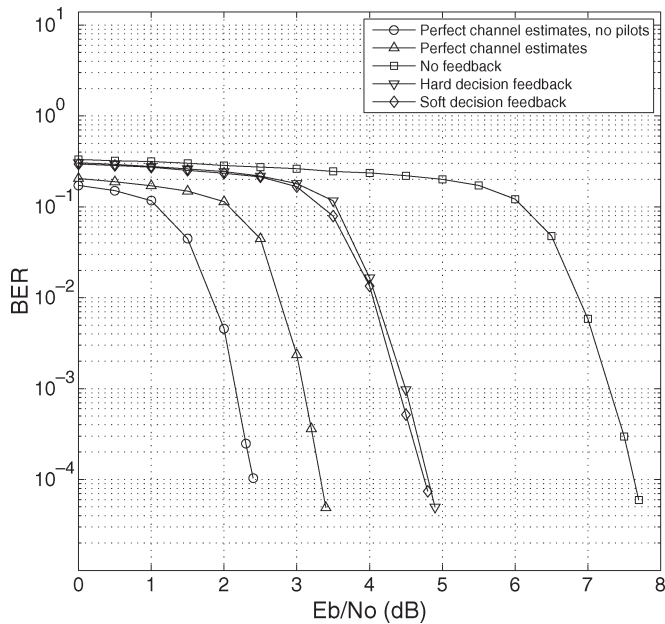


Fig. 3. BER performance of iterative channel estimation, equalization, and decoding over a frequency selective fading channel. The plot is for a channel with three taps of equal average power. A PCC code is used with feedback and feedforward generators 15_o and 17_o, respectively, the code is punctured to rate approximately 0.5. BPSK modulation is used, and the normalized fade rate is $f_d T_s = 0.005$.

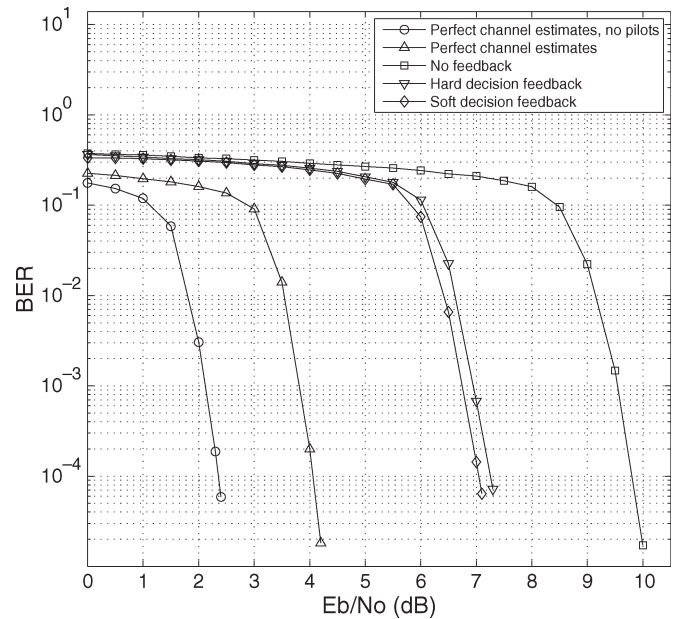


Fig. 4. BER performance of iterative channel estimation, equalization, and decoding over a frequency selective fading channel. The plot is for a channel with three taps of equal average power. A PCC code is used with feedback and feedforward generators 15_o and 17_o, respectively, the code is punctured to rate approximately 0.5. BPSK modulation is used, and the normalized fade rate is $f_d T_s = 0.02$.

which have lengths 5 and 9, also have the property that \mathbf{A}^{-1} consists simply of the elements $-0.5, 0.5,$ and 0 [see (10) and (11)], making the algorithm of Section III-B1 less complex to implement.

1) Three-Tap Channel With PCC Coding and BPSK Modulation: Here, we present simulation results for the case of PCC coded transmission over a channel with $L = 3$ taps of equal average power ($\sigma_j^2 = 1/3 \forall j$). BPSK modulation was assumed, and (2) was used for soft symbol estimation. The pilot sequence used was $\{+1, -1, +1, +1, -1\}$, and the channel estimation filter length was $K = 75$. These five different schemes were compared for the cases of two normalized fade rates: $f_d T_s = 0.005$ and $f_d T_s = 0.02$. A pilot block spacing of $N_P = 20$ was used for the slower fade rate, and the moving average approximation to the optimum Wiener filter was used, while for the faster fade rate, we chose $N_P = 10$, and the Wiener solution was required for adequate performance. In each case, 50 independent frame errors were observed for each point to determine BER.

