Iterative Channel Estimation and Decoding of Pilot Symbol Assisted Turbo Codes Over Flat-Fading Channels

Matthew C. Valenti, *Member, IEEE,* and Brian D. Woerner, *Member, IEEE*

Abstract—A method for coherently detecting and decoding turbo– coded binary phase shift keying (BPSK) signals transmitted over frequency-flat fading channels is discussed. Estimates of the complex channel gain and variance of the additive noise are derived first from known pilot symbols and an estimation filter. After each iteration of turbo decoding, the channel estimates are refined using information fed back from the decoder. Both hard-decision and soft-decision feedback are considered and compared with three baseline turbo-coded systems: 1) a BPSK system that has perfect channel estimates; 2) a system that uses differential phase shift keying and hence needs no estimates; and 3) a system that performs channel estimation using pilot symbols but has no feedback path from decoder to estimator. Performance can be further improved by borrowing channel estimates from the previously decoded frame. Simulation results show the influence of pilot symbol spacing, estimation filter size and type, and fade rate. Performance within 0.49 and 1.16 dB of turbo-coded BPSK with perfect coherent detection is observed at a bit–error rate of 10^{-4} for normalized fade rates of $f_d T_s = 0.005$ and $f_d T_s = 0.02$, respectively.

Index Terms—Channel coding, channel estimation, fading channels, pilot-symbol assisted modulation, turbo codes.

I. INTRODUCTION

TURBO codes, introduced in [1], have been shown to exhibit near-capacity performance over Rayleigh flat-fading channels with coherent detection and perfect knowledge of the channel response [2]. However, mobile communication systems are characterized by channel responses with time-varying magnitude and phase. For mobile communications using coherent detection, special care must be taken to properly estimate and track the carrier phase. For turbo-coded systems, the situation is even more complicated, as the decoder requires estimates of both the magnitude of the fading process and the variance of the additive noise process. The situation is exacerbated by the fact that turbo codes typically operate at such low signal-to-noise ratios that conventional carrier tracking techniques often fail.

In [3], Hall and Wilson consider turbo-coded systems employing frequency shift keying (FSK) with noncoherent detection and differential phase shift keying (DPSK) with differential detection. The receivers for such systems are cheaper and easier to build than their coherent counterparts, primarily due to the elimination of the requirement to estimate and track the phase of the carrier. However, due to the noncoherent combining penalty, the performance of such systems is significantly degraded. Compared to coherent binary phase shift keying (BPSK), a loss of 6.1 and 2.7 dB was observed for FSK and DPSK, respectively, in additive white Gaussian noise (AWGN) with a (2048, 1024) turbo code at a bit-error rate (BER) of 10^{-5}. In [4], a 4.5-dB performance loss was shown for a similar code with DPSK in Rayleigh flat-fading at a BER of 10^{-5}.

One proposal for reducing the noncoherent combining loss is to use multiple-symbol differential detection to detect the DPSK signal [5]. Here the observation window is widened to include more than just the two symbols used for conventional differential detection. Multiple-symbol differential detection is extended to the case of convolutionally coded systems in [6]. The system is considered to contain a serially concatenated turbo code, with the outer code generated by the convolutional encoder and the inner code generated by the differential encoder. As with all serially concatenated turbo codes, the decoder operates in an iterative manner [7]. To account for the time-varying nature of the channel, the techniques of linear prediction [8] and per-survivor processing [9] are used. For M-ary DPSK, the proposed system requires the trellis of the inner decoder to be expanded from M states to M^N states, where N is the order of the predictor. For low fade rates, the system gains back much of the performance degradation due to differential noncoherent combining.

An alternative solution that we explore in this paper is to incorporate pilot symbols into the transmission. Like pilot tones, pilot symbols are used at the receiver to obtain an estimate of the channel response so that coherent detection can be performed. Following the convention of [10], we use the term *pilot symbol assisted modulation* (PSAM) to describe systems that incorporate pilot symbols in the transmission. In [11], it is shown that for uncoded systems, the use of pilot symbols is more power efficient than the use of pilot tones. Reference [12] shows that the same holds true for turbo-coded systems. A branch metric suitable for turbo-coded systems using PSAM is presented in [13], although this metric requires knowledge of the carrier phase. In [12] and [13], all of the estimation was performed prior to the
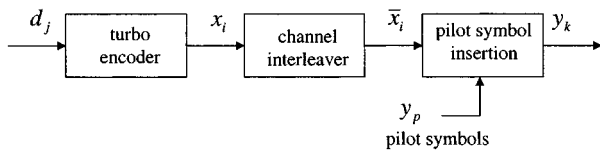
Fig. 1.   Discrete-time transmitter model.

first iteration of turbo decoding. However, since turbo decoding is an iterative process, performance can be improved by re-estimating the channel after each decoder iteration [14]. Iterative estimation and decoding was proposed for convolutional codes in [15], for BPSK modulated turbo codes in [4], and for QAM modulated turbo codes in [16].

This paper presents a detailed study of iterative estimation and decoding of turbo codes over fading channels. In Section II, the system model is presented. Four estimator/decoders are described in Section III: one that does not use feedback from decoder to estimator [12], one that uses hard-decision feedback [17], one that uses soft-decision feedback [18], and one that uses both soft-decision feedback and channel estimates from the previously decoded frame. Simulation results in Section IV show the impact of pilot symbol spacing, filter type and size, and fade rate. Finally, a conclusion is given in Section V.

II. SYSTEM MODEL

A. Transmitter

A block diagram of the discrete-time transmitter model is shown in Fig. 1. A sequence $\{d_j\}$, $1 \leq j \leq L$ of random data in polar form $d_j \in \{-1, 1\}$ is first encoded by a rate r turbo encoder. The encoded bits $\{x_i\}$, $1 \leq i \leq L/r$, $x_i \in \{-1, 1\}$ are then passed to an m by n block channel interleaver. Next the interleaved sequence $\{\overline{x}_i\}$ is parsed into groups of $(M-1)$ contiguous bits, where M is the *pilot symbol spacing* (assumed to be odd) [10]. A known pilot symbol y_p is placed in the center of each group, and the new groups of size M are reassembled into the sequence of symbols $\{y_k\}$, $1 \leq k \leq LM/(r(M-1))$. For ease of exposition, we assume that the pilot symbols all take on the same value y_p, although in practice the pilot symbols may take on different values.

The channel interleaver is required because turbo encoding may not be sufficient to cope with the errors induced by the fading channel. Correlated fading channels tend to produce burst errors, while turbo codes are more effective with uncorrelated errors [2]. Therefore, the channel interleaver scrambles the order of the symbols at the transmitter in order to make the channel appear uncorrelated at the input to the decoder. Note that the channel interleaver is different from the nonuniform interleaver required within the turbo encoder. To distinguish these two interleavers, we refer to the interleaver within the turbo encoder as the *coding interleaver* and the interleaver between the encoder and modulator as the *channel interleaver*.

