EM-Based Sequence Estimation for Wireless Systems with Orthogonal Transmit Diversity

Erdal Panayırcı
Dept. of Electronics Engineering
IŞIK University
Istanbul, 80670, Turkey
Email: eepanay@isikun.edu.tr

Ümit Aygölü
Faculty of Electrical and Electronics Engineering
Istanbul Technical University
Istanbul, 80626, Turkey
Email: aygolu@ehb.itu.edu.tr

Ali Emre Pusane
Dept. of Electrical Engineering
University of Notre Dame
Notre Dame, Indiana 46556
Email: apusane@nd.edu

***Abstract*— In this paper, an optimum sequence estimation algorithm for wireless systems with Alamouti's two transmitter diversity in the presence of multipath fading is proposed. The algorithm is based on a jointly iterative channel and sequence estimation according to the maximum likelihood (ML) criterion, using the Expectation-Maximization (EM) algorithm employing M-PSK modulation scheme with additive Gaussian noise. The discrete multipath channel is represented in terms of the channel gains from each transmit antenna to the receive antenna. EM algorithm derived estimates jointly the complex channel parameters of each channel and the data sequence transmitted, iteratively, which converges to the true ML solution. The channel estimation is achieved in a simple way through the iterative equations by decoupling of the signals transmitted from different antennas. The algorithm is applied to the trellis coded modulation systems and efficiency of the algorithm proposed has been shown by the computer simulations. Simulation results show that the EM algorithm converges quickly for fast fading channels. The performance of the EM-based decoder approaches that of the ML receiver which has perfect knowledge of the channel.**

I. Introduction

Transmitter diversity is an effective technique for combating fading in multipath wireless channels. It has been observed recently that transmitter(spatial) diversity may be the only option when the frequency and time diversity techniques are not always available. Transmit diversity has been studied only recently to reduce the detrimental effects in wireless fading channels because of its relative simplicity of implementation and feasibility of having multiple antennas at the base stations. Several transmit diversity techniques were studied extensively in the past. Wittneben [1] proposed the first bandwidth efficient transmit scheme and subsequently, a delay diversity scheme was introduced by Seshadri and Winters, [2]. More recently, space-time trellis coding has been proposed by Tarokh, Seshadri and Calderbank [3] which combines signal processing at the receiver with coding techniques appropriate to multiple transmit antennas. These so-called space-time codes perform well in slowly-fading channels, assuming perfect channel information(CSI) at the receiver. With the presence of channel mismatch, however, system performance suffers a significant degradation.

This work was supported in part by the Scientific and Technical Research Council of Turkey under grant no. 100E006.

Recently, Alamouti proposed a remarkable transmit diversity scheme for transmission using two transmit antennas, [4]. This scheme has been generalized later in [5], [6] to an arbitrary number of transmit antennas and is able to achieve the full diversity promised by the transmit and receive antennas. Assuming that the channel state information is available somehow, the orthogonal structure of these space-time block codes enables the ML decoding to be implemented in a simple way through decoupling of the signal transmitted from different antennas rather than joint detection. However, channel state information is usually difficult to obtain. In the absence of perfect channel state information, evaluation of the ML function requires the expectation over the joint statistics of the channel fading coefficients, which is usually mathematically intractable. To cope with this technical difficulty, in this paper, we apply the method of Georghiades and Han [7] and use the results of Li,Georghiades and Huang [8] to the sequence and channel estimation for specifically Alamouti's orthogonal space-time coded systems in the presence of multipath fading channels with two-transmitter diversity. The algorithm is based on a jointly iterative channel and sequence estimation according to the ML criterion, using the EM algorithm, [9], [10], [11]. The last part of the paper provides simulation results on the convergence of the EM algorithm. The performance is presented in terms of the bit error rate for a system employing trellis coded 8-PSK signaling. The extensive computer simulations show that a formulation of the sequence estimation based on the EM algorithm is a promising technique for highly efficient data transmission over mobile wireless channels and it performs close to the performance of a maximum likelihood decoder that assumes perfect CSI.

The paper is organized in four sections following this introduction. In Section 2, the system model is introduced, Section 3, includes the EM-based algorithm, Section 4 presents the simulation results and finally conclusions are presented in Section 5.

