Soft Decision-Directed Map Estimate of Fast Rayleigh Flat Fading Channels

Ming Yan and Bhaskar D. Rao

***Abstract*—In this letter, an iterative receiver with soft-decision feedback is derived by using the expectation–maximization algorithm for maximum *a posteriori* estimate of fast Rayleigh flat fading channels. Simulation results indicate that in a fast fading environment, the derived receiver can perform better than an iterative receiver with hard-decision feedback.**

***Index Terms*—Maximum *a posteriori* (MAP) estimation, Rayleigh channels.**

I. Introduction

THE expectation–maximization (EM) algorithm is an effective method for maximizing likelihood parameter estimation in the context of "missing data" [1], [2]. Consequently, it has found wide applications in communications such as synchronization [3], multiuser detection [4], and equalization [5]–[7]. Particularly for the problem of data detection, two typical approaches [5], [8] using EM have been proposed. One approach [8] aims at maximum-likelihood sequence estimation (MLSE) on unknown time-varying flat fading channels, and takes the unknown fading process as missing/auxiliary data. The other [5] aims at maximum-likelihood (ML) estimate of a deterministic intersymbol interference (ISI) channel, and takes the unknown transmitted data as missing/auxiliary data. The similarity is that they both iterate between channel estimation and data detection. The difference is that in [8], the previous ML data sequence estimate is used to reestimate the fading process, while in [5], "soft decisions" of data/code symbols, which are computed via *a posteriori* probabilities (APPs), are used to reestimate the time-invariant channel.

In this letter, we extend the idea in [5] and propose an iterative receiver for estimating time-varying flat fading channels. Because the flat fading channel is correlated over time, it has *a priori* probability density, which we will use for EM-based maximum *a posteriori* (MAP) channel estimation rather than for ML estimation. The derived receiver uses soft decisions (defined as the expected value of data/code symbols) and fixed-interval smoothing by a Kalman filter to reestimate the fading process. For comparison purposes, we refer to the method in [8] as Info EM and to our method as Channel EM. Info EM tries to maximize the likelihood function of the information sequence rather than the *a posteriori* density function of the fading process. So Info EM has a better optimization criterion than Channel EM. However, since global convergence is not guaranteed, better optimization criterion does not necessarily translate into better performance. In fact, in a fast fading environment, it is found by simulation that Channel EM can perform better than Info EM. An intuitive and potential explanation for this relative performance is provided in Section III.

Recently, there has been considerable research activity on iterative detection on unknown fading channels by using soft-decision feedback [9]–[13] (and references therein), partially inspired by the successful application of APPs/soft decisions in decoding concatenated convolutional codes [14]. This letter relates the idea of iterative soft-decision feedback to an optimality criterion, namely the MAP criterion, in a specific context, and hopefully also provides some confidence in the general area of iterative detection using soft-decision feedback. We also note that a similar receiver using EM for MAP channel estimation has been developed in [15]. Since no explicit Gaussian–Markov channel model is assumed in [15], reestimation of the fading process is more complicated. As a result, a Wiener filter of very high dimension usually needs to be approximated, due to complexity considerations. On the other hand, the Channel EM in this letter accomplishes exact reestimation of the fading process via the efficient recursive Kalman filter. Finally, in this letter, channel estimation is formulated as an estimation of a continuous hidden Markov process. The counterpart of this problem, where the fading process is modeled as a discrete Markov process, is investigated in [16].

The organization of the letter is as follows. In Section II, the channel model is established. The EM algorithm for MAP channel estimation is derived in Section III. In Section IV, simulations are carried out to compare Info EM and Channel EM. Finally, conclusions are given in Section V. We will use the following notational convention in the letter: $E\{\star\}$ means the expected value of $\star$; $\Re\{\star\}$ means the real part of $\star$; $(\star)^H$ means the Hermitian transpose of $\star$; bold letters denote vectors and matrices.