For comparison purposes, we also consider ideal coherent detection of BPSK, both with and without pilot symbols. With ideal BPSK, it is assumed that the channel parameters are known precisely at the receiver. Therefore no pilot symbols are required, although we will show the performance both with and without pilot symbols. The benefit of showing the performance
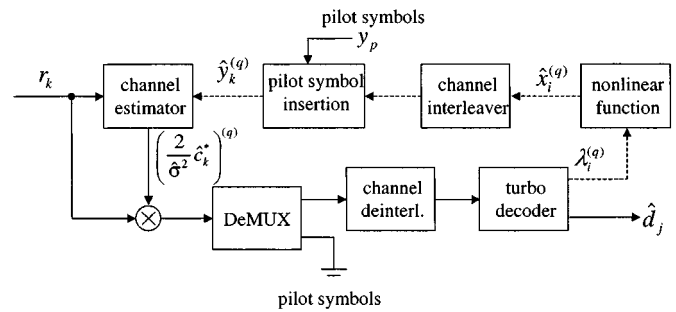

Fig. 2.   Proposed receiver.

of ideal BPSK both with and without pilot symbols is that it allows one to distinguish between the degradation caused by imperfect channel estimation and the loss of energy efficiency due to the use of pilot symbols. In addition, we consider differential detection of binary DPSK. For DPSK, the pilot symbol insertion block in Fig. 1 is replaced by a differential encoder.

B. Channel

The transmitted sequence $\{y_k\}$ passes through a discrete-time slow flat-fading channel with additive white Gaussian noise so that the decision statistic after matched filtering is

$$r_k = c_k y_k + n_k. \tag{1}$$

In the above, $\{n_k\}$ is a set of statistically independent complex-valued Gaussian random variables with zero-mean and variance $\sigma^2 = N_o/2E_s$ in each direction, where E_s is the energy per symbol and N_o is the one-sided noise spectral density. Note that when pilot symbols are used, $E_s = rE_b(M-1)/M$. Otherwise, $E_s = rE_b$ where E_b is the energy per information bit.

For Rayleigh fading, $\{c_k\}$ is modeled as a zero-mean discrete-time complex Gaussian process. In particular, we follow Jakes's isotropic scattering model [19] by assuming that the real and imaginary parts of $\{c_k\}$ are independent with autocorrelation

$$R_c[k] = \tfrac{1}{2} J_o(2\pi f_d T_s k) \tag{2}$$

where
f_d relative Doppler between transmitter and receiver;
T_s symbol period;
$J_o(\cdot)$ zeroth order Bessel function of the first kind.
Note that the power of $\{c_k\}$ is normalized to unity.

C. Receiver

Here we compare two basic receiver structures, one with feedback from decoder to estimator and one without feedback. The receiver without feedback is depicted by Fig. 2 with the dotted lines removed. In this receiver, all of the channel estimation is performed prior to decoding as suggested in [12].

First, the sequence $\{r_k\}$ is sent to a channel estimation algorithm (described in the next section). The algorithm computes estimates $\{\hat{c}_k\}$ of the fading process and $\hat{\sigma}^2$ of the noise variance. The algorithm outputs the sequence $\{2\hat{c}_k^*/\hat{\sigma}^2\}$, where

$\hat{c}^*$ denotes the complex conjugate of $\hat{c}$. The output of the estimator is multiplied by the received sequence $\{r_k\}$. This result is passed to a demultiplexer which strips off the pilot symbols (they are not needed by the decoder). Next, the sequence is passed through a channel deinterleaver and finally to a turbo decoder. We assume that the turbo decoder is implemented using the log-MAP algorithm [20]. The decoder produces estimates $\{\hat{d}_j\}$ of the data sequence.

The receiver shown in Fig. 2 can be improved by including the feedback path indicated by dotted lines. In this case, the turbo decoder outputs log-likelihood ratio (LLR) estimates of the code symbols $\{\lambda_i^{(q)}\}$ after each decoder iteration q. The LLRs are passed through a nonlinearity to be discussed in the next section. Depending on the nature of the nonlinearity, either hard- or soft-decision estimates of the code symbols $\{\hat{x}_i^{(q)}\}$ are produced. The symbol estimates are reinterleaved and pilot symbols are reinserted. The resulting estimated symbol sequence $\{\hat{y}_k^{(q)}\}$ is then fed back into the channel estimator in a decision directed manner. The channel is re-estimated prior to the next decoder iteration. New estimates $\{(2\hat{c}_k^*/\hat{\sigma}^2)^{(q)}\}$ are produced by the channel estimation algorithm and used (after appropriate deletion of pilot symbols and deinterleaving) by the turbo decoder during iteration $(q+1)$.

With turbo-coded DPSK, the receiver structure is much simpler since no channel estimator is needed and there are no pilot symbols to remove [3]. The channel deinterleaver is given the value $r_k r_{k-1}^*$ which is used directly by the decoder after deinterleaving. Note that one could use the principles contained in this paper in conjunction with DPSK. That is, simple differential detection could be used prior to the first iteration of decoding and then tentative bit decisions could be used to coherently detect the signal prior to the remaining iterations. However, due to time limitations and in order to keep our comparison simple, we only compare against simple differential detection in this paper.

III. CHANNEL ESTIMATOR

A. Channel Gain Estimation

If the transmitted sequence $\{y_k\}$ were known at the receiver, then the best linear minimum mean-square-error (MMSE) estimate of the complex channel gain is found according to [21]

$$\hat{c}_k = \sum_{i=-\lfloor K/2 \rfloor}^{\lfloor K/2 \rfloor} w_i y_{k-i} r_{k-i} \tag{3}$$

where K is the size of the filter (assumed to be odd) and w_i is a set of filter coefficients found by solving the Wiener–Hopf equations

$$\sum_{i=-\lfloor K/2 \rfloor}^{\lfloor K/2 \rfloor} w_i R_c[m-i] + 2\sigma^2 w_m = R_c[m], \quad -\left\lfloor \frac{K}{2} \right\rfloor \le m \le \left\lfloor \frac{K}{2} \right\rfloor. \tag{4}$$

However, the only values of $\{y_k\}$ that are known at the receiver *a priori* are the pilot symbols. Because it is assumed that the channel is slowly varying, a good approximation to (3) can be found using

$$\hat{c}_k = \sum_{i=-\lfloor K/2 \rfloor}^{\lfloor K/2 \rfloor} w_i y_p r_{p(k-i)} \tag{5}$$

where we define $r_{p(k-i)}$ to be the received value of the pilot symbol located closest to r_{k-i}, i.e.,

$$p(k-i) = M \left\lfloor \frac{k-i}{M} \right\rfloor + \left(\frac{M+1}{2} \right). \tag{6}$$

Note that when (5) is used by the receiver to compute the initial set of channel gain estimates $\{\hat{c}_k^{(0)}\}$, the filter size can be reduced from K to K/M since only the received pilot symbols are used.

