IMPROVING ERROR PERFORMANCE OF EQUALIZED RECEIVERS FOR BROADBAND RADIO

Maurizio Magarini, Arnaldo Spalvieri, and Guido Tartara

Dipartimento di Elettronica e Informazione, Politecnico di Milano
Piazza Leonardo da Vinci, 32, I-20133 Milano (Italy)
e-mail: spalvier@elet.polimi.it

Abstract

The paper deals with the design of suboptimal receivers for data transmission over frequency selective channels. The complexity of maximum likelihood sequence estimation turns out be exponential in the channel memory. Hence, when dealing with channels with long memory, suboptimal reception must be considered. Among suboptimal receivers, the prefiltered Viterbi detector plays an important role. This receiver consists of a prefilter followed by a Viterbi processor with a number of states lower than that needed for maximum likelihood sequence estimation. Often the parameters of the receiver are optimized according to the minimum mean square error criterion. Our contribution is a stochastic approximation algorithm that optimizes the parameters of the receiver according to a measure of error probability. Simulation results show that our proposed design gives substantial benefits at moderate to high signal to noise ratio.

1 Introduction

The increasing demand for wideband communication services has lead to the development of broadband transmission systems. In the mobile scenario, radio systems are often subject to propagation over multipath channels. The effect of multipath is severe in broadband systems, where the symbol repetition interval may be shorter than the delay between the paths. In this case, one or more deep notches affect the spectrum of the received signal. In modern radio systems two methods are adopted to make the transmission robust against multipath: multicarrier modulation or equalization of the received signal. Multicarrier modulation is based on the idea of dividing the wideband signal in several narrowband signals, the narrowband signal being less sensitive to multipath. Equalization attempts to recover the transmitted data from the received waveform by suitable processing of the received signal. The concern of the present paper is a suboptimal technique for signal equalization. The receiver that guarantees minimum Bit Error Rate (BER) is the maximum a posteriori probability receiver. However, one often renounces to this receiver, because it is too complex. A simpler receiver is obtained if the probability of sequence error is considered. This approach leads to Maximum Likelihood Sequence Estimation (MLSE) [1, 2]. Unfortunately, even MLSE is often too complex. Actually, the MLSE receiver is realized by a Viterbi algorithm with a number of states that is exponential in the channel memory. Hence, when dealing with channels with long memory, one is forced to consider suboptimal receivers. Several architectures of suboptimal receivers have been proposed and studied in the huge literature of channel equalization. The more common and simple is the FIR filter followed by a threshold detector. Better performance can be achieved if memory is introduced in the detector. One popular example is the decision feedback equalizer. A second example is the prefiltered Viterbi detector, which is the object of our investigation. The idea behind the prefiltered Viterbi detector is to introduce a prefilter, which takes the form of a FIR filter, before the conventional Viterbi detector. The prefilter should be designed in such a way that the prefiltered impulse response has shorter memory

than the impulse response of the channel. The signal is then processed by a Viterbi algorithm with fixed complexity. The receiver is suboptimal because the noise present in the prefiltered signal can be colored, and noise coloration is not taken into account in the metric used in the conventional Viterbi algorithm, and because the memory of the prefiltered impulse response can still be too long. Two sets of parameters are designed in the prefiltered Viterbi detector: the taps of the FIR prefilter and the Desired Impulse Response (DIR) used in the computation of the metrics for the Viterbi algorithm. In the design of these parameters one has to optimize the compromise between the detrimental effects of excess of memory in the prefiltered impulse response and noise coloration. The Mean Square Error (MSE) has received in the past large attention as a design criterion [5, 7, 8, 9, 11]. The idea is to adjust receiver' s parameters in such a way that the MSE between the output of the prefilter and the transmitted sequence filtered by the DIR is minimized. Note that the MSE criterion does not guarantee the best error performance. Hence there is room to improve the design method. To this aim, in the following section we introduce a new design criterion based on a measure of error probability that we call Last Error Event (LEE) probability, and a stochastic gradient algorithm that optimizes receiver' s parameters according to this criterion. Our idea is that, by optimizing the probability of LEE, better BER performance is obtained with respect to the MSE receiver. In section 3 simulation results are presented. As expected, our results show that the BER of the receiver based on the LEE is better than the BER of the classical receiver based on the MSE. From our simulations it turns out that substantial improvement can be obtained for a receiver operating on a severely distorted signal at moderate to high SNR. In section 4 conclusions are drawn.