II. System Model

We consider the wireless communication system as shown in Figure 1 with transmitter diversity using a space-time block coded transmit diversity scheme first proposed by Alamouti, [4]. The scheme is described with 2 transmit and 1 receive

0-7803-7802-4/03/$17.00 © 2003 IEEE

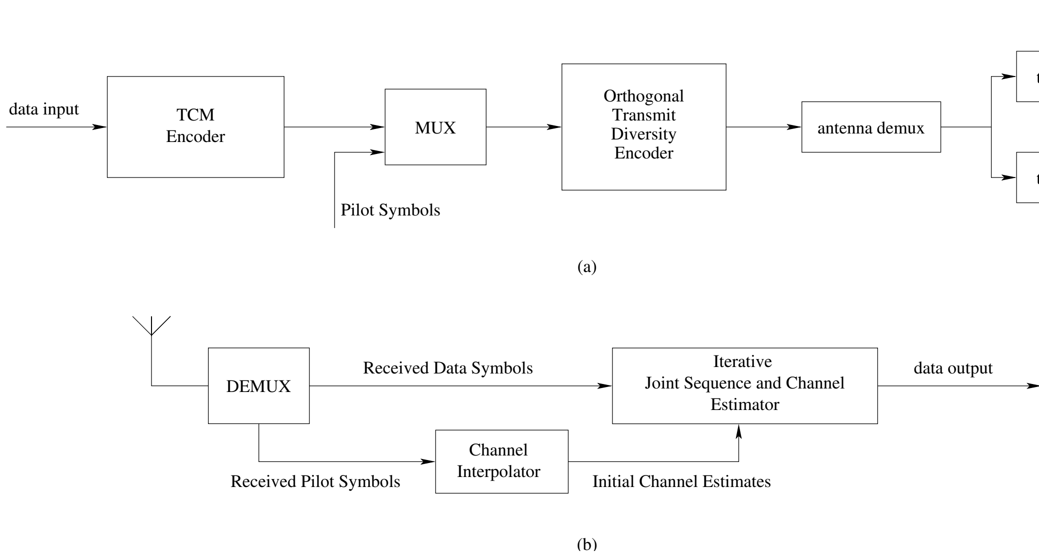

Fig. 1. (a) Transmitter and (b) Receiver block diagrams of the transmit diversity system

antennas to provide a diversity of order 2. Note that, the method can be easily extended to the more general orthogonal space-time block coded systems introduced by Tarokh *et al.*[5] involving more than two transmit and one receive antennas.

The information data can be either uncoded or encoded by a TCM encoder, then fed into the space-time block encoder. At each time slot, the output symbols are modulated and transmitted simultaneously each from a different transmit antenna. At the receiver end, the space-time block decoder followed by symbol-by-symbol decoder or by Viterbi decoder, for uncoded and coded cases, respectively, can be used to decode the received sequence. The generated complex constellation symbols characterizing the input bits are fed to the space-time block encoder proposed by Alamouti whose transmission matrix is given as

$$\begin{array}{cc} & space \rightarrow \\ time \downarrow & \left[\begin{array}{cc} x_1 & x_2 \\ -x_2^* & x_1^* \end{array}\right] \end{array} \tag{1}$$

whose rows are transmitted in successive time intervals with the first and second symbol in a given row sent simultaneously through the first and second antenna, respectively. Based on this scheme, at each time slot k ($k = 0, 1, \cdots, L-1$), the signal transmitted from the first antenna is s_{2k} and the signal transmitted from the second antenna is s_{2k+1}. In the next time slot $(k+1)$, the signal $-s_{2k+1}^*$ is transmitted from the first antenna, and the signal s_{2k}^* is transmitted from the second antenna. Coherent phase-shift keying(PSK) modulation is used here to enhance the system performance.