II. System Model

A temporally correlated Rayleigh flat fading process usually can be modeled as the output of a narrowband shaping filter driven by temporally white complex Gaussian noise. In our case, we assume that the shaping filter is an autoregressive (AR) filter of order one. With m Rayleigh flat fading channels (e.g., multiple receive antennas), the equation describing such a model is

$$\mathbf{x}_{n+1} = \mathbf{F}\mathbf{x}_n + \mathbf{u}_n$$

where $\mathbf{F}$ is a m by m matrix whose diagonal entries are the coefficients of the AR-1 shaping filters. $\mathbf{x}_n$ is the spatial signature at time n, whose elements are the corresponding complex

Paper approved by K. Chugg, the Editor for Signal Processing and Iterative Detection of the IEEE Communications Society. Manuscript received March 1, 2001; revised August 30, 2002. This work was supported by CoRe under Grant C 98-07 and Grant cor99-10050. This paper was presernted in part at MilComm'2000, Los Angeles, CA, October 2000.

The authors are with the Center for Wireless Communications, University of California at San Diego, La Jolla, CA 92093 USA (e-mail: yanming@ece.ucsd.edu; brao@ece.ucsd.edu).

Digital Object Identifier 10.1109/TCOMM.2003.820725

fading coefficients of m channels. $\mathbf{u}_n$ is the m-dimension complex white Gaussian vector exciting m shaping filters, whose correlation matrix is $\mathbf{Q}_1$. Though we use a simplified AR-1 filter for easy representation, the EM method we propose below is directly applicable to shaping filters of higher order.

For linearly modulated signals, the received signal vector at the output of m Rayleigh flat fading channels can be expressed as

$$\mathbf{y}_n = \mathbf{x}_n d_n + \mathbf{v}_n, \quad n = 1, \ldots, N$$

where $\mathbf{y}_n$ is the m-dimensional signal vector received at time n and N is the number of symbols in the packet. d_n is the nth transmitted symbol, which can be either a known pilot symbol inserted for the purpose of channel estimation or an information-bearing data/code symbol unknown to the receiver. M-ary phase-shift keying (MPSK) is assumed in this letter, with constellation points lying on the unit circle. $\mathbf{v}_n$ is a vector representing spatially white complex Gaussian noise with spatial covariance matrix $\mathbf{Q}_2 = \eta_0 \mathbf{I}_m$.

III. MAP Estimate of the Channel Using EM

The channel parameters of interest are the vectors $\mathbf{x}_1$ up to $\mathbf{x}_N$. For notational simplicity, we abbreviate the collection of vectors $\mathbf{x}_1$ up to $\mathbf{x}_N$ by $\mathbf{x}_1^N$. This notational convention also applies to quantities such as $\mathbf{y}$ and d. The observations are the measurement vectors $\mathbf{y}_1^N$. The presence of the unknown information symbols d_1^N is the source of complication preventing the use of standard Kalman filtering. However, they form a good candidate for the "missing data," With the identification of the parameters of interest, observations, and missing data, the EM approach can be adapted to derive a MAP estimate of the channel parameters $\mathbf{x}_1^N$. That is, when d_1^N is not known, we can maximize the density function $P(\mathbf{x}_1^N | \mathbf{y}_1^N)$ by iteratively maximizing the cost function $E_{d_1^N | \hat{\mathbf{x}}_1^N, \mathbf{y}_1^N} \{\log[p(\mathbf{y}_1^N, d_1^N | \mathbf{x}_1^N) p(\mathbf{x}_1^N)] | \hat{\mathbf{x}}_1^N, \mathbf{y}_1^N\}$, where $\hat{\mathbf{x}}_1^N$ is the most recent estimate of $\mathbf{x}_1^N$, and the expectation is over the distribution of missing data d_1^N conditioned on ($\hat{\mathbf{x}}_1^N$, $\mathbf{y}_1^N$). Due to the fact that the EM approach is now standard, and due to space limitations, the details of the derivation of the receiver are omitted. Interested readers are referred to [17] for details. The derived iterative receiver using Channel EM involves the expectation and maximization steps, which are as follows.