The BER performance of the five different schemes is shown in Fig. 3 for the slower fade rate $f_d T_s = 0.005$ and in Fig. 4 for the faster fade rate $f_d T_s = 0.02$. The transmission of pilot symbols causes a loss in energy efficiency of $10 \log_{10}((N_P + 2L - 1)/N_P)$ dB, i.e., 0.97 dB for the slower fade rate and 1.76 dB for the faster fade rate. However, the performance degradation observed in the simulations was only 0.92 dB for the slower fade rate and 1.70 dB for the faster fade rate at a BER of 10^{-4} . This is because the inserted messages corresponding to the known pilot symbols aid the equalizer SISO in determining new extrinsic information.

In the case of channel estimation with no feedback from the PCC decoder to the estimator, there is a performance

degradation of 4.31 dB for the slower fade rate and 5.74 dB for the faster fade rate over the performance for perfect channel estimates at a BER of 10^{-4} . This degradation can be lessened with the inclusion of feedback from PCC decoder to estimator. Hard decision feedback allows us to gain back 2.83 dB for the slower fade rate and 2.64 dB for the faster fade rate at a BER of 10^{-4} . Soft decision feedback gives an extra 0.05 dB for the slower fade rate and 0.12 dB for the fast fade rate (at a BER of 10^{-4}). For both fade rates, the performance improvement when feedback is incorporated into the channel estimator is approximately the same, as reported in [11]. Soft decision feedback is observed to give only marginal improvement in performance over hard decision feedback. This is a phenomenon reported by other authors [27]. Therefore, we advocate the use of hard decision feedback in practice.

2) Three-Tap Channel With LDPC Coding and QPSK Modulation: Fig. 5 presents simulation results for the case of LDPC coded transmission over a channel with $L = 3$ taps of equal average power ($\sigma_j^2 = 1/3 \forall j$). Gray-coded QPSK modulation was assumed, and (4) was used for soft symbol estimation. The pilot sequence used was $((1 + j)/\sqrt{2})^* \{+1, -1, +1, +1, -1\}$, and the channel estimation filter length was $K = 75$. The normalized fade rate was $f_d T_s = 0.02$. Again, we chose a pilot block spacing of $N_P = 10$, and the Wiener solution was required for adequate performance. In each case, 80 independent frame errors were observed for each point to determine BER.

The BER performance of the five different schemes is shown in Fig. 5. In the case of channel estimation with no feedback from the LDPC decoder to the channel estimator, there is a performance degradation of 6.00 dB over the performance for perfect channel estimates at a BER of 10^{-4} .

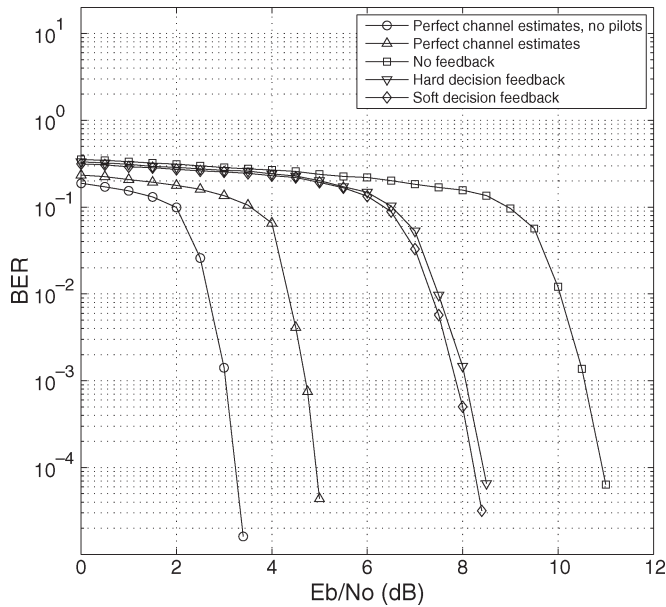


Fig. 5. BER performance of iterative channel estimation, equalization, and decoding over a frequency selective fading channel. The plot is for a channel with three taps of equal average power. The (1320, 2640) Margulis LDPC code is used in conjunction with Gray-coded QPSK modulation. The normalized fade rate is $f_d T_s = 0.02$.