The wireless channel is assumed to be a fast fading channel where the maximum Doppler spread normalized by the symbol rate is of the order of 10^{-2}. Since we use Alamouti's scheme, it means that channel fading is required to be constant over two consecutive symbol periods($2T$), but varies from one time interval $2T$ to another. Define $\boldsymbol{h}_0 = [h_{0,0}, h_{0,2}, \cdots, h_{0,(2L-2)}]^T$ and $\boldsymbol{h}_1 = [h_{1,0}, h_{1,2}, \cdots, h_{1,(2L-2)}]^T$, where $h_{i,j}$ denotes the channel gains from the first and second transmit antennas to receive antenna, respectively, at the jth symbol period, $j = 0, 2, \cdots, 2L-2$. They are modeled as complex zero-mean Gaussian random variables with autocorrelation $r_l = E[h_{i,2k} h_{i,\ 2k+2l}^*]$, $i = 0, 1$; $l = 0, 1, \cdots, L-1$ and that $\boldsymbol{h}_0$ and $\boldsymbol{h}_1$ are independent of each other. For mobile fading channels, the autocorrelations are given by $r_l = v^2 J_0(2\pi f_d T l)$ where v^2 us the unnormalized variance of the fading gains, $J_0(.)$ is the zero-order Bessel function of the first kind, f_d is the maximum Doppler frequency in Hz and T represents the signaling interval. Thus, for $i = 0, 1$, vector $\boldsymbol{h}_i$ has a normalized Toeplitz covariance matrix $\boldsymbol{R} = (1/v^2)[r_l]$. For $k = 0, 1, \cdots, L-1$, each pair of the two consecutive received signals can then be expressed as

$$\begin{aligned} r_{2k} &= s_{2k} h_{0,2k} + s_{2k+1} h_{1,2k} + n_{2k} \\ r_{2k+1} &= -s_{2k+1}^* h_{0,2k} + s_{2k}^* h_{1,2k} + n_{2k+1} \end{aligned} \tag{2}$$

where n_{2k} and n_{2k+1} are independent samples of an additive Gaussian random variable with variance σ^2, representing the additive white Gaussian noise entering the system.

Letting $\boldsymbol{r} = [\boldsymbol{r}_0^T \;\; \boldsymbol{r}_1^T]^T$ where $\boldsymbol{r}_0 = [r_0, r_2, \cdots, r_{2L-2}]^T$ and $\boldsymbol{r}_1 = [r_1, r_3, \cdots, r_{2L-1}]^T$, (2) can be expressed into a matrix form

$$\boldsymbol{r} = \boldsymbol{S}\boldsymbol{h} + \boldsymbol{n} \tag{3}$$

where, $\boldsymbol{h} = [\boldsymbol{h}_0^T \;\; \boldsymbol{h}_1^T]^T$, $\boldsymbol{n} = [\boldsymbol{n}_0^T \;\; \boldsymbol{n}_1^T]^T$,

$$\boldsymbol{S} = \left[\begin{array}{cc} \boldsymbol{S}_0 & \boldsymbol{S}_1 \\ -\boldsymbol{S}_1^\dagger & \boldsymbol{S}_0^\dagger \end{array}\right] \tag{4}$$

and, $\boldsymbol{S}_0 = \text{diag}\{s_0, s_2, \cdots, s_{2L-2}\}$, $\boldsymbol{S}_1 = \text{diag}\{s_1, s_3, \cdots, s_{2L-1}\}$. $\dagger$ denotes conjugated transpose.

III. Sequence Estimation with EM Algorithm

Now consider the classical problem of estimating data sequence $\boldsymbol{s} = (s_0, s_1, \cdots, s_{2L-1})$ from the observations of received data $\boldsymbol{r} = (r_0, r_1, \cdots, r_{2L-1})$. A ML receiver then performs

$$\max_{\boldsymbol{s}} p(\boldsymbol{r}|\boldsymbol{s}) = \max_{\boldsymbol{s}} E_{\boldsymbol{h}}\left[p(\boldsymbol{r}|\boldsymbol{s},\boldsymbol{h})\right]. \quad (5)$$

Note that evaluation of the likelihood function above requires the expectation over the joint statistics of the random channel parameters $\boldsymbol{h}$, a task that more often is mathematically intractable. Even if the likelihood function can be obtained analytically off line, however, it is invariably a nonlinear function of $\boldsymbol{s}$, which makes the maximization step computationally infeasible in real time. Especially for long and/or coded sequences transmitted over fading channels, the problem of optimum sequence estimation is known to be difficult or intractable. In such cases, an iterative formulation of the sequence estimation problem based on the EM algorithm can provide an implementable solution. The ML estimate $\widehat{\boldsymbol{s}}_{ML}$ is given by

$$\widehat{\boldsymbol{s}}_{ML} = \arg\max_{\boldsymbol{s}} p(\boldsymbol{r}|\boldsymbol{s}). \quad (6)$$