1) *E step*: Compute the conditional expected value of d_i, or say soft decision s_i, which can be expressed as

$$s_i = E\left\{d_i | \hat{\mathbf{x}}_1^N(k-1), \mathbf{y}_1^N\right\} \tag{1}$$

where $k \geq 1$ is the EM iteration index and $\hat{\mathbf{x}}_1^N(k-1)$ is the estimate of $\mathbf{x}_1^N$ from the $(k-1)$th iteration. More details on this computation are provided later in this section.

2) *M step*: Compute $\hat{\mathbf{x}}_1^N(k)$ from Kalman fixed-interval smoother [18] applied to the following state and measurement equations:

$$\begin{aligned} \mathbf{x}_{n+1} &= \mathbf{F}\mathbf{x}_n + \mathbf{u}_n && (2)\\ \mathbf{y}_n s_n^H &= \mathbf{x}_n + \mathbf{v}_n. && (3) \end{aligned}$$

The relevant equations for fixed-interval smoothing are listed at the end of this section.

Initialization: We initialize $\hat{\mathbf{x}}_1^N(k)$, $k = 0$ with some decent estimate of fading process $\mathbf{x}_1^N$. A more specific initialization method is discussed in Section IV.

Termination: During the kth iteration, we go through two steps, namely, the E step and the M step. If convergence of $\hat{\mathbf{x}}_1^N(k)$ is achieved or the preset number of iterations has been run, the procedure is stopped. Otherwise, the iteration is continued by going back to the E step.

Remarks: In (3), the soft decision, which accounts for probabilities of different hypothesis about the decision, is used to demodulate the received signal. On the other hand, in Info EM, the most likely information sequence is reencoded to demodulate the received signal [8]. In other words, hard-decisioned data/code symbols without any reliability information are used to remove the modulation. In a fast fading environment where the Kalman filter has to average over only a few symbols in reestimating the channel state at a given time, Channel EM using soft decision can be advantageous, because an unreliable decision has small magnitude, and therefore, relatively low impact on the channel estimate. On the other hand, the effect of a wrong hard decision is difficult to average out, because only a limited number of symbols are used in reestimating the channel state. However, in a slow fading environment, Kalman filter can average over many symbols and thus average out the effect of a few wrong hard decisions. So the use of soft decision is not very compelling in a slow fading environment.

EM step details: Below we discuss how to obtain the soft decision needed in the E step. If d_i is a known pilot symbol, then $s_i = d_i$. In the uncoded binary phase-shift keying (BPSK) case, the conditional expected value of an unknown symbol d_i can be computed as follows:

$$\begin{aligned} s_i &= p\left(d_i = 1 | \mathbf{y}_1^N, \hat{\mathbf{x}}_1^N(k-1)\right)\\ &\quad - p\left(d_i = -1 | \mathbf{y}_1^N, \hat{\mathbf{x}}_1^N(k-1)\right) && (4)\\ &= \tanh\left(\frac{2\Re\left\{\hat{\mathbf{x}}_i^H(k-1)\mathbf{y}_i\right\}}{\eta_0}\right). && (5) \end{aligned}$$

In the case of binary convolutionally coded data, the Bahl–Cocke–Jelinek–Raviv (BCJR) algorithm [19] can be used to generate the soft decision of code symbols s_i, by assuming that the true fading process is $\hat{\mathbf{x}}_1^N(k-1)$. The computation of s_i discussed here can be easily extended to MPSK with a higher order alphabet size. In fact, since no specific data structure is assumed in the derivation, in principle Channel EM applies to any coded data. For example, we can combine Channel EM with concatenated convolutional code, which has great potential for future wireless data networks because of its large coding gain. However, because of the huge memory introduced by the scrambler, currently the only known practical way to obtain approximate APPs of code symbols is the ad hoc iterative decoder [14], [20]. So the E step in Channel EM can only be approximated rather than exactly calculated.