For the iterative receiver, refined channel estimates are found for $q > 0$ using

$$\hat{c}_k^{(q)} = \sum_{i=-\lfloor K/2 \rfloor}^{\lfloor K/2 \rfloor} w_i \hat{y}_{k-i}^{(q)} r_{k-i}. \tag{7}$$

In the above equation, $\hat{y}_{k-i}^{(q)} = y_p$ when $(k-i)$ is the index of a pilot symbol, otherwise it is the interleaved symbol estimate whose value depends on the nature of the nonlinearity in Fig. 2. For hard-decision feedback, the nonlinearity operates according to [17]

$$\hat{x}_i^{(q)} = \begin{cases} 1, & \text{if } \lambda_i^{(q)} > 0 \\ -1, & \text{if } \lambda_i^{(q)} \le 0 \end{cases} \tag{8}$$

where $\lambda_i^{(q)}$ is the LLR output of the turbo decoder after iteration q corresponding to code symbol i. If, on the other hand, soft-decision feedback is desired, then [22]

$$\hat{x}_i^{(q)} = \tanh\left(\frac{\lambda_i^{(q)}}{2} \right). \tag{9}$$

Note that if the normalized fade rate $f_d T_s$ is slow ($f_d T_s \ll 1$) and the filter size K is sufficiently small ($K \ll (f_d T_s)^{-1}$), the filter coefficients are all approximately equal

$$w_i \approx \frac{1}{K + 2\sigma^2} \quad \forall i. \tag{10}$$

When $K \gg 2\sigma^2$ this becomes

$$w_i \approx \frac{1}{K} \quad \forall i. \tag{11}$$

If the approximation of (11) is used with equality, then the filter is a simple moving average (MA). The benefit of using a moving average is that it is simpler than a Wiener filter and does not require knowledge of the fade rate or autocorrelation of the channel. In the simulation results presented in the next section, we compare the performance of the Wiener filter with the performance of the moving average.

B. Noise Variance Estimation

Noise variance estimation and its effect on the performance of turbo codes has been an important topic of interest in the liter-

ature. Several studies illustrate the sensitivity of noise variance estimate errors on performance [23] and present various blind estimators [24], [25]. The consensus of these studies is that the performance of turbo codes is not extremely sensitive to noise variance estimation errors. In particular, estimation errors that are less than about 3 dB do not noticeably degrade the turbo code's performance.

To determine the noise variance, first assume that the set of channel gains $\{c_k\}$ and transmitted symbols $\{y_k\}$ are known at the receiver. Then form the random variable

$$z_k = r_k - c_k y_k \qquad (12)$$
$$= n_k. \qquad (13)$$

The sequence $\{z_k\}$ is thus a set of independent Gaussian random variables with zero mean and variance σ^2. The best estimate of σ^2 can be found by simply taking the sample variance of z_k.

Of course $\{c_k\}$ and $\{y_k\}$ are not perfectly known at the receiver. Again assuming that the channel is slowly varying, (13) can be approximated by

$$z_k = r_{p(k)} - \hat{c}_{p(k)} y_p \qquad (14)$$

where $\hat{c}_{p(k)}$ is the estimated channel gain associated with the pilot symbol located closest to r_k.

The sample variance of (14) is used by the receiver to determine the initial noise variance estimate $(\hat{\sigma}^2)^{(0)}$. The iterative receiver recomputes the noise variance after each iteration $q > 0$ by calculating the sample variance of

$$z_k^{(q)} = r_k - \hat{c}_k^{(q)} \hat{y}_k^{(q)}. \qquad (15)$$

Thus, the variance estimator uses the output of the channel gain estimator along with the tentative decisions of the turbo decoder and knowledge of the pilot symbols.

C. Interframe Estimation

Until now, we have assumed that estimation occurs on a frame-by-frame basis. Thus the channel estimates computed using (5) and (7) only use received samples r_k and symbol estimates $\hat{y}_k^{(q)}$ belonging to the turbo code word of interest. Information from previous and subsequent frames are not used. Thus when $k - i < 1$ or $k - i > LM/(r(M-1))$ in (5) and (7), $r_{p(k-i)}$ or r_{k-i} is set to zero.

The problem with performing estimation only on a frame-by-frame basis is that the estimates at the beginning and end of the interleaved turbo code word tend to be rather poor. As will be shown in the simulation results, this contributes to a higher error floor than with perfectly coherent detection.

Ideally, the estimation algorithm should be initialized using information from the previously decoded frame as well as from the next decoded frame. In practice, normally only information from the previous frame is available, although it may be possible to acquire information from the next frame at the cost of increased latency and complexity. Thus a simple, but effective, improvement to the proposed algorithm is to use the channel estimates of the last $K/2$ symbols of the previous frame to assist with channel estimation in the current frame. This can be accomplished by using the final channel estimates from the previous frame $\hat{c}_{k-i}^{(Q)}$ in place of $y_p r_{p(k-i)}$ in (3) and in place of $y_{k-i} r_{k-i}$ in (5) whenever $k - i < 1$ (Q is the total number of decoder iterations).

IV. SIMULATION STUDY

A. Simulation Description

An extensive simulation study was undertaken to investigate the coded performance of the various receiver structures and channel estimation techniques discussed in Sections III and IV. The turbo code under consideration is composed of two rate 1/2 recursive convolutional (RSC) encoders, each with constraint length 4 (with feedback generator polynomial $1 + D + D^3$ and feedforward generator polynomial $1 + D + D^2 + D^3$, both in octal). The trellis of the upper encoder is terminated with 3 tail bits, while the trellis of the lower encoder is left open. The data frame consists of 1247 data bits and 3 tail bits. The turbo encoder uses a $L = 1250$ bit S-random interleaver, with $S = 20$ [26]. Puncturing is used to increase the overall code rate to 1/2. In particular, the even indexed parity bits from the upper encoder and odd indexed parity bits from the lower encoder are deleted prior to transmission. The channel interleaver has a depth of 50 and is implemented with a 50 by 50 matrix (just one code word is interleaved at a time).

Slow fading is assumed, and two normalized fade rates are considered: $f_d T_s = 0.005$ and $f_d T_s = 0.02$. The slower fade rate corresponds to a typical digital cellular system operating at 900 MHz with 19.2 kbaud symbol rate and a relative mobile velocity of approximately 70 mph, while the faster code rate corresponds to a PCS system operating at 1.9 GHz with 9.6 kbaud symbol rate and 70 mph mobile velocity. Twelve decoder iterations are performed using the log-MAP algorithm. Enough trials are run to generate 40 independent frame errors for each value of E_b/N_o considered.