The EM algorithm inductively reestimates $\widehat{\boldsymbol{s}}_{ML}$ so that a monotonic increase in the *a posteriori* conditional pdf above is guaranteed. The monotonic increase is realized via the maximization of the auxiliary function

$$Q(\boldsymbol{s}|\boldsymbol{s}^{(i)}) = E[\log p(\boldsymbol{r}|\boldsymbol{s},\boldsymbol{h})|\boldsymbol{r},\boldsymbol{s}^{(i)}]. \quad (7)$$

Given the received signal $\boldsymbol{r}$, the EM algorithm starts with an initial value $\boldsymbol{s}^{(0)}$ of the unknown channel parameters $\boldsymbol{s}$. The $(i+1)th$ estimate of $\boldsymbol{s}$ is obtained by the maximization step described by

$$\boldsymbol{s}^{(i+1)} = \arg\max_{\boldsymbol{s}} Q(\boldsymbol{s}|\boldsymbol{s}^{(i)}). \quad (8)$$

The log-likelihood function of $\boldsymbol{r}$ given $\boldsymbol{s}$ and $\boldsymbol{h}$ needed in (7) to compute the expectation step is easily obtained from (3) as follows

$$\ell(\boldsymbol{r}|\boldsymbol{s},\boldsymbol{h}) \equiv \log p(\boldsymbol{r}|\boldsymbol{s},\boldsymbol{h}) \sim p(||\boldsymbol{r} - \boldsymbol{S}\boldsymbol{h}||^2)$$

Dropping unnecessary terms and rearranging slightly it follows that

$$\ell(\boldsymbol{r}|\boldsymbol{s},\boldsymbol{h}) = \mathcal{R}e[\boldsymbol{r}^\dagger \boldsymbol{S}\boldsymbol{h}] - \frac{1}{2}||\boldsymbol{S}||^2. \quad (9)$$

Assuming the PSK signaling is used we can drop the second term in the right hand side of (8).

Then, the expectation step of the EM algorithm at the ith iteration yields,

$$\begin{aligned} Q(\boldsymbol{s}|\boldsymbol{s}^{(i)}) =& \mathcal{R}e[\boldsymbol{r}^\dagger \boldsymbol{S}\widehat{\boldsymbol{h}}^{(i)}] \\ =& \sum_{k=0}^{L-1}\Big[\mathcal{R}e\left\{(r_{2k}^* s_{2k} - r_{2k+1}^* s_{2k+1}^*)\widehat{h}_{0,2k}^{(i)}\right\} + \\ & \mathcal{R}e\left\{(r_{2k}^* s_{2k+1} + r_{2k+1}^* s_{2k}^*)\widehat{h}_{1,2k}^{(i)}\right\}\Big] \end{aligned} \quad (10)$$

where

$$\widehat{\boldsymbol{h}}^{(i)} = E[\boldsymbol{h}|\boldsymbol{r},\boldsymbol{s}^{(i)}]. \quad (11)$$

After some algebra, the above conditional mean can be obtained as follows: It can be shown that

$$p(\boldsymbol{h}|\boldsymbol{r},\boldsymbol{s}^{(i)}) \sim \exp\left[-(\boldsymbol{h} - \widehat{\boldsymbol{h}}^{(i)})^\dagger \boldsymbol{\Psi}^{-1}(\boldsymbol{h} - \widehat{\boldsymbol{h}}^{(i)})\right], \quad (12)$$

where,

$$\widehat{\boldsymbol{h}}^{(i)} = (v^2/\sigma^2)\boldsymbol{\Psi}\boldsymbol{S}^{\dagger(i)}\boldsymbol{r},$$

and

$$\boldsymbol{\Psi} = \left(\boldsymbol{R}_{\text{h}}^{-1} + (v^2/\sigma^2)\boldsymbol{I}\right)^{-1}.$$

Here, $\boldsymbol{R}_{\text{h}}$ is a $2L \times 2L$ block diagonal matrix defined by $\boldsymbol{R}_{\text{h}} = \text{diag}\{\boldsymbol{R}\ \ \boldsymbol{R}\}$, where $\boldsymbol{R}$ is the normalized autocorrelation matrix of the random fading vector, as defined earlier, whose main diagonal elements are unity. v^2 is the unnormalized variance of the random fading gains. σ^2 is the variances of the noise.