At last, from the equation above, we can see that the noise variance η_0 is needed in computing the soft decision. When the noise variance is not known exactly, the EM algorithm can employ the Gauss–Seidel method in the M step to include the estimation of noise variance η_0 [17].

For the M step, the fixed-interval smoother is used, which is well documented in the literature [18]. For easy reference and completeness, the equations relevant to the Kalman fixed-interval smoothing involving forward/backward recursions are listed below, where the EM iteration index $k(\geq 1)$ is omitted for brevity. The forward recursion has the following two update equations:

$$\begin{aligned}\hat{\mathbf{x}}_{j+1|j} =& \mathbf{F}\hat{\mathbf{x}}_{j|j-1} + \mathbf{F}\mathbf{K}_{j|j-1}(\mathbf{K}_{j|j-1}+\mathbf{Q}_2)^{-1} \\ & \times \left(\mathbf{y}_j s_j^H - \hat{\mathbf{x}}_{j|j-1}\right) \qquad (6)\\ \mathbf{K}_{j+1|j} =& \mathbf{F}\mathbf{K}_{j|j-1}\mathbf{F}^H - \mathbf{F}\mathbf{K}_{j|j-1}(\mathbf{K}_{j|j-1}+\mathbf{Q}_2)^{-1} \\ & \times \mathbf{K}_{j|j-1}^H\mathbf{F}^H + \mathbf{Q}_1, \quad 1 \leq j \leq N \qquad (7)\end{aligned}$$

where $\hat{\mathbf{x}}_{j+1|j} = E\{\mathbf{x}_{j+1}|\mathbf{y}_1^j, s_1^j\}$ is the forward prediction of $\mathbf{x}_{j+1}$ and is initialized with $\hat{\mathbf{x}}_{1|0} = \mathbf{0}$. $\mathbf{K}_{j+1|j} = E\{(\mathbf{x}_{j+1} - \hat{\mathbf{x}}_{j+1|j})(\mathbf{x}_{j+1} - \hat{\mathbf{x}}_{j+1|j})^H|\mathbf{y}_1^j, s_1^j\}$ is the covariance matrix of the prediction error, and is initialized with $\mathbf{K}_{1|0} = E\{\mathbf{x}_n\mathbf{x}_n^H\}$. The backward recursion is

$$\begin{aligned}\hat{\mathbf{x}}_{j-1|N} = \mathbf{F}^{-1}\Big[\hat{\mathbf{x}}_{j|j-1} + \Big(\mathbf{I}_m - \mathbf{Q}_1\mathbf{K}_{j|j-1}^{-1}\Big) \\ \times(\hat{\mathbf{x}}_{j|N} - \hat{\mathbf{x}}_{j|j-1})\Big], \quad 2 \leq j \leq N \qquad (8)\end{aligned}$$

where $\hat{\mathbf{x}}_{j|N} = E\{\mathbf{x}_j|\mathbf{y}_1^N, s_1^N\}$ is the fixed-interval smoothed estimate of $\mathbf{x}_j$ generated by the kth EM iteration, i.e., $\hat{\mathbf{x}}_j(k)$. $\hat{\mathbf{x}}_{N|N}$ can be obtained from $\hat{\mathbf{x}}_{N|N} = \mathbf{F}^{-1}\hat{\mathbf{x}}_{N+1|N}$, where $\hat{\mathbf{x}}_{N+1|N}$ can be obtained from the forward recursion.