Hard decision feedback allows us to gain back 2.50 dB at a BER of 10^{-4} , and soft decision feedback gives an extra 0.20 dB at this BER. These gains in power efficiency are similar to those for BPSK modulation; however, with QPSK modulation, an additional gain in spectral efficiency has been achieved.

3) *Five-Tap Channel With LDPC Coding and QPSK Modulation:* Fig. 6 presents simulation results for the case of LDPC coded transmission over a channel with $L = 5$ taps and the exponentially decaying power profile $\sigma_j^2 = (a - 1)a^j / (a^L - 1)$ for $j = 0, 1, \dots, L - 1$. The power-profile decay factor was chosen to be $a = 0.7$. Gray-coded QPSK modulation was assumed, and (4) was used for soft symbol estimation. The pilot sequence used was $((1 + j)/\sqrt{2}) * [+1, +1, +1, +1, -1, -1, +1, -1, +1]$. The channel estimation filter length was $K = 75$, and the normalized fade rate was $f_d T_s = 0.005$. The pilot block spacing was chosen to be $N_P = 20$, and the moving average approximation to the optimum Wiener filter was used. In each case, 100 independent frame errors were observed for each point to determine BER.

It is shown in Fig. 6 that the diversity provided by the extra channel taps means that the system exhibits high performance in the case of perfect channel estimation, both with and without pilot transmission; however, the larger number of taps also makes the task of channel estimation more difficult. Therefore, the cost of imperfect channel estimation is high (> 10.7 dB over the BER range of interest). Iterative channel estimation with hard decision feedback allows us to gain back a significant portion of this loss (> 5.2 dB over the BER range of interest). Note also that an error floor is observed in the curve for feedback-free channel estimation. This is due to near-codewords of the Margulis code [28]. An error floor due to these near-codewords will eventually manifest itself in all simulations;

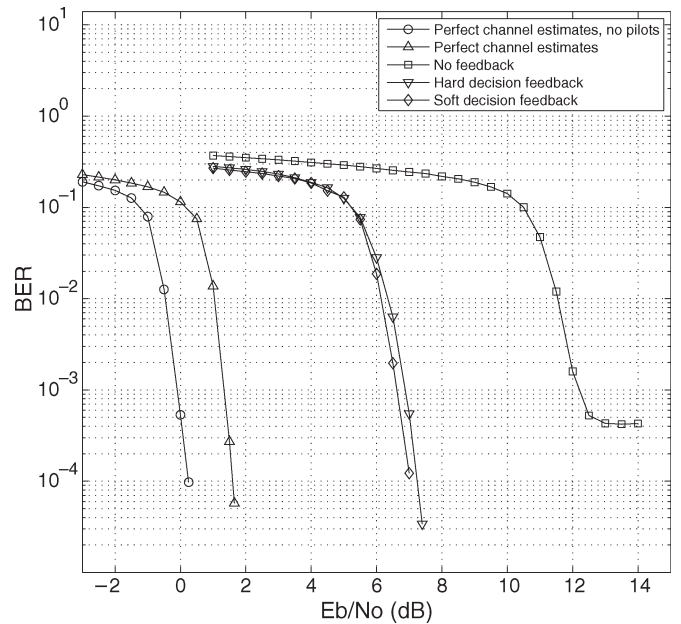


Fig. 6. BER performance of iterative channel estimation, equalization, and decoding over a frequency selective fading channel. The plot is for a five-tap channel with an exponentially decaying power profile (decay factor 0.7). The (1320, 2640) Margulis LDPC code is used in conjunction with Gray-coded QPSK modulation. The normalized fade rate is $f_d T_s = 0.005$.

however, in the case of no feedback, the channel estimation error acts as an unmitigated noise term on the equalizer SISO branch metrics, which raises this error floor. This effect is more noticeable for the longer channel as the equalizer branch metric error contains contributions from all of the channel tap errors (cf. [18, eq. (3)]). Note that the curves for the schemes which use iterative channel estimation evince no error floor down to a BER of 10^{-4} .