The EM algorithm starts with an initial estimate of the channel estimates $\{\widehat{h}_{0,2k}^{(0)}, \widehat{h}_{1,2k}^{(0)}\}$ and uses them in (6) to produce, by maximization, a sequence estimate. This sequence estimate is then used in (8) to produce the next sequence estimate, and so on, until convergence within two to three iterations. At convergence, $\boldsymbol{s}^{(i+1)} = \boldsymbol{s}^{(i)}$, the algorithm produces both a sequence estimate and a fading channel estimate.

We now turn to the maximization step of the EM algorithm, where we distinguish between the coded and the uncoded transmission. First we observe from (10) that in the absence of coding, maximizing $Q(\boldsymbol{s}|\boldsymbol{s}^{(i)})$ with respect to sequence $\boldsymbol{s}$ is equivalent to maximizing each individual term in the sum, i.e., making symbol-by-symbol decisions. Then, if $\boldsymbol{s}^{(i+1)}$ is the maximizing sequence, for $k = 0, 1, \cdots, L-1$, its components are given by

$$\begin{aligned} s_{2k}^{(i+1)} &= \arg\max_{s_{2k}} \mathcal{R}e\left\{r_{2k}^* s_{2k}\widehat{h}_{0,2k}^{(i)} + r_{2k+1}^* s_{2k}^*\widehat{h}_{1,2k}^{(i)}\right\} \\ s_{2k+1}^{(i+1)} &= \arg\max_{s_{2k+1}} \mathcal{R}e\left\{-r_{2k+1}^* s_{2k+1}^*\widehat{h}_{0,2k}^{(i)} + r_{2k}^* s_{2k+1}\widehat{h}_{1,2k}^{(i)}\right\} \end{aligned} \quad (13)$$

where we have used the expression for $Q(\boldsymbol{s}|\boldsymbol{s}^{(i)})$ in (10).

When trellis coding is used, the maximization over all trellis sequences can be done efficiently using the Viterbi algorithm. It is seen that in contrast to directly evaluating the likelihood function in (9), the EM algorithm yields at each step of iteration a likelihood function that allows the use of the Viterbi algorithm for efficient computations.

Initialization

In order to be able to choose good initial values for $\boldsymbol{s}^{(0)}$, the N_{PS} data symbols $\{s_{2k}, s_{2k+1}\}$ for $k \in S_{PS}$, in each observation block are generally used as pilot symbols known by the receiver. They are inserted periodically in the sequence. Here, S_{PS} denotes the set of pilot symbols indices. To interpolate the channel estimates, initially, there exist a minimum spacing,

l_{SC}, between pilots given by $l_{SC} < 1/\tau_{max}$, where τ_{max} is the maximum delay spread of the channel ($B_{coh} = 1/\tau_{\max}$, channel coherent bandwidth).

To initialize the receiver we determine $\widehat{h}_{0,2k}^{(0)} = \widehat{\boldsymbol{h}}_0^{(0)}[2k]$ in terms of the pilot symbols and the received signals corresponding to the pilot symbols from the following equations. $\widehat{h}_{1,2k}^{(0)} = \widehat{\boldsymbol{h}}_1^{(0)}[2k], k \in S_{PS}$, where