IV. Simulation

In all simulations below, an initial channel estimate, $\hat{\mathbf{x}}_1^N(0)$, is obtained using a fixed-interval smoothed estimate of the fading processing based only on the pilot symbols. This Kalman filter has a forward and a backward recursion to obtain fixed-interval smoothed estimates of the fading process. In particular, in the forward recursion when a received sample modulated by a known pilot symbol is encountered, this Kalman filter works as the common Kalman filter; when a received sample with unknown modulation is met, the Kalman filter artificially sets noise variance η_0 to be plus infinity, and thus neglects this sample in forming the channel estimate. The forward update equations of this Kalman filter are as follows, where the EM iteration index $k = 0$ is omitted for brevity:

If d_j is a known pilot symbol

$$\begin{aligned}\hat{\mathbf{x}}_{j+1|j} =& \mathbf{F}\hat{\mathbf{x}}_{j|j-1} + \mathbf{F}\mathbf{K}_{j|j-1}(\mathbf{K}_{j|j-1}+\mathbf{Q}_2)^{-1} \\ & \times \left(\mathbf{y}_j d_j^H - \hat{\mathbf{x}}_{j|j-1}\right) \qquad (9)\\ \mathbf{K}_{j+1|j} =& \mathbf{F}\mathbf{K}_{j|j-1}\mathbf{F}^H - \mathbf{F}\mathbf{K}_{j|j-1}(\mathbf{K}_{j|j-1}+\mathbf{Q}_2)^{-1} \\ & \times \mathbf{K}_{j|j-1}^H\mathbf{F}^H + \mathbf{Q}_1 \qquad (10)\end{aligned}$$

otherwise

$$\hat{\mathbf{x}}_{j+1|j} = \mathbf{F}\hat{\mathbf{x}}_{j|j-1} \qquad (11)$$

$$\mathbf{K}_{j+1|j} = \mathbf{F}\mathbf{K}_{j|j-1}\mathbf{F}^H + \mathbf{Q}_1. \qquad (12)$$

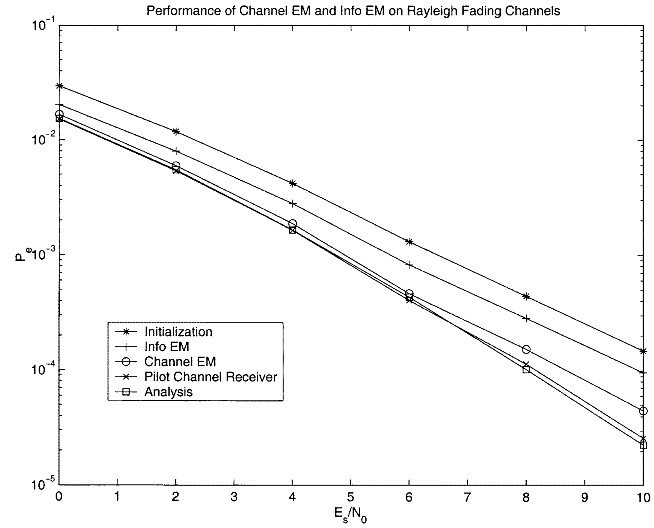

Fig. 1. Comparison of Info EM and Channel EM for uncoded data. BPSK, $f_d = (1/30)$, four antennas, pilot-symbol insertion rate 1/15.

The backward recursion is the same as the common backward recursion in fixed-interval smoothing [c.f. (8)]. Obviously, this Kalman filter gives the optimal linear estimation of the fading process, based only on samples modulated by known pilot symbols. Therefore, this initialization is the same as a Wiener interpolator [21] whose window spans the whole data packet. However, the computation is much faster than a separate Wiener filter for each sample. In all simulations, five iterations are carried out between the Kalman smoother and the demodulator/decoder, which turns out to achieve convergence in terms of bit-error rate (BER) in all cases. For clarity, we only plot the BER performance curve of the initial channel estimate and that of the last iteration. All simulations use BPSK modulation.