B. Comparison of Reception Techniques

Figs. 3 and 4 compare the BER performance of seven different transmission and reception techniques for the normalized fade rates of $f_d T_s = 0.005$ and $f_d T_s = 0.02$, respectively. For each figure, a family of seven curves showing BER versus E_b/N_o is shown. From most power efficient to least power efficient, these curves are: 1) performance using ideal BPSK without pilot symbols; 2) performance using PSAM with ideal coherent detection; 3) performance using PSAM with iterative estimation using both soft-decision feedback and channel estimates from the previously decoded turbo code word; 4) performance using PSAM with frame-by-frame iterative estimation and soft-decision feedback; 5) performance using PSAM with frame-by-frame iterative estimation and hard-decision feedback; 6) performance using PSAM with no feedback from decoder to estimator; and 7) DPSK with no channel estimation. For the five curves in the middle of each figure, pilot symbols are multiplexed into the transmitted stream. For the slower fade rate (Fig. 3), a pilot symbol spacing of $M = 21$ was used, while for the faster fade rate (Fig. 4), $M = 11$. For the slower fade rate, (11) holds and thus the channel estimation filter is implemented with a moving average. For the faster
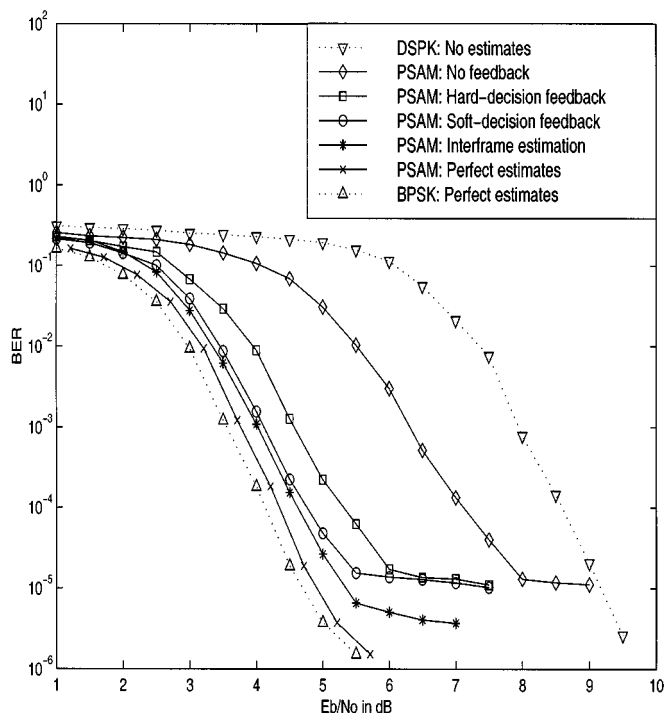
Fig. 3. Comparison of coded bit error performance versus E_b/N_o for several transmission/reception techniques over a complex Rayleigh flat fading channel with normalized fade rate $f_d T_s = 0.005$ and block channel interleaving. A turbo code with rate 1/2, constraint length 4, and 1250 bit S-random interleaver is used. The five solid curves in the middle use PSAM with pilot symbol spacing $M = 21$ and a $K = 61$ sample moving average estimation filter.
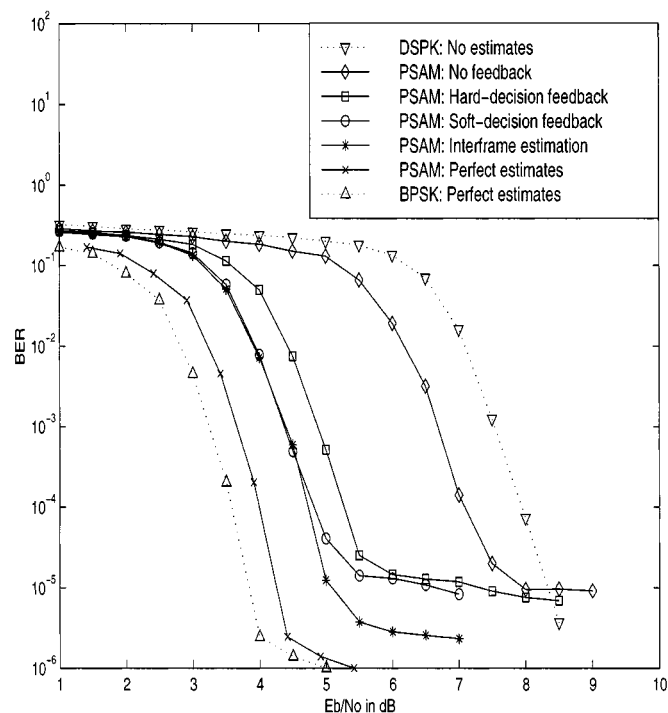

Fig. 4. Comparison of coded bit error performance versus E_b/N_o for several transmission/reception techniques over a complex Rayleigh flat fading channel with normalized fade rate $f_d T_s = 0.02$ and block channel interleaving. A turbo code with rate 1/2, constraint length 4, and 1250 bit S-random interleaver is used. The five solid curves in the middle use PSAM with pilot symbol spacing $M = 11$ and a $K = 61$ Wiener estimation filter.

fade rate, (11) does not hold and thus a Wiener filter must be used. For both fade rates the size of the channel estimator was $K = 61$. The choice of filter type, pilot symbol spacing, and filter size were made on the basis of a set of simulations (described below) that explored the impact of these parameters.

Several observations can be made from Figs. 3 and 4. First, the performance of both ideal BPSK and DPSK improves with increasing fade rate. For instance, the BER is 10^{-4} at 4.13 dB for ideal BPSK at the slower fade rate but only 3.58 dB for the faster fade rate. This difference can be attributed to the ability of the channel interleaver to better break up correlated fading for the faster fade rate [25]. Thus, the channel errors for the faster fade rate are more randomly distributed at the input to the decoder and performance is improved. Another observation is that the use of DPSK with differential detection imposes a severe noncoherent combining penalty of approximately 4.4 dB for both fade rates at a BER of 10^{-4}.

Performance can be improved over the DPSK case by using PSAM. When PSAM is used without any feedback from decoder to estimator, performance is improved by 1.5 dB at the slower fade rate and 0.9 dB at the faster fade rate (at BER 10^{-4}). This performance improvement can be extended by incorporating feedback from the decoder to the estimator. Hard-decision feedback provides an additional 1.8 dB of coding gain over PSAM without feedback at both fade rates. Soft-decision feedback improves performance by an additional 0.5 dB over hard-decision feedback (again at both fade rates). An incremental gain (approximately 0.1 dB at 10^{-4} BER) is observed when estimates from the previous frame are used to assist the

estimator. For the slower fade rate, the loss at 10^{-4} BER due to using pilot symbols and the best channel estimation technique is 0.49 dB, of which 0.21 dB is due to the use of pilot symbols and 0.28 dB is due to imperfect channel estimation. For the faster fade rate, the total loss increases to 1.16, 0.42 dB due to using pilot symbols and 0.74 dB due to imperfect channel estimation.