2 Optimization based on the last error event

We consider the model of a binary uncoded data sequence transmitted over a baseband linear channel and corrupted by additive white Gaussian noise. The common model of a baud-spaced FIR filter is adopted for the linear system, hence
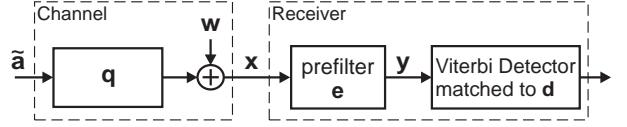

Figure 1: Channel and receiver block diagram.

the observed signal is

$$\mathbf{x} = \tilde{\mathbf{a}} \otimes \mathbf{q} + \mathbf{w}, \qquad (1)$$

where $\tilde{\mathbf{a}} = (\tilde{a}_0, \tilde{a}_1, \ldots, \tilde{a}_l)$ is the transmitted sequence, $\otimes$ indicates the discrete convolution, $\mathbf{q} = (q_0, q_1, \ldots, q_\nu)$ is the impulse response of the FIR channel, and $\mathbf{w}$ is the row vector containing $l + \nu + 1$ samples of white Gaussian noise. Among reduced complexity receivers, we examine the prefiltered Viterbi detector. It is derived from the MLSE receiver, and it is based on the idea of shortening the impulse response $\mathbf{q}$ to a desired impulse response by a prefilter. The block diagram of the system is shown in figure 1, where $\mathbf{d} = (d_0, d_1, \cdots, d_\mu)$, $\mu \leq \nu$, is the DIR, and $\mathbf{e} = (e_0, e_1, \cdots, e_n)$ is the prefilter. Let $\mathbf{y} = \mathbf{x} \otimes \mathbf{e}$ be the output of the prefilter, and consider a truncated version of $\mathbf{y}$ made by $l + \mu + 1$ samples as $\mathbf{y}_D = (y_D, y_{D+1}, \cdots, y_{D+l+\mu})$, where D is the delay between the prefiltered signal sequence and the data sequence. The receiver computes

$$m(\mathbf{a}, \mathbf{y}_D) = (\mathbf{y}_D - \mathbf{a} \otimes \mathbf{d})(\mathbf{y}_D - \mathbf{a} \otimes \mathbf{d})^T, \quad (2)$$

and decides in favour of the sequence that has minimum metric. Since the time spanning of the DIR is $\mu + 1$, the Viterbi algorithm has 2^μ states. The choice of $\mathbf{e}$, $\mathbf{d}$, and D is up to the designer. Often, the Mean Square Error

$$MSE = \lim_{l \to \infty} \frac{1}{l+1}(\mathbf{y}_D - \tilde{\mathbf{a}} \otimes \mathbf{d})(\mathbf{y}_D - \tilde{\mathbf{a}} \otimes \mathbf{d})^T,$$
$$(3)$$