Case 1: This is a simulation for uncoded data. Four antennas with identical independent fading are used to provide diversity, i.e., $E\{\mathbf{x}_n\mathbf{x}_n^H\} = E_s\mathbf{I}_4$, where E_s is the symbol energy. The normalized Doppler frequency is 1/30 and the shaping filter is an AR filter of order 1. The information packet size is 510. Known pilot symbols are inserted every 15 information symbols. It can be seen from Fig. 1 that at BER of 10^{-3}, Channel EM (the curve with "o") has a gain of about 0.9 dB over Info EM (the curve with "+").

The channel estimate formed by the iterative receiver (with either hard- or soft-decision feedback) can not be better than the channel estimate from a separate pilot channel of equal power and signal-to-noise ratio (SNR), whose symbols are completely known to the receiver. So, the BER performance of a receiver aided by such a separate pilot channel can be used to lower bound the performance of the iterative receiver. For comparison, we also simulate the performance of such a pilot-channel-aided receiver, which is the curve with "x". In this receiver, first the channel estimate is formed by Kalman smoothing in the pilot channel alone. Then data in the traffic channel is detected based solely on the channel estimate obtained from the separate pilot channel. It can be seen that when E_s/N_0 is less than 8 dB, which corresponds to BER more than 10^{-4}, the performance of Channel EM and that of the pilot receiver are very close. The performance of the receiver aided by a pilot channel can be an-

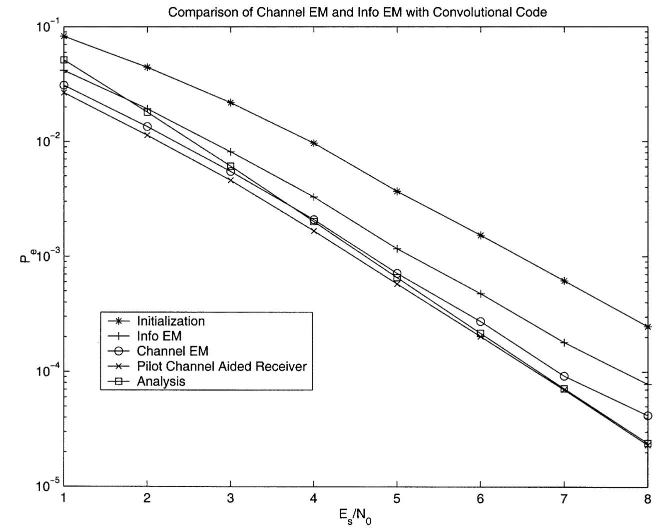

Fig. 2. Comparison of Info EM and Channel EM for convolutional coded data. BPSK, $f_d = (1/40)$, one antenna, pilot-symbol insertion rate 1/20, generator polynomials $[1 + D^2, 1 + D + D^2]$, channel interleaver size 600 × 100.

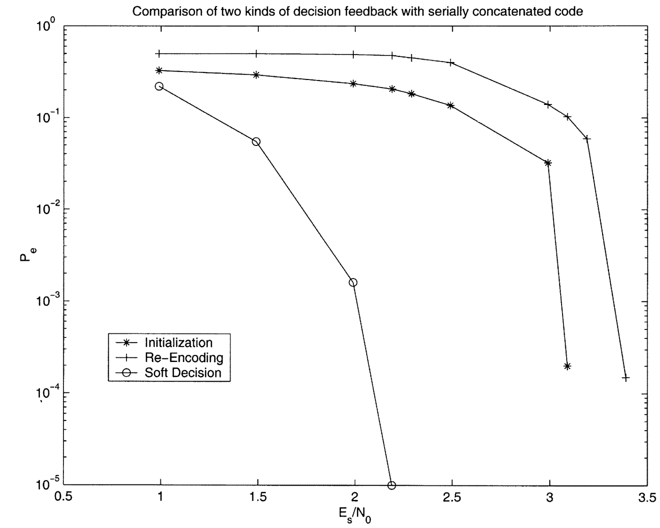

Fig. 3. Comparison of Info EM and Channel EM for serially concatenated code. BPSK, $f_d = (1/40)$, one antenna, pilot-symbol insertion rate 1/20, outer code generator polynomials $[1 + D^2, 1 + D + D^2]$, inner code generator polynomial $[1/(1 + D)]$, channel interleaver size 600 × 100.

alyzed [22] and the analysis result is shown by the curve with "□," which overlaps with the simulation result.