$$\begin{aligned} \widehat{\boldsymbol{h}}_0^{(0)} &= \boldsymbol{\Psi}_{11}^{(0)}(\boldsymbol{s}_0^{\dagger(0)}\boldsymbol{r}_0 - \boldsymbol{s}_1^{(0)}\boldsymbol{r}_1) + \boldsymbol{\Psi}_{12}^{(0)}(\boldsymbol{s}_1^{\dagger(0)}\boldsymbol{r}_0 + \boldsymbol{s}_0^{(0)}\boldsymbol{r}_1) \\ \widehat{\boldsymbol{h}}_1^{(0)} &= \boldsymbol{\Psi}_{21}^{(0)}(\boldsymbol{s}_0^{\dagger(0)}\boldsymbol{r}_0 - \boldsymbol{s}_1^{(0)}\boldsymbol{r}_1) + \boldsymbol{\Psi}_{22}^{(0)}(\boldsymbol{s}_{\dagger 0}^{(1)}\boldsymbol{r}_0 + \boldsymbol{s}^{(0)}\boldsymbol{r}_1), \end{aligned} \tag{14}$$

and

$$\boldsymbol{\Psi}^{(0)} = \begin{bmatrix} \boldsymbol{\Psi}_{11}^{(0)} & \boldsymbol{\Psi}_{12}^{(0)} \\ -\boldsymbol{\Psi}_{21}^{(0)} & \boldsymbol{\Psi}_{22}^{(0)} \end{bmatrix}.$$

The complete initial channel gains $h_{0,2k}^{(0)}, h_{1,2k}^{(0)}$ for $k = 0, 1, \cdots, L-1$ can be easily determined using an interpolation technique, i.e., Lagrange interpolation algorithm.

The EM algorithm can be summarized briefly as follows:
Step 1. Set $i = 0$ and choose the initial values $\boldsymbol{s}^{(0)}$, and determine $\widehat{\boldsymbol{h}}_0^{(0)}, \widehat{\boldsymbol{h}}_1^{(0)}$, as explained above

Step 2. Compute $\boldsymbol{s}^{(i+1)}$ by maximizing $Q(\boldsymbol{s}|\widehat{\boldsymbol{s}}^{(i)})$ in (8)and (10) over all sequences by Viterbi algorithm if trellis coding is present. Use (13) to perform the maximization if coding is not present.

Step 3. Compute $\widehat{\boldsymbol{h}}_0^{(i+1)}, \widehat{\boldsymbol{h}}_1^{(i+1)}$ from (11) and goto Step 2, repeat until the algorithm converges, in which case the last sequence estimate is produced as the ML estimate.

Note that a computation of the number of iterations needed to implement the EM algorithm indicates that it increases linearly in the sequence length compared to the more than exponential increase for direct implementation. Also, the maximization step in (8) can be implemented easily due to the fact that $Q(\boldsymbol{s}|\boldsymbol{s}^{(i)})$ can be expressed as in recursive form as in (10), and thus, the Viterbi algorithm can be employed.

IV. Simulation Results

Error performance of the proposed iterative decoder has been investigated via computer simulations. The fading channel is modeled as the Jakes fading with autocorrelation $v^2 = 1$, $r_l = J_0(2\pi f_d T l)$ where $J_0(.)$ is the zero-order Bessel function of the first kind, f_d is the maximum Doppler frequency in Hz and T represents the signaling interval. Data bits are first encoded by a rate $2/3$, 4-state 8-PSK TCM encoder to produce the coded data symbol sequence of length 100. The encoder used was recently proposed in [12] and has optimum performance when used in combination with Alamouti's transmit diversity scheme.

In order to initialize the EM algorithm, the receiver has to have good estimates of the channel. These estimates have been provided using pilot symbol assisted modulation (PSAM), [13]. Six pairs of pilot symbols, which are already known at the receiver, are added periodically to the data symbol sequence with a period of 20. At the receiver, channel fading coefficients are first estimated at the pilot symbol positions. The unknown data fading coefficients are then estimated by applying Lagrange interpolation technique on the pilot fading coefficients, according to the initialization procedure as explained in Section 3. The EM algorithm uses these channel estimates to initialize and converge to the maximum likelihood decoding within two or three iterations. The maximization step of the EM algorithm is efficienty performed using the Viterbi algorithm. Bit error probability curves have been presented for a channel with normalized maximum Doppler frequency of 0.01 in Fig. 2.