4) *Three-Tap Channel With LDPC Coding and 16-QAM Modulation:* Fig. 7 presents simulation results for the case of LDPC coded transmission over a channel with $L = 3$ taps of equal average power ($\sigma_j^2 = 1/3 \forall j$). Gray-coded 16-QAM modulation was assumed, and the general form of (1) was used for soft symbol estimation. The pilot sequence used was $((3 + j)/\sqrt{10}) * \{+1, -1, +1, +1, -1\}$; thus, the pilots maintain the average power of the transmitted sequence. The channel estimation filter length was $K = 75$, and the normalized fade rate was $f_d T_s = 0.005$. The pilot block spacing was chosen to be $N_P = 20$, and the moving average approximation to the optimum Wiener filter was used. In each case, 100 independent frame errors were observed for each point to determine BER.

As shown in Fig. 7, the much increased spectral efficiency of the 16-QAM system comes at the cost of much increased difficulty of channel estimation; the BER curves in the case of imperfect channel estimation are no longer approximately parallel to those with perfect channel estimates. The reason for this behavior is as follows. Although the channel estimation MSE in this case is not significantly higher than in the case of BPSK, the constellation minimum distance is much reduced (by a factor of $\sqrt{10}$), and so, while perfect channel estimation exhibits steep BER curves for 16-QAM, the presence of channel estimation error causes the soft demapper output to improve only gradually with increasing E_b/N_0 . However, a large performance gain

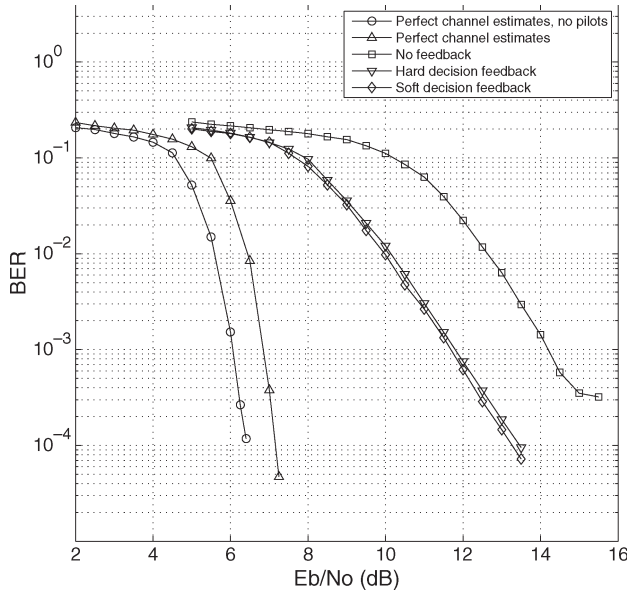


Fig. 7. BER performance of iterative channel estimation, equalization, and decoding over a frequency selective fading channel. The plot is for a channel with three taps of equal average power. The (1320, 2640) Margulis LDPC code is used in conjunction with Gray-coded 16-QAM modulation. The normalized fade rate is $f_d T_s = 0.005$.

(> 2.3 dB over the BER range of interest) is still achieved through the use of the iterative channel estimation scheme. As in Section IV-A3, an error floor is observed in the curve for feedback-free channel estimation, again due to initial channel estimation error exacerbating the effect of near codewords. Increasing the channel estimation filter length K does not provide an effective means to combat this error floor, since MSE decreases slowly with increasing K . However, the use of iterative channel estimation significantly lowers the MSE and, thus, the error floor.

B. MSE Performance of the Channel Estimator

The evolution of the MSE of the channel tap estimator with increasing number of iterations is plotted against E_b/N_0 in Fig. 8 for the simulation environment previously described in Section IV-A1, with the faster fade rate $f_d T_s = 0.02$. This plot is for the case of hard decision feedback, and the MSE is averaged over all taps. Also shown is the MMSE given by (7), which forms a lower bound for the observed MSE. In this plot, the MSE decreases with increasing number of iterations, for all values of E_b/N_0 . The MSE for the second iteration shows a dramatic improvement over the first, regardless of the value of E_b/N_0 . This is because the availability of the PCC decoder output gives the second channel estimation a distinct advantage over the first. Also, at lower values of E_b/N_0 , the MSE never approaches the lower bound because of the poor quality of the bit estimates fed back by the PCC decoder. At higher E_b/N_0 , however, the bit estimates improve with successive decoder iterations until eventually the feedback consists almost entirely of correct bits; thus, the MSE approaches the lower bound very tightly. It can be seen that the set of E_b/N_0 -values for which the MSE starts to approach the lower bound corresponds to the waterfall region of the corresponding plot in Fig. 4.

