Iterative Channel Estimation for Turbo Codes over Fading Channels*

Matthew C. Valenti
Dept. of Comp. Sci. & Elec. Eng.
West Virginia University
Morgantown, WV 26506-6109
mvalenti@wvu.edu

Abstract- **A method for coherently detecting and decoding turbo coded 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 simulation results compare the proposed technique with DPSK, ideal BPSK, and non-iterative PSAM based reception.**

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, conventional coherent detection is difficult to implement since the phase must be estimated and tracked. 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.

Several solutions have been presented in the literature to the problem of transmitting turbo codes over fading channels. One possibility is to use noncoherently detected FSK or differentially detected DPSK [3]. However, due to the noncoherent combining penalty, performance is significantly degraded with respect to ideal coherently detected BPSK. Performance can be drastically improved by using iterative multiple-symbol differential detection and decoding [4]. This technique regains much of the loss due to noncoherent combining but requires implicit channel tracking through the use of per-survivor processing and linear prediction.

*This work was supported by the Office of Naval Research.

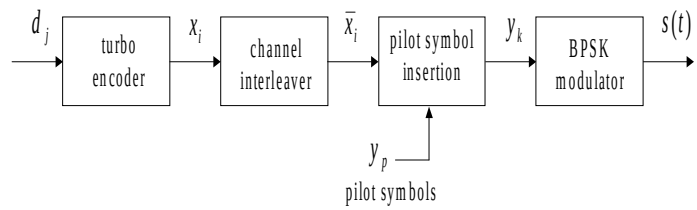

Figure 1: Transmitter.

An alternative to using noncoherent modulation is to use PSK and then attempt to track the channel at the receiver. Such techniques typically employ pilot symbol assisted modulation (PSAM) [5]. Two general approaches can be taken: (1) the channel can be tracked using a MMSE approach involving the Wiener filtering of the pilot symbols [6], and (2) the phase of the channel can be tracked by modeling it as a Markov process [7]. In this paper we pursue the first approach, but allow the estimates to be refined by feeding back information from the decoder to the channel estimator as suggested in [8]. Under moderate fading conditions (i.e. a normalized fade rate of $f_d T_s = 0.005$), performance within about 0.6 dB of ideal BPSK detection (i.e. with perfect channel state information) can be achieved by using the proposed technique.

II. System Model

A. Transmitter

As shown in Fig. 1, a sequence $\{d_j\}, 1 \leq j \leq L$, of random data $d_j \in \{-1, 1\}$ 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 through a $m \times n$ block channel interleaver. Next the sequence $\{\bar{x}_i\}$ is parsed into groups of $(M-1)$ contiguous bits, where M is the *pilot symbol spacing* [5]. A pilot symbol y_p is placed in the center of each group, and the new groups of size M are reassembled into the sequence $\{y_k\}, 1 \leq k \leq LM/(r(M-1))$. This sequence is then fed into a BPSK modulator which produces the output $s(t) = y_k g(t)$, where $g(t)$ is an arbitrary pulse shape with energy E_s.

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 random 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. With ideal BPSK, it is assumed that the channel parameters are known precisely at the receiver and thus no pilot symbols are required. In this case, the pilot symbol insertion block in Fig. 1 is removed and the code symbols are modulated immediately following the interleaver. Furthermore, we also 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 signal $s(t)$ passes through a fading channel with additive white Gaussian noise so that

$$r(t) \quad = \quad c(t)s(t) + n(t), \tag{1}$$

where c(t) is a complex-valued process that describes the multiplicative distortion of the frequency-flat fading channel and $n(t)$ is complex zero-mean white-Gaussian noise. The two-sided noise spectral density of $n(t)$ is $N_o/2$ in each of the real and imaginary directions.

For Rayleigh fading, $c(t)$ is modeled as a zero-mean complex Gaussian low-pass process. The bandwidth of this process is equal to the Doppler shift due to the relative motion between transmitter and receiver. In particular, we adopt an isotropic scattering model and assume that the real and imaginary parts of $c(t)$ are independent with autocorrelation [9]

$$R_c(\tau) \quad = \quad \frac{1}{2} J_o(2\pi f_d \tau), \tag{2}$$

where f_d is the relative Doppler between transmitter and receiver, and $J_o(\cdot)$ is the zeroth order Bessel function of the first kind. Note that the power of $c(t)$ 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