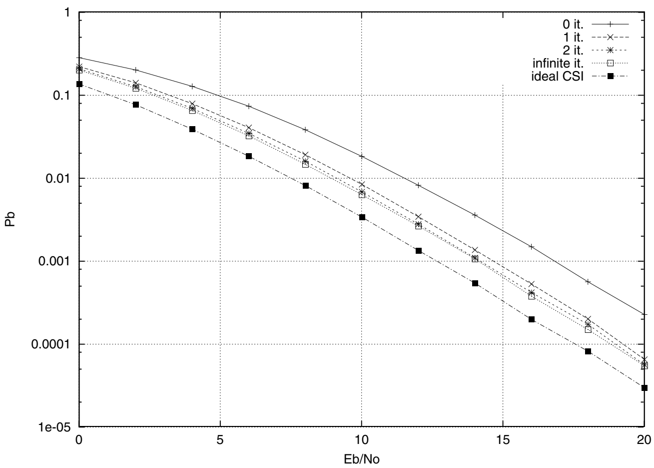

Fig. 2. Bit error performance of trellis coded 8-PSK code for $f_D T = 0.01$

The proposed scheme seems to converge to the ML decoding in two iterations. This provides an SNR gain of $3\,dB$ in the high SNR region. The performance improvement is caused by the reduction in the channel estimation error which can be seen in Fig. 3, where the minimum square estimation error (MSEE) values versus iteration numbers are presented for different SNR values.

The channel estimation errors converge to the maximum likely estimates in two iterations. For a channel with higher Doppler frequency ($f_D T = 0.03$), the bit error probability curves again converge in two iterations (Fig. 4), but this time resulting in an error floor. Alamouti's transmit diversity scheme loses its orthogonality property in the presence of channel estimation error and an error floor is observed.

Since, in the fast fading channel, PSAM with a pilot separation of 20 loses its effectiveness in estimating the channel fading coefficients, the algorithm converges to a local maximum which results in a high estimation error (Fig. 5). In both cases, the proposed decoder is shown to converge to the ML decoding in just two iterations.

V. Conclusion

In this paper, we proposed an optimum sequence estimation algorithm for wireless communications systems employing a