Case 2: This is a simulation for convolutionally coded data (Fig. 2). One antenna is used. The normalized Doppler frequency is 1/40 and the shaping filter is an AR filter of order 1. The code is a rate-1/2 nonrecursive convolutional code with generator polynomials $[1 + D^2, 1 + D + D^2]$ and is terminated in the zero state. The information packet size is 30 000 and the channel interleaver size is 600 by 100. Known pilot symbols are inserted every 20 code symbols. It can be seen that at BER $= 10^{-4}$, Channel EM (the curve with "o") has a gain of about 0.7 dB over Info EM (the curve with "+").

As with uncoded data, we also compare the performance with that of a pilot-channel-aided receiver. In this particular case, the performance of the Channel EM is close to that of the pilot-channel-aided receiver (the curve with "x") when BER is greater than 10^{-3}. In the presence of channel estimation error, the performance of the pilot-channel-aided receiver can be upper bounded by following the approaches in [22] and [23]. The curve with "□" shows such an upper union bound truncated to six terms, which overlaps with the simulation result when BER is less than 10^{-3}.

Case 3: In this last example, we show the combination of Channel EM with serially concatenated convolutional code. In this simulation, the code employed is the same as that in [20]. The outer code is a rate-1/2 nonsystematic convolutional code with generator $[1 + D^2, 1 + D + D^2]$, and the inner code is a rate-1 differential encoder. The outer code is terminated in the zero state, while the inner code is left unterminated. Such a code is shown to outperform the $[1 + D^2, 1 + D + D^2]$ convolutional code because of a large interleaver gain [20]. In the simulation, each packet has 30 000 information bits and the size of the channel interleaver is 600 by 100. The scrambler between the outer encoder and the inner encoder is generated randomly at the beginning of the simulation. One antenna is used, with the normalized fading rate being 1/40 and the order of the shaping filter being one. Known pilot symbols are inserted every 20 code symbols. At the receiver, the number of iterations of the soft-input–soft-output (SISO) decoder is 10 for each iteration between the decoder and the Kalman filter. In Fig. 3, the curve with "o" is obtained via the Channel EM. For comparison, we also simulate a conceptual counterpart of Info EM, where the decoded information sequence is reencoded for reestimating the channel. The curve with "+" shows the performance of reencoding, where the the number of iterations of the SISO decoder is also 10 for each iteration between the decoder and the Kalman filter. At BER of 10^{-3}, the soft-decision feedback approach of the Channel EM can be seen to be about 1 dB better than the reencoding approach. The result seems to parallel that in [10] and [13]. However, the soft-decision feedback approach in this letter is derived more systematically.

V. Conclusion

An iterative receiver with soft-decision feedback is derived by using the EM algorithm for MAP estimation of fast flat fading channels. Simulations indicate that in a fast fading environment, the proposed receiver can outperform an iterative receiver with hard-decision feedback. Finally, we note that the approach is quite general, since no particular trellis structure of the data is assumed in the derivation. So the Channel EM algorithm developed can be easily extended to other scenarios, such as decoding space–time codes and multiuser detection on time-varying Rayleigh fading channels.

Acknowledgment

The authors would like to thank three anonymous reviewers for their detailed review, which greatly improves the presentation of this letter. The authors would also like to thank Prof. P. Siegel and his research group at the University of California at San Diego for helpful discussions on concatenated convolutional codes.

REFERENCES

[1] N. M. Laird, A. P. Dempster, and D. B. Rubin, "Maximum-likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc., Ser. B*, vol. 39, no. 1, pp. 1–38, 1977.

[2] M. A. Tanner, *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions*. New York: Springer, 1996.