is adopted as a design criterion. In [5, 8] the DIR is fixed, and the prefilter is optimized for minimum MSE. In [7, 9, 11] it is proposed to optimize both the DIR and the prefilter for minimum MSE. To discard the trivial solution $\mathbf{d} = \mathbf{0}$, $\mathbf{e} = \mathbf{0}$, it is proposed in [9] to constrain the energy of the DIR. The receiver of [7, 11] is similar to [9], but the constraint is imposed by fixing one of the samples of the DIR to 1. It is worth noting that, for $\mu = 0$, the MSE criterion leads to the classical minimum MSE linear equalizer followed by an instantaneous threshold detector. Usually, the delay D is optimized by repeated trials. The main weakness of the MSE criterion is
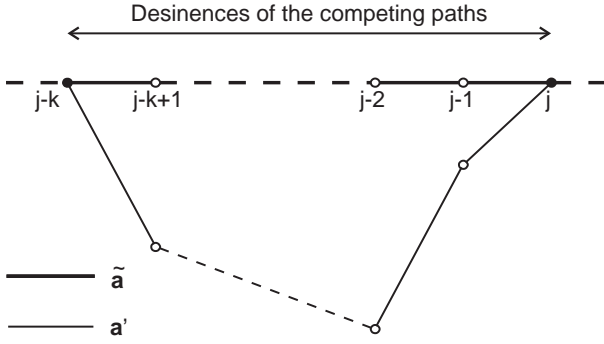
Figure 2: Desinences of $\tilde{\mathbf{a}}$ and $\mathbf{a}'$.

that it does not guarantee low error probability. Hence we propose a new design criterion based on the Last Error Event (LEE). The LEE is similar to the First Error Event (FEE) proposed in [1] as a measure of performance of the sequence detector. As the FEE, the LEE is a conditional event. We introduce the LEE because the conditions for the LEE are simpler to check than the conditions for the FEE, leading to a simpler optimization algorithm. Roughly speaking, the condition for the LEE at time j in the trellis is that the desinence of the transmitted sequence appears among the survivors at time j. In figure 2, the desinences of $\tilde{\mathbf{a}}$ and of a competing path $\mathbf{a}'$ are illustrated. More precisely, the conditions for the LEE at time j are:

- *the two competitors that merge in the state visited by $\tilde{\mathbf{a}}$ at time j diverge from some state visited, say at time $j - k$, by $\tilde{\mathbf{a}}$;*

- *one of the two competitors is the desinence of $\tilde{\mathbf{a}}$, that is $(\tilde{a}_{j-k-\mu}, \tilde{a}_{j-k-\mu+1}, \cdots, \tilde{a}_{j-1})$.*

Given the conditions, the LEE occurs if

- *the desinence of $\tilde{\mathbf{a}}$ loses the competition.*

Note that when $\mu = 0$ the trellis has only one state and only $k = 1$ is allowed. In this case the LEE is the bit error event, and $P(LEE)$ is the BER. We cannot find a closed form for the prefilter and the DIR that minimize the probability of LEE. Therefore we resort to a stochastic approximation technique, based on the difference between the metric of the desinence of the transmitted sequence and the metric of the competitor at time j:

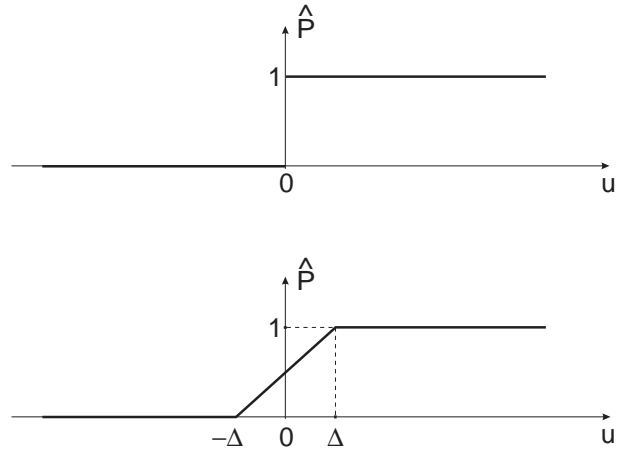$$u^{(j)} = \sum_{l=j-k}^{j-1} (y_{D+l} - \sum_{m=0}^{\mu} d_m \tilde{a}_{l-m})^2$$

Figure 3: Representation of the hard and the smooth estimate.

$$- \sum_{l=j-k}^{j-1} (y_{D+l} - \sum_{m=0}^{\mu} d_m a'_{l-m})^2, \qquad (4)$$

where $(a'_{j-k-\mu}, a'_{j-k-\mu+1}, \cdots, a'_{j-1})$ is the competitor. Given the conditions, a last error event occurs at time j if $u^{(j)} > 0$. Note that a hard estimate of $P(LEE)$ based on u would not be differentiable. Hence we introduce a smooth estimate of $P(LEE)$:

If the conditions for LEE at time j are satisfied, then

$$\hat{P}^{(j)}(LEE) = \begin{cases} 0 & if & u^{(j)} \le -\Delta^{(j)}, \\ \frac{u^{(j)}+\Delta^{(j)}}{2\Delta^{(j)}} & if & -\Delta^{(j)} < u^{(j)} \le \Delta^{(j)}, \\ 1 & if & u^{(j)} > \Delta^{(j)}, \end{cases}$$
$$(5)$$

where $\{\Delta^{(j)}\}$ (the smoothing parameter) is a sequence of real, positive, decreasing numbers. The relation between the smooth and the hard estimate is illustrated in figure 3. The iterative algorithm for minimizing $P(LEE)$ with respect to the generic vector of parameters $\mathbf{p}$ ($\mathbf{p} = \mathbf{e}, \mathbf{d}$ in our contest) is

$$\mathbf{p}^{(j)} = \mathbf{p}^{(j-1)} - \gamma^{(j)} \nabla_{\mathbf{p}} \hat{P}^{(j)}(LEE), \qquad (6)$$

where $\{\gamma^{(j)}\}$ (the step size) is a sequence of real, positive, decreasing numbers. Recalling $\mathbf{y} = \mathbf{x} \otimes \mathbf{e}$, one easily obtains from (4) and (5)

if $-\Delta^{(j)} < u^{(j)} \le \Delta^{(j)}$ then

$$e_i^{(j)} = e_i^{(j-1)} + \frac{\gamma^{(j)}}{\Delta^{(j)}} \sum_{l=j-k}^{j-1} x_{l+D-i}$$

$$\cdot \left(\sum_{m=0}^{\mu} d_m \tilde{a}_{l-m} - \sum_{m=0}^{\mu} d_m a'_{l-m} \right),$$
$$i = 0, \cdots, n, (7)$$

and

$$d_i^{(j)} = d_i^{(j-1)} + \frac{\gamma^{(j)}}{\Delta^{(j)}} \left(\sum_{l=j-k}^{j-1} y_{l+D} \left(\tilde{a}_{l-i} - a'_{l-i} \right) \right.$$

$$\left. -\tilde{a}_{l-i} \sum_{m=0}^{\mu} d_m \tilde{a}_{l-m} + a'_{l-i} \sum_{m=0}^{\mu} d_m a'_{l-m} \right),$$

$$i = 0, \cdots, \mu, (8)$$

else

$$e_i^{(j)} = e_i^{(j-1)}, \quad i = 0, \cdots, n, \qquad (9)$$

$$d_i^{(j)} = d_i^{(j-1)}, \quad i = 0, \cdots, \mu. \qquad (10)$$

It is apparent indeed that the proposed receiver has at least one redundant degree of freedom. As a matter of fact, the same $P(LEE)$ is obtained from all the receivers based on $\alpha \mathbf{e}$, $\alpha \mathbf{d}$, where $\alpha \neq 0$ is a real constant. To discard this redundant degree of freedom, we fix one of the samples of the DIR. Iteration of algorithm (6) guarantees a local minimum of $P(LEE)$ with probability one as $j \to \infty$, if the following conditions on the sequences $\gamma^{(j)}$ and $\Delta^{(j)}$ are satisfied:

$$\sum_{j=1}^{\infty} \gamma^{(j)} = \infty, \ \sum_{j=1}^{\infty} \gamma^{(j)} \Delta^{(j)} \leq \infty, \ \sum_{j=1}^{\infty} \frac{\gamma^{(j)2}}{\Delta^{(j)}} \leq \infty.$$

$$(11)$$

The proposed algorithm can be seen as a variant of the classical stochastic gradient algorithm, but here a smooth version of the stochastic gradient is considered, the smoothing being progressively reduced down to zero during the optimization. The complexity of the algorithm is the same as that of the least mean-square algorithm. Stochastic approximation techniques that have a form similar to (5), (6), (11), have found wide application in neural networks for pattern recognition. Specifically, it is shown in [6] that conditions (11) are sufficient to guarantee a local minimum for a family of stochastic approximation techniques where ours is included. An optimization algorithm based on the bit error probability has been recently derived by other means for the instantaneous detector ($\mu = 0$) in [3, 4]. Note that our receiver is more general than the receiver studied in [3, 4], because we consider the detector with memory.