It is important to note that for all four cases using PSAM and channel estimation, a noticeable BER floor at about 10^{-5} is present. For the frame-by-frame estimation techniques, the BER floor is higher than the BER floor encountered by ideal BPSK by about an order of magnitude. This BER floor is apparently due to residual error in the estimation process and the fact that estimates from adjacent frames are not used. Note that the floor is reduced by about half an order of magnitude simply by using estimates from the previously decoded frame. We speculate that this floor could be reduced further by also using estimates from the next frame, but such a strategy would increase latency and complexity.

C. Influence of Pilot Symbol Spacing

The influence of the pilot symbol spacing M on BER performance is illustrated in Fig. 5 for the normalized fade rates of $f_d T_s = 0.005$ and $f_d T_s = 0.02$. The performance of (frame-by-frame) hard-decision feedback is compared with that of soft-decision feedback. In each case, the signal-to-noise ratio is fixed at $E_b/N_o = 4.5$ dB and the size of the channel estimation filter is $K = 61$. For the slower fade rate a moving average is used, while for the faster fade rate a Wiener filter is used. Each of these curves takes on a bowl shape. For low values of M, performance
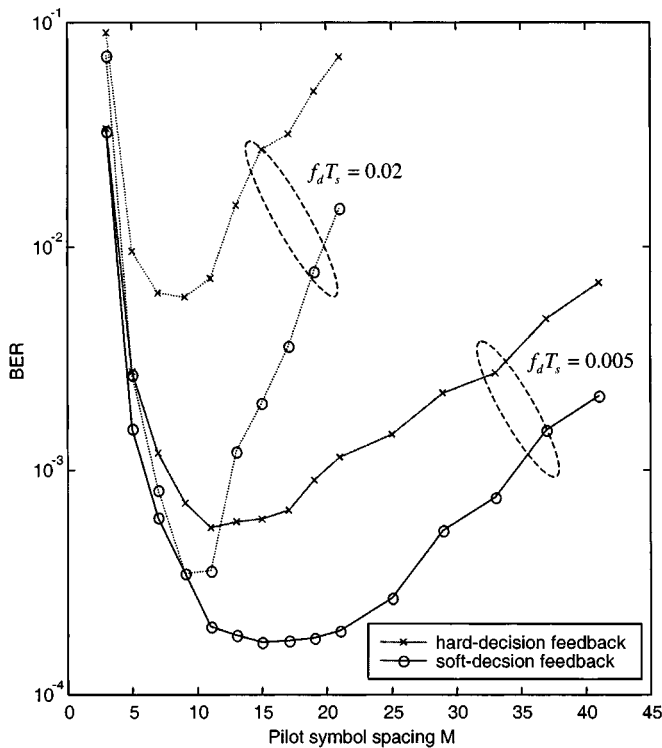
Fig. 5. Bit error performance as a function of pilot symbol spacing M for a turbo-coded system operating over a Rayleigh fading channel with $E_b/N_o = 4.5$ dB, soft-decision feedback, filter length $K = 61$, and both $f_dT_s = 0.005$ and $f_dT_s = 0.02$.

Fig. 6. Bit error performance as a function of filter order K for a turbo-coded system operating over a Rayleigh fading channel with $f_dT_s = 0.005$ and $E_b/N_o = 4.5$ dB. The pilot symbol spacing is $M = 21$. Both hard-decision and soft-decision feedback are considered with both Wiener and moving average estimation filters.

Fig. 7. Bit error performance as a function of filter order K for a turbo-coded system operating over a Rayleigh fading channel with $f_dT_s = 0.02$ and $E_b/N_o = 4.5$ dB. The pilot symbol spacing is $M = 11$. Both hard-decision and soft-decision feedback are considered with both Wiener and moving average estimation filters.

is poor because an unnecessarily large portion of the overall energy is being devoted to the transmission of pilots. At first, performance improves with increasing M because the energy per channel symbol $E_s = rE_b(M - 1)/M$ increases. However, for larger values of M, performance begins to degrade with increasing M because the pilot symbol insertion rate is no longer sufficient to track the channel. According to the Nyquist Sampling theorem, the pilot symbol insertion rate $R_M = 1/(MT_s)$ must be at least twice the bandwidth of the fading process. This implies that for $f_dT_s = 0.005$, $M < 100$, and for $f_dT_s = 0.02$, $M < 25$. However, in practice the best pilot symbol spacing may be significantly lower than the limit imposed by the Sampling theorem. Our simulation results indicate that the best pilot symbol spacings are in the range $9 \leq M \leq 25$ for the slower rate and $7 \leq M \leq 13$ for the faster rate. On the basis of these curves, we have selected to use $M = 21$ for the slow fade rate and $M = 11$ for the fast rate. It is interesting to note that the system is more forgiving at the slower fade rate in the sense that the range of acceptable M is larger than it is for the faster fade rate.