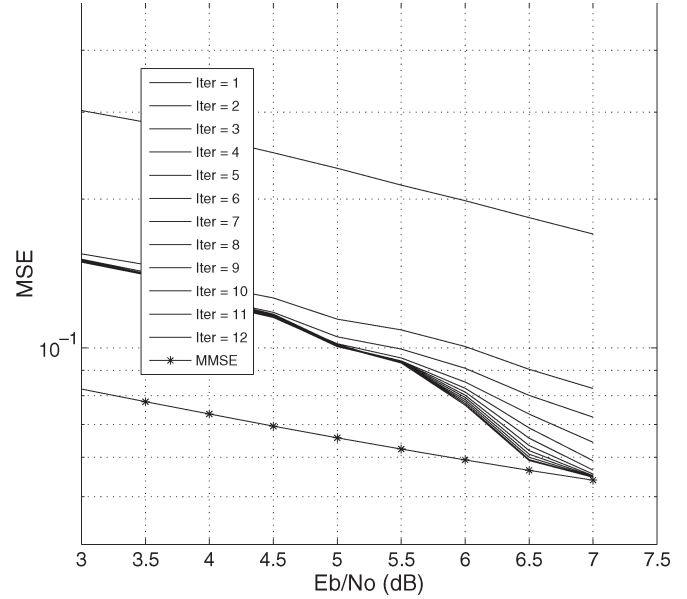


Fig. 8. Evolution of the MSE with increasing number of iterations. The plot is for a channel with three taps of equal average power. A PCC code is used with feedback and feedforward generators 15_o and 17_o , respectively, the code is punctured to rate approximately 0.5. BPSK modulation is used, and the normalized fade rate is $f_d T_s = 0.02$. The MSE decreases with increasing number of iterations for all values of E_b/N_0 .

C. Performance Versus Complexity Tradeoff for Channel Estimation Scheme

In the simulation results presented previously, the channel estimation filter length K and the pilot block spacing N_P were both optimized for performance versus complexity. The details of the effect of variation of these parameters are described previously.

1) Effect of Length and Type of Channel Estimation Filter:

Fig. 9 shows the effect of variation of length (K) and type (moving average/Wiener) of channel estimation filter, for the simulation environment described in Section IV-A1 (iterative channel estimation with hard decision feedback). For the cases of slower fade rate ($f_d T_s = 0.005$) and faster fade rate ($f_d T_s = 0.02$), the E_b/N_0 values chosen for simulations were 4.7 and 7.0 dB, respectively, as each of these E_b/N_0 values lies in the waterfall region of the corresponding BER curve (see Figs. 3 and 4). Also, the value of N_P was chosen to be 20 and 10 for the cases of slower and faster fade rates, respectively, based on the optimization of Section IV-C2.

For the case of the slower fade rate ($f_d T_s = 0.005$), the BER performances of the moving average filter and the Wiener filter are approximately equal for $K \leq 75$ and increase with increasing K in this range. For $K > 75$, the performance of the moving average filter actually deteriorates with increasing K ; this is because the approximation (9) no longer holds. The performance of the Wiener filter improves beyond $K = 75$; however, it yields little extra gain beyond this value. On the basis of this analysis, the value $K = 75$ and the moving average channel estimation filter were chosen for the slower fade rate; however, if necessary, a shorter channel estimation filter could be chosen at some expense in BER performance.