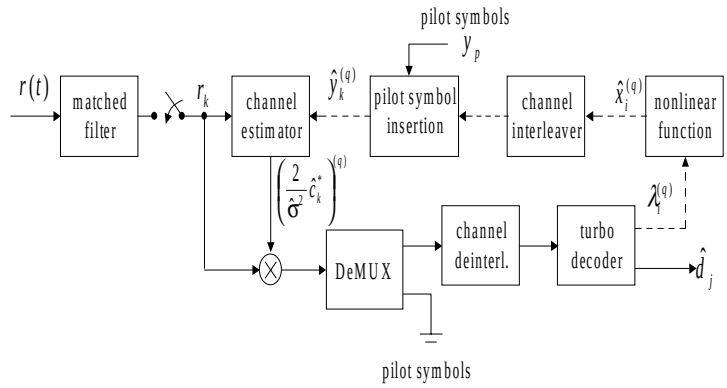

Figure 2: Proposed receiver.

with the dotted lines removed. In this receiver, all of the channel estimation is performed prior to decoding as suggested in [6]. The received signal $r(t)$ is passed through a matched filter and then sampled at the symbol rate $1/T_s$. We assume that $c(t)$ varies slowly with respect to the symbol period, perfect timing is available, and there is no intersymbol interference. The sequence at the output of the sampler can be expressed as

$$r_k \quad = \quad c_k y_k + n_k, \tag{3}$$

where $c_k = c(kT_s)$ and $\{n_k\}$ 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. 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.

After sampling, 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 $\hat{c}$. This output is multiplied by the received sequence $\{r_k\}$, and the 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, which produces estimates 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 $(q+1)$.

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 [10]

$$\hat{c}_k = \sum_{i=-\lfloor K/2 \rfloor}^{\lfloor K/2 \rfloor} w_i y_{k-i} r_{k-i}, \tag{4}$$

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.

Note that 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 (4) 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 $r_{p(k-i)}$ is the received value of the pilot symbol located closest to r_{k-i}.

Equation (5) is used by the receiver to compute the initial set of channel gain estimates $\left\{\hat{c}_k^{(0)}\right\}$. For the iterative receiver, refined channel estimates are found for $q > 0$ using

$$\hat{c}_k^{(q)} = \sum_{i=-\lfloor K/2 \rfloor}^{\lfloor K/2 \rfloor} w_i \hat{y}_{k-i}^{(q)} r_{k-i}. \tag{6}$$

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 [11]

$$\hat{x}_i^{(q)} = \begin{cases} 1 & \text{if } \lambda_i^{(q)} > 0 \\ -1 & \text{if } \lambda_i^{(q)} \leq 0. \end{cases} \tag{7}$$

If, on the other hand, soft-decision feedback is desired, then [12]

$$\hat{x}_i^{(q)} = \tanh\left(\frac{\lambda_i^{(q)}}{2}\right). \tag{8}$$

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

$$w_i \approx \frac{1}{K} \quad \forall i. \tag{9}$$

If the approximation of (9) 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.

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 literature. Several studies illustrate the sensitivity of noise variance estimate errors on performance [13], and present various blind estimators [14], [15]. 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

$$z_k = y_k r_k - c_k \tag{10}$$

$$= y_k n_k. \tag{11}$$

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, (11) can be approximated by

$$z_k = y_p r_{p(k)} - \hat{c}_{p(k)}, \tag{12}$$

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

The sample variance of (12) 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

$$z_k^{(q)} = \hat{y}_k^{(q)} r_k - \hat{c}_k^{(q)}. \tag{13}$$

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.

IV. Simulation Study

A. Simulation Description

The performance of the proposed receiver structures and channel estimators was determined by simulation. We considered a turbo code composed of two rate 1/2, constraint length 4 RSC encoders (feedback generator $(15)_o$ and feedforward generator $(17)_o$). The upper encoder's trellis was terminated with 3 tail bits, while the lower encoder's trellis was left open. A $L = 1250$ bit S-random interleaver [16], with $S = 20$, was used in conjunction with frames of 1247 data bits and 3 tail bits. An overall code rate of $r = 1/2$ was achieved by deleting the even indexed parity bits from the upper encoder and odd indexed parity bits from the lower encoder. The channel interleaver depth was 50, and was implemented with a 50 by 50 matrix (i.e. just one code word is interleaved at a time).

Slow fading was assumed, and two normalized fade rates were considered: $f_dT_s = 0.005$ and $f_dT_s = 0.02$. 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 were performed using the log-MAP algorithm [17]. Enough trials were run to generate 40 independent frame errors for each value of E_b/N_o considered.