D. Influence of Filter Type and Size

In Figs. 6 and 7 the influence of filter size and type (moving average or Wiener) is illustrated for $f_dT_s = 0.005$ and $f_dT_s = 0.02$, respectively. Each curve compares the performance of (frame-by-frame) hard-decision feedback with soft-decision feedback and the two filter types. In each case, the signal-to-noise ratio is fixed at $E_b/N_o = 4.5$ dB. The pilot symbol spacing for the slower fade rate is $M = 21$, while
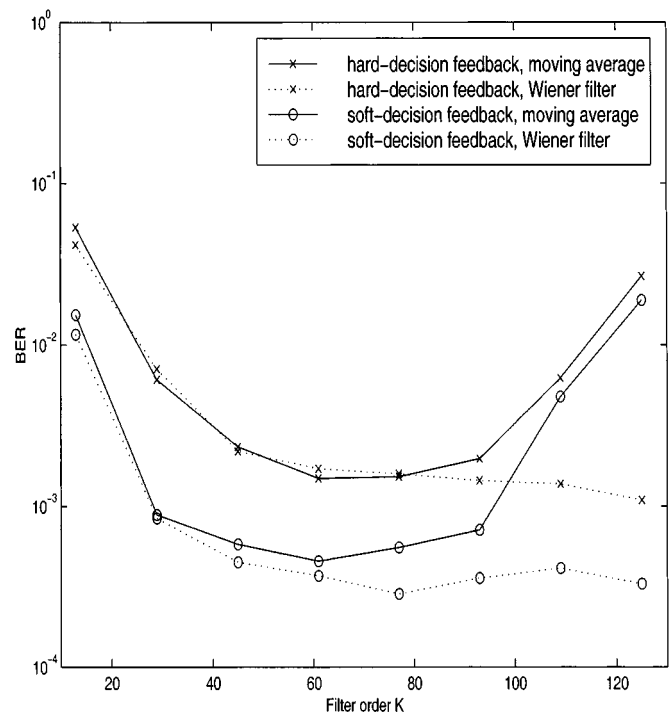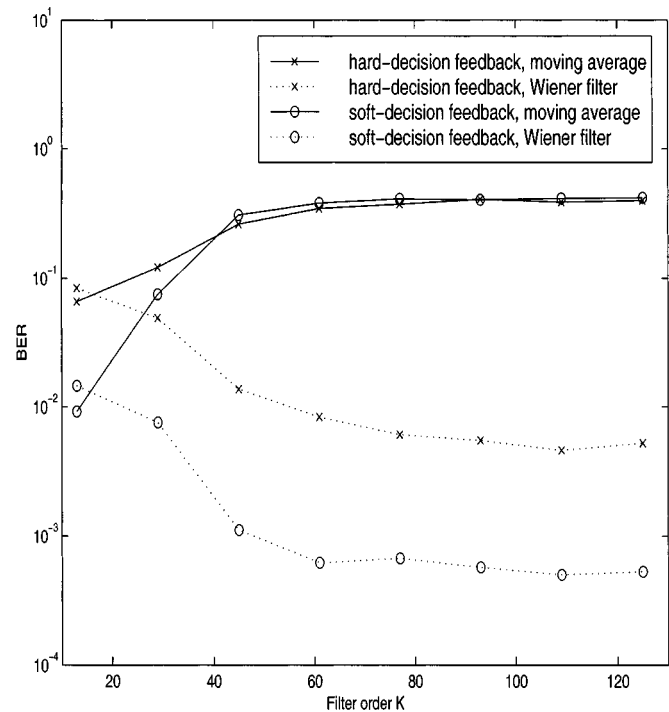
the pilot symbol spacing for the faster fade rate is $M = 11$. As expected, the performance when using the Wiener filter
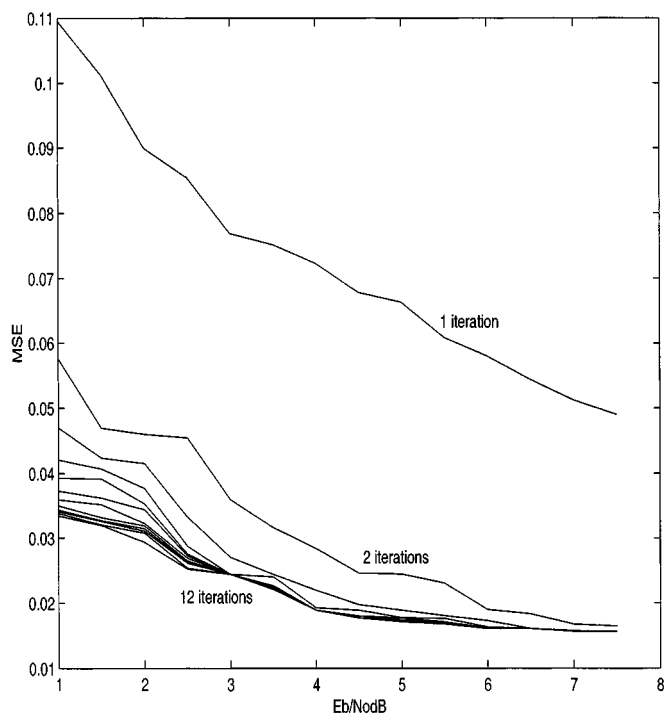
Fig. 8. Mean square error (MSE) versus E_b/N_o as parameterized by number of iterations. A rate 1/2, constraint length 4 turbo code with 1250 bit S-random interleaver is used over a complex Rayleigh fading channel with normalized fade rate $f_d T_s = 0.005$. PSAM is used in conjunction with soft-decision feedback at the receiver. The pilot symbol spacing is $M = 21$ and a length $K = 61$ moving average estimation filter is used.
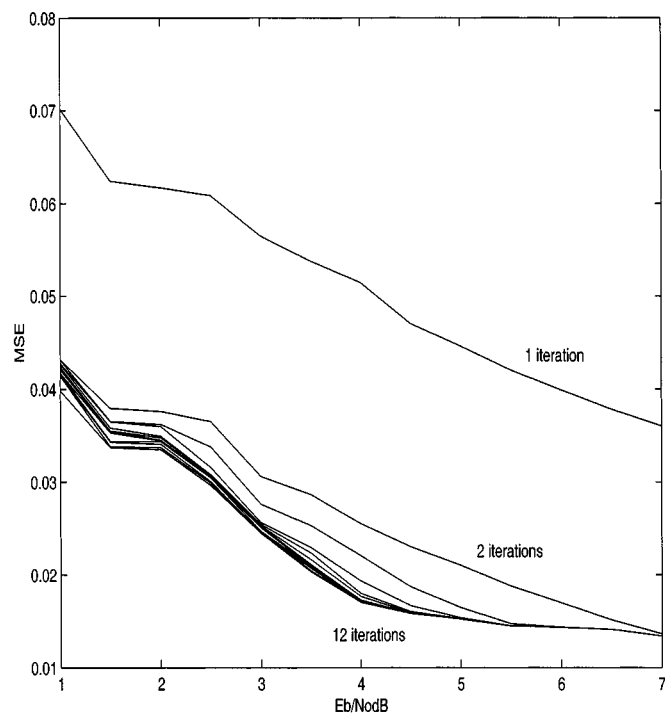

Fig. 9. MSE versus E_b/N_o as parameterized by number of iterations. A rate 1/2, constraint length 4 turbo code with 1250 bit S-random interleaver is used over a complex Rayleigh fading channel with normalized fade rate $f_d T_s = 0.02$. PSAM is used in conjunction with soft-decision feedback at the receiver. The pilot symbol spacing is $M = 11$ and a length $K = 61$ Wiener estimation filter is used.

improves with increasing filter size K. However, the curves tend to flatten out, indicating that there is little to be gained by using filters larger than $K \approx 61$. If a smaller filter is desired, reasonable performance can be achieved with $K \approx 45$. For the slower fade rate and small K, the performance using a moving average is nearly identical to that of the Wiener filter. However, as K increases beyond $K = 61$ the two curves diverge and the performance of the moving average actually gets worse with increasing K. This is because (11) no longer holds and hence the channel changes significantly within the window of the MA. For the faster fade rate, the performance of the moving average quickly deteriorates with increasing K. This is because (11) does not hold for even small values of K, such as $K = 21$. On the basis of these curves, we have chosen to use $K = 61$ and to use a moving average for the slower fade rate and a Wiener filter for the faster fade rate.