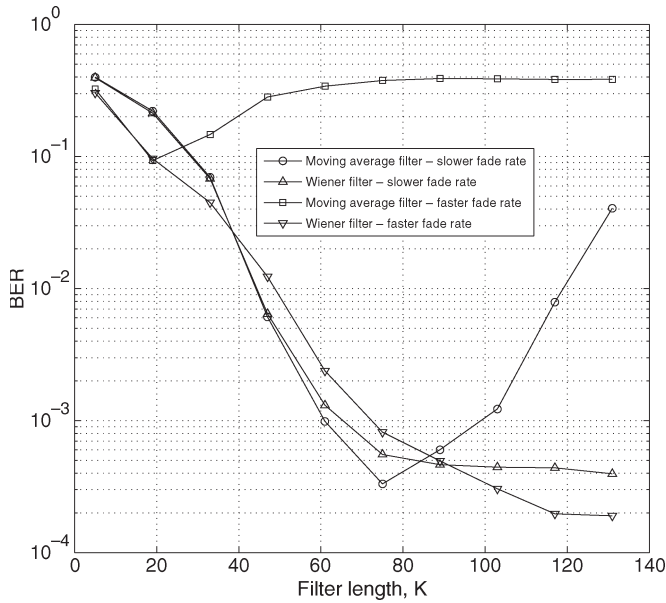


Fig. 9. Effect of channel estimation filter length and type on BER performance of iterative channel estimation, equalization, and decoding over a frequency selective fading channel. The plot is for BPSK transmission over a channel with three taps of equal average power. Hard decision feedback is assumed. A PCC code is used with feedback and feedforward generators 15_o and 17_o , respectively, and the code is punctured to rate approximately 0.5.

For the case of the faster fade rate ($f_d T_s = 0.02$), the moving average filter again exhibits an optimum filter length; however, in this case, the corresponding BER is too high to be of practical use. The Wiener filter again shows appreciable performance improvement with increasing K up to $K \approx 75$; beyond this value, diminishing returns are again observed. On the basis of this analysis, the value $K = 75$ and the Wiener channel estimation filter were chosen for the faster fade rate.

2) *Effect of Pilot Block Spacing:* Fig. 10 shows the effect of variation of the pilot block spacing N_P for the simulation environment described in Section IV-A1. For the cases of slower fade rate ($f_d T_s = 0.005$) and faster fade rate ($f_d T_s = 0.02$), the E_b/N_0 values chosen for simulations were 4.5 and 6.7 dB, respectively, as each of these E_b/N_0 values lies in the waterfall region of the corresponding BER curve (see Figs. 3 and 4). Results are included for the cases of hard and soft decision feedback. In all cases, the length of the channel estimation filter was chosen to be $K = 75$ based on the optimization of Section IV-C1.

Note that all four cases exhibit two qualitative effects. First, BER is degraded when the pilot block spacing is too low. This is because insertion of more pilots than are necessary to adequately track the channel incurs performance degradation through the loss in energy efficiency due to pilot insertion (this is $10 \log_{10}((N_P + 2L - 1)/N_P)$ dB). Second, BER is degraded when the pilot block spacing is too high. This is because the pilot insertion frequency is not sufficiently high to acquire the channel with precision, and the degraded channel estimation increases BER. Also, the interval of N_P -values for which the BER is acceptable is smaller in the case of faster fading, mainly because of the latter effect. On the basis of this analysis, we chose $N_P = 20$ for the slower fade rate, and

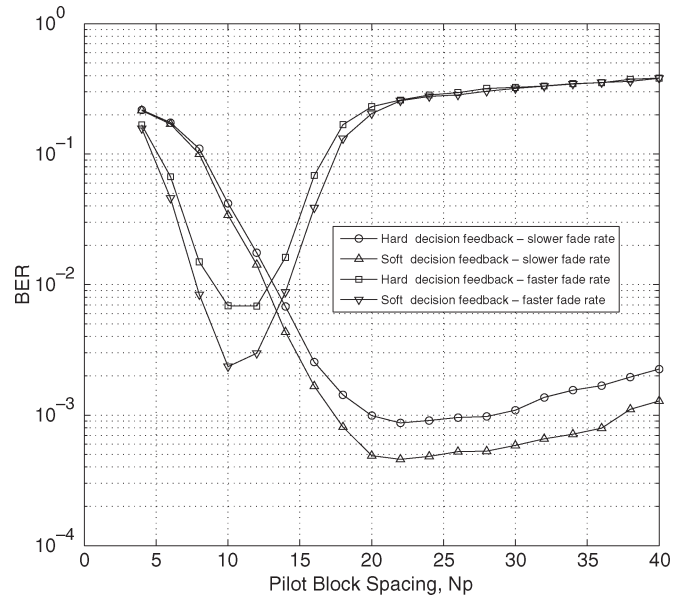


Fig. 10. Effect of pilot block spacing N_P on the BER performance of iterative channel estimation, equalization, and decoding over a frequency selective fading channel. The plot is for BPSK transmission over a channel with three taps of equal average power. A PCC code is used with feedback and feedforward generators 15_o and 17_o , respectively, and the code is punctured to rate approximately 0.5.