[3] C. N. Georghiades and D. L. Snyder, "The expectation-maximization algorithm for symbol unsynchronized sequence detection," *IEEE Trans. Commun.*, vol. 39, pp. 54–61, Jan. 1991.

[4] L. B. Nelson and H. V. Poor, "Iterative multiuser receivers for CDMA channels: An EM-based approach," *IEEE Trans. Commun.*, vol. 44, pp. 1700–1710, Dec. 1996.

[5] G. K. Kaleh and R. Vallet, "Joint parameter estimation and symbol detection for linear or nonlinear unknown channels," *IEEE Trans. Commun.*, vol. 42, pp. 2406–2413, July 1994.

[6] H. Zamiri-Jafarian and S. Pasupathy, "Adaptive MLSDE using the EM algorithm," *IEEE Trans. Commun.*, vol. 47, pp. 1181–1193, Aug. 1999.

[7] ——, "EM-based recursive estimation of channel parameters," *IEEE Trans. Commun.*, vol. 47, pp. 1297–1302, Sept. 1999.

[8] C. N. Georghiades and J. C. Han, "Sequence estimation in the presence of random parameters via the EM algorithm," *IEEE Trans. Commun.*, vol. 45, pp. 300–308, Mar. 1997.

[9] C. Luschi, M. Sandell, and P. Yan, "Iterative channel estimation using soft decision feedback," in *Proc. IEEE GLOBECOM 1998*, Sydney, Australia, Nov. 1998, pp. 3728–3733.

[10] M. C. Valenti and B. D. Woerner, "Iterative channel estimation and decoding of pilot-symbol-assisted turbo codes over flat-fading channels," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 1697–1705, Sept. 2001.

[11] A. Anastasopoulos and K. M. Chugg, "Adaptive iterative detection for phase tracking in turbo-coded systems," *IEEE Trans. Commun.*, vol. 49, pp. 2135–2144, Dec. 2001.

[12] S. Bhashyam and A. Hottinen, "Performance of iterative multiuser decoding and channel estimation in WCDMA systems," in *Proc. IEEE Vehicular Technology Conf. 2000*, vol. 1, Tokyo, Japan, May 2000, pp. 425–429.

[13] J. Heo and M. Chugg, "Adaptive iterative detection for turbo codes on flat fading channels," in *Proc. IEEE Wireless Communications and Networking Conf.*, vol. 1, Chicago, IL, 2000, pp. 134–139.

[14] C. Berrou and A. Glavieux, "Near-optimum error-correcting coding and decoding: turbo codes," *IEEE Trans. Commun.*, vol. 44, pp. 1261–1271, Oct. 1996.

[15] E. Chiavaccini and G. M. Vitetta, "MAP symbol estimation on frequency-flat Rayleigh fading channels via a Bayesian EM algorithm," *IEEE Trans. Commun.*, vol. 49, pp. 1869–1872, Nov. 2001.

[16] C. Komninakis and R. D. Wesel, "Joint iterative channel estimation and decoding in flat correlated Rayleigh fading," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 1706–1717, Sept. 2001.

[17] M. Yan, "Antenna array receiver design and analysis in fast fading environments," Ph.D. dissertation, Dept. of ECE, Univ. of Calif., San Diego, La Jolla, CA, 2001.

[18] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979.

[19] F. Jelinek, L. R. Bahl, J. Cocke, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 284–287, Mar. 1974.

[20] K. R. Narayanan and G. L. Stuber, "A serial concatenation approach to iterative demodulation and decoding," *IEEE Trans. Commun.*, vol. 47, pp. 956–961, July 1999.

[21] 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.

[22] J. G. Proakis, *Digital Communications*, 2nd ed. New York: McGraw-Hill, 1989.

[23] C. Schlegel, "Trellis-coded modulation on time-selective fading channels," *IEEE Trans. Commun.*, vol. 42, pp. 1617–1627, Feb.–Apr. 1994.