E. Influence of Noise Variance Estimation

In all the simulation results shown in this paper, it is assumed that neither the channel gains $\{c_k\}$ nor the noise variance σ^2 are known. Thus, both must be estimated at the receiver. We ran additional simulations to explore the impact of estimating each of these parameters. In particular, we ran many of the same simulations shown in this paper under the assumption that the noise variance is known at the receiver and thus only the channel gains must be estimated. The results showed that knowledge of the noise variance only improves performance slightly (never more than 0.1 dB), and therefore we have not included the results in
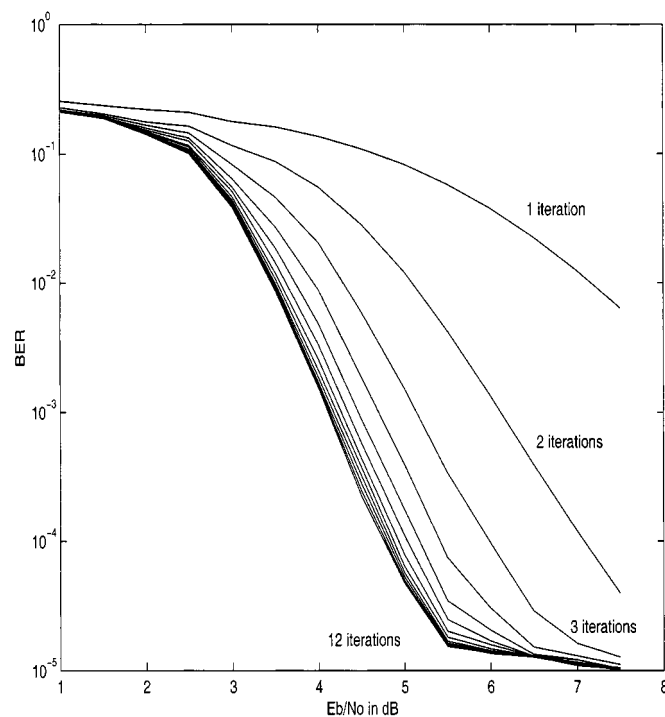

Fig. 10. BER versus E_b/N_o as parameterized by number of iterations. A rate 1/2, constraint length 4 turbo code with 1250 bit S-random interleaver is used over a complex Rayleigh fading channel with normalized fade rate $f_d T_s = 0.005$. PSAM is used in conjunction with soft-decision feedback at the receiver. The pilot symbol spacing is $M = 21$ and a length $K = 61$ moving average estimation filter is used.
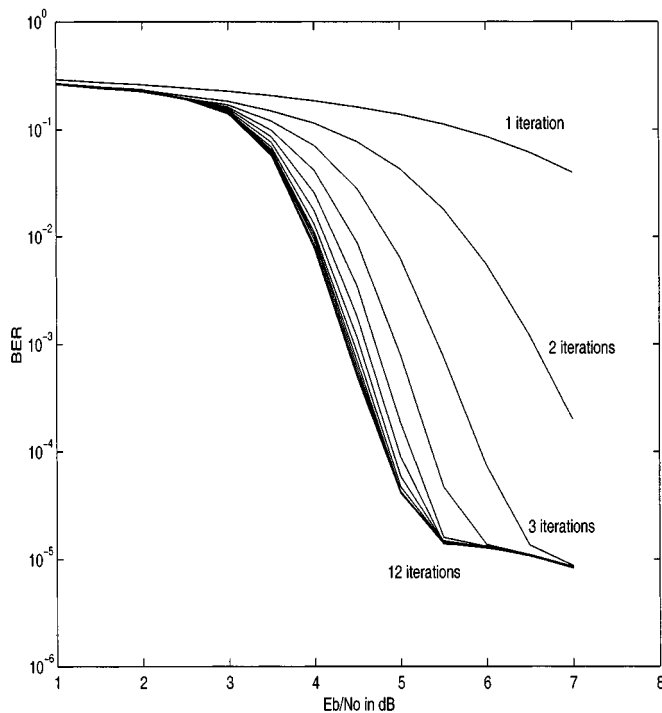
Fig. 11.   BER versus E_b/N_o as parameterized by number of iterations. A rate 1/2, constraint length 4 turbo code with 1250 bit S-random interleaver is used over a complex Rayleigh fading channel with normalized fade rate $f_d T_s = 0.02$. PSAM is used in conjunction with soft-decision feedback at the receiver. The pilot symbol spacing is $M = 11$ and a length $K = 61$ Wiener estimation filter is used.

this paper because they would appear repetitious. From these experiments, we can safely draw the conclusion that most of the performance degradation comes from estimating the channel gain, while estimating the noise variance only has a minimal impact on performance.

F. MSE of the Estimator

Finally, we wish to show the performance of the estimator and decoder as a function of the number of decoder iterations. Figs. 8 and 9 show the MSE after each iteration of frame-by-frame soft-decision feedback estimation for $f_d T_s = 0.005$ and $f_d T_s = 0.02$, respectively. The BER curves that correspond to Figs. 8 and 9 are shown in Figs. 10 and 11. Notice that there is a very large improvement in MSE from the first iteration to the second. The MSE improvement from the second to the third iteration is still noticeable, but not nearly as dramatic as the first improvement. Performance continues to improve with increasing iterations, but as with iterative processing in general, obeys a law of diminishing returns. For sufficiently large E_b/N_o, no improvement in MSE is achieved after the third decoder iteration. These curves indicate that perhaps it is not necessary to recalculate the estimates after each iteration. For instance, it may be sufficient to calculate the channel estimates only after the first three iterations of turbo decoding.

V. Conclusion

Turbo codes can be coherently detected over flat-fading channels with the help of pilot symbols. Dramatic performance

improvements can be achieved by iteratively estimating the channel and decoding the turbo code. Iterative estimation and decoding can be implemented using either hard-decision or soft-decision feedback, with soft-decision feedback outperforming hard-decision by about 0.5 dB. The pilot symbol spacing has a major role in determining the overall performance of the system. Too many pilot symbols result in wasted energy, but too few limits the ability of the estimator to track the channel. The length and type of the channel estimation filter also plays a role in performance. One drawback of the proposed system is that the BER floor is about an order of magnitude higher than that of a turbo-coded system with ideal BPSK detection. However, this floor can be significantly reduced by using estimates from the previously decoded frame.

The proposed technique assumes that the fade rate is both known at the receiver and remains constant during each frame. A natural extension of this work is to incorporate an adaptive estimation filter that can accommodate unknown or varying fade rates. The filter could be implemented with a Kalman filter instead of a Wiener filter. Furthermore, the concepts discussed in this paper could be applied to other modulation techniques such as M-ary PSK, QAM, and DPSK.