$N_P = 10$ for the faster fade rate, for use with both hard and soft decision feedbacks.

V. CONCLUSION

A scheme has been proposed for iterative channel estimation, equalization, and decoding for pilot symbol assisted transmission over a frequency selective Rayleigh fading channel. The scheme uses a channel estimator which has low complexity and is decoupled from the equalizer SISO module. The scheme is compatible with any memoryless QAM modulation scheme and with both PCC and LDPC coding. Significant performance improvement is observed for the scheme with iterative as opposed to noniterative channel estimation. Hard decision feedback of bit estimates derived from LLRs is simple to implement compared to factor-graph-based methods and is very effective. The channel estimation scheme may be regarded as a direct generalization of Valenti and Woerner's scheme to multitap channels, higher order modulations and LDPC and PCC coding. Because the scheme is demonstrated to work with both PCC and LDPC coding, a natural implementation arises using PCC encoding in the mobile transmitter and LDPC encoding in the BS transmitter. This maintains the scheme's performance while relegating complexity to the BS transmitter and receiver.

REFERENCES

- [1] C. Douillard, M. Jézéquel, C. Berrou, A. Picart, P. Didier, and A. Glavieux, "Iterative correction of intersymbol interference: Turbo-equalization," *Eur. Trans. Telecommun.*, vol. 6, no. 5, pp. 507–511, Sep./Oct. 1995.
- [2] G. Bauch, H. Khorrarn, and J. Hagenauer, "Iterative equalization and decoding in mobile communications systems," in *Proc. Eur. Pers. Mobile Commun. Conf.*, 1997, pp. 307–312.