B. Comparison of Reception Techniques

Fig. 3 and 4 compare the BER performance of five transmission and reception techniques for the normalized fade rates of $f_dT_s = 0.005$ and $f_dT_s = 0.02$, respectively. For each figure, a family of five curves showing BER vs. E_b/N_o is shown. From most power efficient to least power efficient, these curves are (1) performance using ideal BPSK (i.e. perfect channel estimates), (2) performance using PSAM with iterative estimation and soft-decision feedback, (3) performance using PSAM with iterative estimation and hard-decision feedback, (4) performance using PSAM with no feedback from decoder to estimator, and (5) DPSK with no channel estimation. For the three 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, (9) holds and thus the channel estimation filter is implemented with a moving average. For the faster fade rate, (9) 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 that explored the impact of these parameters.

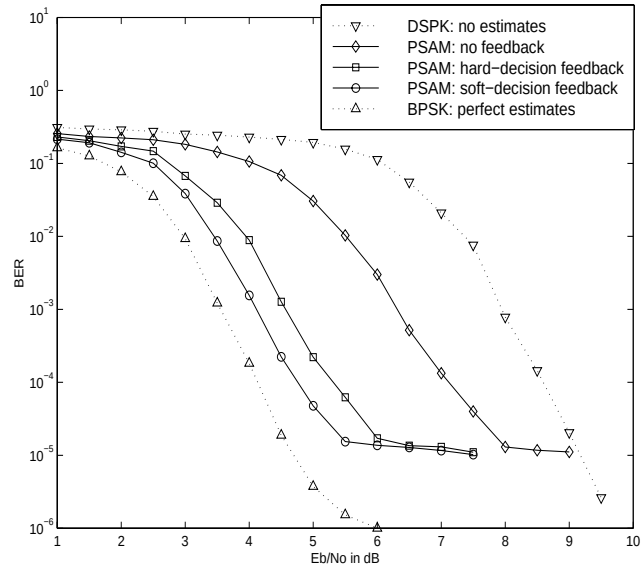

Figure 3: Comparison of coded bit error performance versus E_b/N_o for several transmission/reception techniques over a complex Rayleigh flat fading channel with normalized fade rate $f_dT_s = 0.005$ and block channel interleaving.

Several observations can be made from Fig. 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 at 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 [15]. 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}, which agrees with the observations made in [3]. 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).

It is important to note that for all three PSAM-based techniques, a noticeable BER floor at about 10^{-5} is present. This 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

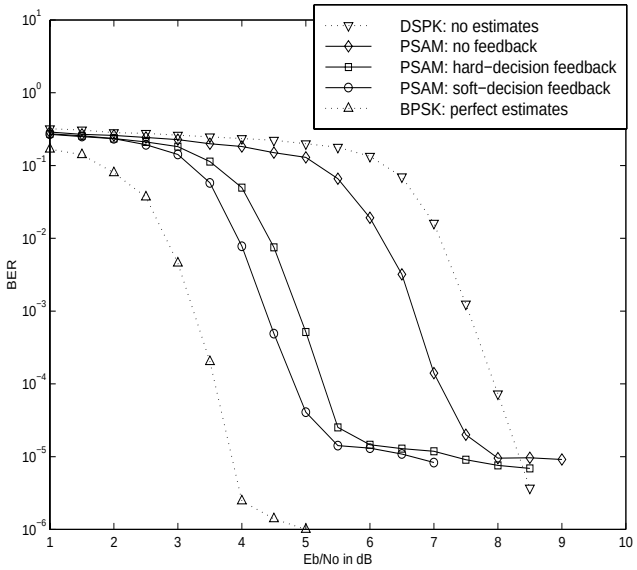

Figure 4: Comparison of coded bit error performance versus E_b/N_o for several transmission/reception techniques over a complex Rayleigh flat fading channel with normalized fade rate $f_dT_s = 0.02$ and block channel interleaving.

process. Our simulations show that the mean squared error (MSE) of the estimator does not go to zero even for the soft-decision feedback decoder. Since we only estimate on a frame-by-frame basis, the symbols at the beginning and end of the frame tend to have poor estimates, and thus impose a lower limit on the performance of the estimator. The estimator performance could be improved by using channel estimates from adjacent frames, which would in turn lower the BER floor.