References

[1] C. Berrou, A. Glavieux, and P. Thitimasjshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes(1)," in *Proc. IEEE Int. Conf. Commun.*, Geneva, Switzerland, May 1993, pp. 1064–1070.

[2] E. K. Hall and S. G. Wilson, "Design and analysis of turbo codes on Rayleigh fading channels," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 160–174, Feb. 1998.

[3] ——, "Turbo codes for noncoherent channels," in *Proc. IEEE GLOBECOM, Communication Theory Mini-Conf.*, Phoenix, AZ, Nov. 1997, pp. 66–70.

[4] M. C. Valenti and B. D. Woerner, "Refined channel estimation for coherent detection of turbo codes over flat-fading channels," *Electron. Lett.*, vol. 34, pp. 1648–1649, Aug. 20, 1998.

[5] D. Divsalar and M. K. Simon, "Multiple-symbol differential detection of MPSK," *IEEE Trans. Commun.*, vol. 38, pp. 300–308, Mar. 1990.

[6] P. Hoeher and J. Lodge, "Turbo DPSK: Iterative differential PSK demodulation and channel decoding," *IEEE Trans. Commun.*, vol. 47, pp. 837–843, June 1999.

[7] S. Benedetto, D. Divsalar, D. Montorsi, and F. Pollara, "Serial concatenation of interleaved codes: Performance analysis, design, and iterative decoding," *IEEE Trans. Inform. Theory*, vol. 44, pp. 909–926, May 1998.

[8] J. Makhoul, "Linear prediction: A tutorial review," *Proc. IEEE*, vol. 63, pp. 561–580, Apr. 1975.

[9] R. Raheli, A. Polydoros, and C. K. Tzou, "Per-survivor processing: A general approach to MLSE in uncertain environments," *IEEE Trans. Commun.*, vol. 43, pp. 354–364, Feb./Mar./Apr. 1995.

[10] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 686–693, Nov. 1991.

[11] J. K. Cavers and M. Liao, "A comparison of pilot tone and pilot symbol techniques for digital mobile communications," in *Proc. IEEE GLOBECOM*, Orlando, FL, Dec. 1992, pp. 915–921.

[12] L.-D. Jeng, Y. T. Su, and J.-T. Chiang, "Performance of turbo codes in multipath fading channels," in *Proc. IEEE Vehicular Technology Conf.*, Ottawa, Canada, May 1998, pp. 61–65.

[13] K. L. Li and S. W. Cheung, "Modified MAP algorithm incorporated with PSA technique for turbo codes in Rayleigh fading channels," *Electron. Lett.*, vol. 35, pp. 537–539, Apr. 1, 1999.

[14] M. C. Valenti, "Iterative detection and decoding for wireless communications," Ph.D. dissertation, Virginia Tech., Blacksburg, VA, July 1999.

[15] H. El Gamal, M. M. Khairy, and E. Geraniotis, "Iterative decoding and channel estimation of DS/CDMA over slow Rayleigh fading channels," in *Proc. IEEE PIMRC*, Boston, MA, Sept. 1998, pp. 1299–1303.

[16] Y. Xu, H. J. Su, and E. Geraniotis, "Iterative decoding and channel estimation of DS/CDMA over slow Rayleigh fading channels," in *Proc. IEEE MILCOM*, Atlantic City, NJ, Nov. 1999, pp. 86–91.

[17] M. C. Valenti and B. D. Woerner, "A bandwidth efficient pilot symbol technique for coherent detection of turbo codes over fading channels," in *Proc. IEEE MILCOM*, Atlantic City, NJ, Nov. 1999.

[18] M. C. Valenti, "Iterative channel estimation for turbo codes over fading channels," in *Proc. IEEE Wireless Commun. Network. Conf.*, Chicago, IL, Sept. 2000, invited paper, pp. 1019–1024.

[19] W. C. Jakes, *Mobile Microwave Communication*. New York: Wiley, 1974.

[20] 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. 119–125, Mar./Apr. 1997.

[21] C. W. Therrien, *Discrete Random Signals and Statistical Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1992.

[22] M. Sandell, C. Luschi, P. Strauch, and R. Yan, "Iterative channel estimation using soft decision feedback," in *Proc. IEEE GLOBECOM*, Sydney, Australia, Nov. 1998.

[23] M. A. Jordan and R. A. Nichols, "The effects of channel characteristics on turbo code performance," in *Proc. IEEE MILCOM*, 1996, pp. 17–21.

[24] T. A. Summers and S. G. Wilson, "SNR mismatch and online estimation in turbo decoding," *IEEE Trans. Commun.*, vol. 46, pp. 421–423, Apr. 1998.

[25] M. C. Valenti and B. D. Woerner, "Performance of turbo codes in interleaved flat fading channels with estimated channel state information," in *Proc. IEEE Vehic. Technol. Conf.*, Ottawa, Canada, May 1998, pp. 66–70.

[26] S. Dolinar and D. Divsalar, "Weight distributions for turbo codes using random and nonrandom permutations,", JPL TDA Progress Rep., vol. 42, Aug. 15, 1995.

Matthew C. Valenti (M'99) received the B.S.E.E. degree from Virginia Polytechnic Institute and State University (Virginia Tech), Blacksburg, VA, in 1992, the M.S.E.E. degree from The Johns Hopkins University, Baltimore, MD, in 1995, and the Ph.D. degree in electrical engineering from Virginia Tech in 1999, where he was a Bradley Fellow.

He is currently an Assistant Professor with the Department of Computer Science and Electrical Engineering at West Virginia University, Morgantown. His research interests include the areas of communication theory, error correction coding, applied information theory, and wireless multiple-access networks. He also acts as a Consultant to several companies engaged in various aspects of turbo codec design, including software radio, FPGA, and ASIC implementations for military, satellite, and third-generation cellular applications. Prior to attending graduate school at Virginia Tech, he was an Electronics Engineer with the United States Naval Research Laboratory, Washington, DC, where he was engaged in the design and development of a space-borne adaptive antenna array and a system for the collection and correlation of maritime ELINT signals.

Brian D. Woerner (S'88–M'91) was born on October 11, 1964. He received the B.S. degree in computer and electrical engineering from Purdue University, West Lafayette, IN, in 1986, and the M.S. and Ph.D. degrees from the University of Michigan, Ann Arbor, in 1987 and 1991, respectively, where he was a Unisys Fellow. He also received the Master of Public Policy degree from the University of Michigan's School of Public Policy Studies, with an emphasis in telecommunications policy.

He is currently a Professor with the Bradley Department of Electrical Engineering at Virginia Tech, Blacksburg. He has served two-year terms as Director of the Mobile and Portable Radio Research Group and is currently an Associate Director. He has been a member of the group since 1991.