- [3] J. I. Park and Y. Choi, "Turbo equalization for wireless cellular systems," *IEICE Trans. Fundamentals*, vol. E83-A, no. 6, pp. 1184–1185, Jun. 2000.
- [4] E. Tungsrisaguan and R. M. A. P. Rajatheva, "Turbo equalization with sequential sequence estimation over fading multipath channels," *IEEE Commun. Lett.*, vol. 6, no. 3, pp. 93–95, Mar. 2002.
- [5] R. Otnes and M. Tüchler, "Iterative channel estimation for turbo equalization of time-varying frequency-selective channels," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1918–1923, Nov. 2004.
- [6] S. Song, A. C. Singer, and K. M. Sung, "Soft input channel estimation for turbo equalization," *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 2885–2894, Oct. 2004.
- [7] 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.
- [8] E. Baccarelli and R. Cusani, "Combined channel estimation and data detection using soft statistics for frequency-selective fast-fading digital links," *IEEE Trans. Commun.*, vol. 46, no. 4, pp. 424–426, Apr. 1998.
- [9] A. Anastasopoulos and A. Polydoros, "Adaptive soft decision algorithms for mobile fading channels," *Eur. Trans. Commun.*, vol. 9, no. 2, pp. 183–190, Mar./Apr. 1998.
- [10] L. M. Davis, I. B. Collings, and P. Hoeher, "Joint MAP equalization and channel estimation for frequency-selective and frequency-flat fast-fading channels," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2106–2114, Dec. 2001.
- [11] M. C. Valenti and B. D. Woerner, "Iterative channel estimation and decoding of pilot symbol assisted turbo codes over flat-fading channels," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 9, pp. 1697–1705, Sep. 2001.
- [12] E. Zehavi, "8-PSK trellis codes for a Rayleigh channel," *IEEE Trans. Commun.*, vol. 40, no. 5, pp. 873–884, May 1992.
- [13] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927–946, May 1998.
- [14] X. Li and J. A. Ritcey, "Bit-interleaved coded modulation with iterative decoding," *IEEE Commun. Lett.*, vol. 1, no. 6, pp. 169–171, Nov. 1997.
- [15] M. Flanagan, A. D. Fagan, and M. Herro, "Pilot symbol assisted turbo equalization over a frequency selective Rayleigh fading channel," in *Proc. 3rd Int. Symp. Turbo Codes Related Topics*, Brest, France, Sep. 1–5, 2003, pp. 359–362.
- [16] H. Mai and A. G. Burr, "Iterative channel estimation for turbo equalization," in *Proc. IEEE Int. Symp. PIMRC*, Barcelona, Spain, Sep. 2004, pp. 1327–1331.
- [17] A. Lampe, "Iterative multiuser detection with integrated channel estimation for coded DS-CDMA," *IEEE Trans. Commun.*, vol. 50, no. 8, pp. 1217–1223, Aug. 2002.
- [18] M. Flanagan and A. D. Fagan, "Factor graph based derivation of a receiver for PCC coded transmissions over a frequency selective fading channel," in *Proc. 4th Int. Symp. Turbo Codes Related Topics*, Munich, Germany, Apr. 3–7, 2006.
- [19] J. Hou, P. H. Siegel, and L. B. Milstein, "Performance analysis and code optimization of low density parity-check codes on Rayleigh fading channels," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 5, pp. 924–934, May 2001.
- [20] W. C. Jakes, *Microwave Mobile Communications*. Hoboken, NJ: Wiley, 1974.
- [21] A. P. Worthen and W. E. Stark, "Unified design of iterative receivers using factor graphs," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 843–849, Feb. 2001.
- [22] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Soft-input soft-output modules for the construction and distributed iterative decoding of code networks," *Eur. Trans. Telecommun.*, vol. 9, no. 2, pp. 155–172, Mar./Apr. 1998.
- [23] S. Haykin, *Adaptive Filter Theory*, ser. Information and System Science Series. Englewood Cliffs, NJ: Prentice-Hall, 1996.
- [24] S. Dolinar and D. Divsalar, "Weight distributions for turbo codes using random and nonrandom permutations," Jet Propulsion Lab., California Inst. of Technol., Pasadena, pp. 56–65, TDA Progress Rep. 42-122, Aug. 1995.
- [25] G. A. Margulis, "Explicit constructions of graphs without short cycles and low-density codes," *Combinatorica*, vol. 2, no. 1, pp. 71–78, 1982.
- [26] S. Mertens, "Exhaustive search for low-autocorrelation binary sequences," *J. Phys. A, Math. Gen.*, vol. 29, no. 18, pp. 473–481, 1996.
- [27] M. Tüchler, "Estimation with non-ideal training information," in *Proc. 3rd Int. Symp. Turbo Codes Related Topics*, Brest, France, Sep. 1–5, 2003, pp. 35–38.
- [28] D. J. C. MacKay and M. Postol, "Weaknesses of Margulis and Ramanujan–Margulis low-density parity-check codes," *Electronic Notes in Theoretical Computer Science*, vol. 74, 2003.



Mark F. Flanagan (S'04–M'06) received the Ph.D. degree in electronic and electrical engineering from University College Dublin (UCD), Dublin, Ireland, in 2005.

He has been working as a Postdoctoral Researcher with the digital signal processing research group at UCD. His research interests include the design of iterative or "turbo" receivers (with emphasis on incorporating low-complexity signal processing into such receivers), algebraic constructions of low-density parity check (LDPC) codes, and low-complexity

LDPC encoder and decoder architectures. He also has research interests in the area of combined coding and modulation for wireline and wireless communications.



Anthony D. Fagan (S'76–M'77) received the Ph.D. degree from University College Dublin (UCD), Dublin, Ireland, in 1978.

He was a Research Engineer at Marconi Research Laboratories, Essex, U.K., from 1977 to 1980, where he worked on digital signal processing (DSP) for advanced communication systems. In 1980, he took up a position as a Lecturer with the Department of Electronic and Electrical Engineering, UCD, where he established the DSP Research Group. The group carries out a balanced mix of theoretical and applied

research in the areas of digital communications (wireline and wireless), speech and audio processing, image processing, pattern recognition, and biomedical signal processing. He is a member of the editorial board of the *Elsevier Journal Digital Signal Processing* and is a reviewer for many international journals. He frequently acts as a reviewer of projects and an evaluator of proposals for the European Commission. He has been instrumental in establishing a number of Irish companies that carry out work at the highest level in the area of physical layer communications, and he is widely consulted by many international communication companies. His main research interest is in the application of DSP techniques to advanced digital fixed line and wireless communications, with a particular interest in synchronization techniques, adaptive modulation, equalization, detection methods, and flexible use of the radio spectrum. He also has research interests in the areas of speech, audio, and image processing.