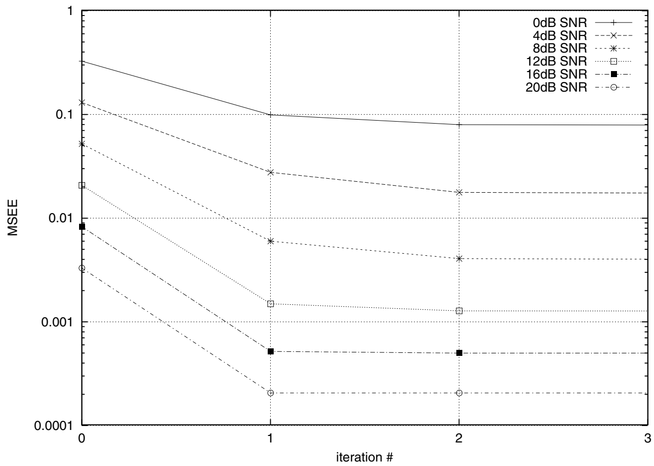

Fig. 3. Mean square estimation error for $f_D T = 0.01$

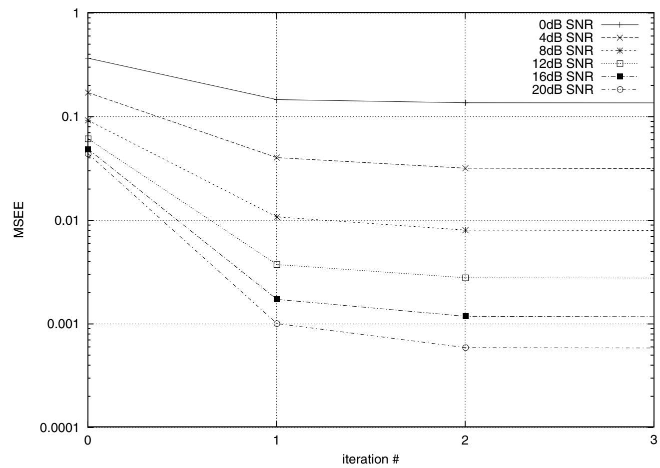

Fig. 5. Mean square estimation error for $f_D T = 0.03$

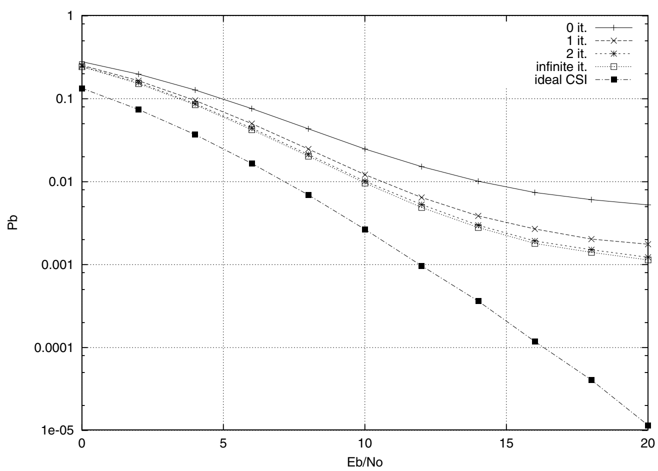

Fig. 4. Bit error performance of trellis coded 8-PSK code for $f_D T = 0.03$

transmit diversity. This algorithm performs an iterative estimation of the transmitted sequence of data symbols according to the ML criterion, using the EM algorithm employing M-PSK modulation scheme with additive Gaussian noise. The discrete multipath channel was represented in terms of the channel gains from each transmit antenna to the receive antenna. EM algorithm derived estimates jointly the complex channel parameters of each channel and the data sequence transmitted, iteratively, which converges to the true ML solution. The algorithm is applied to the trellis coded 8-PSK modulated wireless systems and efficiency of the algorithm proposed has been shown by the computer simulations. Simulation results show that the EM algorithm converges quickly for fast fading channels. The performance of the EM-based decoder approaches that of the ML receiver which has perfect knowledge of the channel. In addition, the EM-based detector is rather simple to implement since the maximization step of the algorithm can be done using the Viterbi algorithm.

REFERENCES

[1] A. Wittneben, "Base station modulation diversity for digital SIMULCAST," *in Proc. IEEE'VTC*, pp. 505-511, May 1993.

[2] N. Seshadri and J. H. Winters,"Two signaling schemes for improving the error performance of frequency-division-duplex (FDD) transmission systems using transmitter diversity," *Int. J. Wireless Inform. Networks*, vol.1, no.1, 1994.

[3] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communications: Performance analysis and code construction,"*IEEE Trans. Inform. Theory*, vol. 44, pp. 744-765, March 1998.

[4] S. M. Alamouti, "A simple transmitter diversity technique for wireless communications,"*IEEE J. Selected Areas Commun.*, vol. 16, pp. 1451-1458, October 1998.

[5] V. Tarokh, Hamid Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs,"*IEEE Trans. Inform. Theory* , vol. 45, pp. 1456-1467, July 1999.

[6] V. Tarokh, Hamid Jafarkhani, and A. R. Calderbank, "Application of orthogonal design to wireless communication,"*Proc. IEEE Information Theory Workshop* , pp. 46-47, Killarney, Ireland,June 1998.

[7] 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. 300-3007, March 1997.

[8] Y. Li, C. N. Georghiades and G. Huang, "Iterative maximum likelihood sequence estimation for space-time coded systems," *IEEE Trans. Commun.*, vol. 49 , pp. 948-951, June 2002.

[9] A. P. Dempster, N. M. Laird and D. R. Rubin, "Maximum-likelihood estimation from incomplete data via the EM algorithm," *J. Roy. Statist. Soc.*, vol. 39, pp. 1-17, 1977.

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

[11] E. Panayı rcı and C. N. Georghiades, " Carrier phase synchronization of OFDM systems over frequency selective channels via EM algorithm," *Proc. IEEE Vehicular Technology Conference (VTC99)*, 1999, Houston, Texas, USA.

[12] Y. Gong and K. B. Letaief, "Analysis and design of trellis coded modulation with transmit diversity for wireless communications," *Proc. IEEE Wireless Communications and Networking Conference (WCNC'2000)*, pp. 1356-1361, 2000, Chicago, Illinois, USA.

[13] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Vehicular Tech.*, vol 40, November 1991.