C. Influence of Pilot Symbol Spacing

The influence of the pilot symbol spacing M on BER performance is illustrated in Fig. 5 and 6 for the normalized fade rates of $f_dT_s = 0.005$ and $f_dT_s = 0.02$, respectively. Each curve compares the performance of hard-decision feedback with 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 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

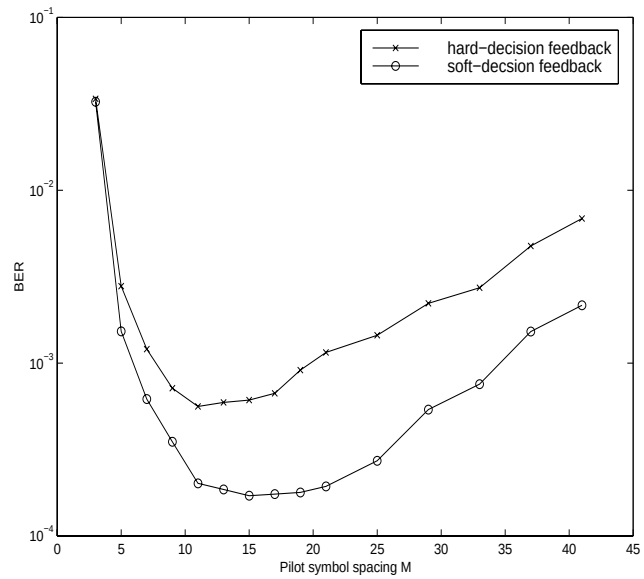

Figure 5: Bit error performance as a function of pilot symbol spacing M for a turbo coded system operating over a Rayleigh fading channel with $f_dT_s = 0.005$ and $E_b/N_o = 4.5$ dB.

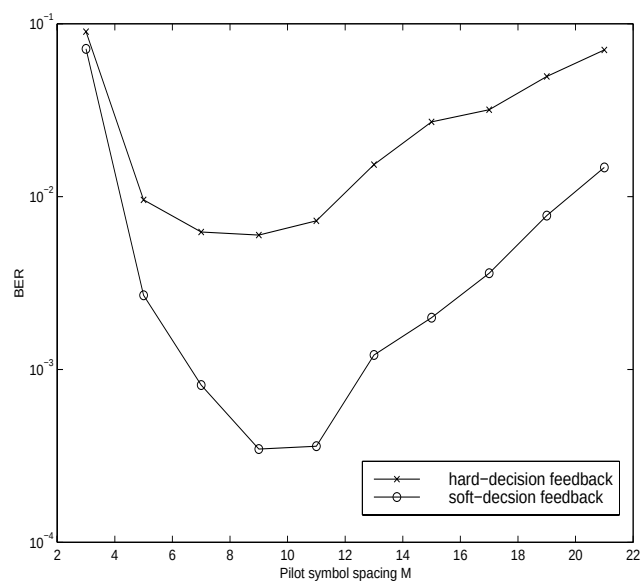

Figure 6: Bit error performance as a function of pilot symbol spacing M for a turbo coded system operating over a Rayleigh fading channel with $f_dT_s = 0.02$ and $E_b/N_o = 4.5$ dB.

fading process. This implies that for $f_d T_s = 0.005$, $M < 100$ and for $f_d T_s = 0.02$, $M < 25$. 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.

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. 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 may be reduced by borrowing channel estimates from adjacent frames.

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. on Commun.*, (Geneva, Switzerland), pp. 1064–1070, May 1993.

[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] E. K. Hall and S. G. Wilson, "Turbo codes for noncoherent channels," in *Proc., IEEE GLOBECOM, Communication Theory Mini-Conference*, (Phoenix, AZ), pp. 66–70, Nov. 1997.

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

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

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

[7] C. Komminakis and R. D. Wessel, "Pilot-aided joint iterative data and channel estimation in flat Rayleigh fading," in *Proc., IEEE GLOBECOM, Communication Theory Mini-Conference*, (Rio de Janero, Brazil), Dec. 1999.

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

[9] W. C. Jakes, *Mobile Microwave Communication.* New York: John Wiley & Sons, 1974.

[10] C. W. Therrien, *Discrete Random Signals and Statistical Signal Processing.* Prentice Hall, 1992.

[11] 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.

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

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

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

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

[16] S. Dolinar and D. Divsalar, "Weight distributions for turbo codes using random and nonrandom permutations," *JPL TDA Progress Report*, vol. 42, pp. 56–65, Aug. 15th 1995.

[17] P. Robertson, P. Hoeher, and E. Villebrun, "Optimal and sub-optimal maximum a posteriori algorithms suitable for turbo decoding," *European Trans. on Telecommun.*, vol. 8, pp. 119–125, Mar./Apr. 1997.