3 Experimental results

To demonstrate the effectiveness of our criterion we show here the results obtained by computer simulation. Although our method allows to improve the error performance of MSE receivers over all the channels that we considered, we found that, to obtain substantial improvement, a severely distorted channel has to be considered. Specifically, we considered the channels of [3, 7, 10], and a channel having an impulse response that decreases exponentially. Among these, the channels studied in [10] are the most severe, in the sense that they give the lower minimum distance for a fixed length of the impulse response. We focused on the channel with $\nu = 6$. Since this channel has spectral nulls, linear equalization is not adequate. MLSE can be performed with a Viterbi processor of 64 states, and leads to BER= 10^{-3} at SNR= $\sum_k q_k^2/\sigma^2 = 18 \, dB$, where σ^2 is the noise variance. The impulse response of the FIR channel is $Q(z) = 0.176 + 0.316z^{-1} + 0.476z^{-2} + 0.532z^{-3} + 0.476z^{-4} + 0.316z^{-5} + 0.176z^{-6}$. The shape of the impulse response resembles a bell, a shape that is often found in channels from the real world.

To demonstrate the effectiveness of our method, we have to compare it with a rival method. We examined the methods proposed in [7, 9]. Both are based on constrained minimization of MSE, but, as mentioned in the previous section, they differ in the constraint. It should be noted that no clear cut choice can be made among the two methods, because none outperforms the other over all the channels that we considered. However, when the channel [10] is considered, [9] outperforms [7], at least with the receiver parameters that we considered. Hence, in the results to be presented, the rival method is [9]. We choose for the decision delay D the one that gives minimum MSE, calculate the parameters of the MSE receiver as in [9], and use them as initial guess for our optimization method. After the optimization, the BER is measured by a random pattern of $5 \cdot 10^6$ data.

In table 1 we report BER versus SNR considering both MMSE and LEE receivers for DIR length $\mu = 0$, 2, 4, 6, and for prefilter length $n = 14$. The results in the table show that the BER improvement at fixed SNR is higher as the DIR length increases. A clear proof of the superiority of our method is apparent from the table comparing the BER of the two methods with that of MLSE ($\mu = 6$). While our method achieves the performance of MLSE, the MSE method does not. Table 2 reports BER versus SNR for pre-

TABLE 1

SNR (dB)	$\mu = 0$		$\mu = 2$		$\mu = 4$		$\mu = 6$		
	MSE	LEE	MSE	LEE	MSE	LEE	MSE	LEE	MLSE
10	2.4E-1	2.4E-1	2.9E-1	2.6E-1	2.5E-1	2.2E-1	2.5E-1	2.0E-1	1.9E-1
15	2.0E-1	2.0E-1	2.3E-1	1.9E-1	1.3E-1	8.8E-2	6.4E-2	3.5E-2	3.3E-2
20	1.5E-1	1.5E-1	1.7E-1	1.2E-1	2.8E-2	9.6E-3	3.2E-4	1.0E-4	8.0E-5
25	1.0E-1	1.0E-1	1.1E-1	5.7E-2	2.8E-3	6.5E-4	-	-	-
30	8.2E-2	7.7E-2	8.5E-2	3.9E-2	4.6E-4	9.2E-5	-	-	-

Table 1: BER of LEE and MSE for fixed prefilter length $n = 14$.

TABLE 2

SNR (dB)	$n = 6$		$n = 10$		$n = 14$		$n = 20$	
	MSE	LEE	MSE	LEE	MSE	LEE	MSE	LEE
10	2.6E-1	2.2E-1	2.6E-1	2.2E-1	2.5E-1	2.2E-1	2.5E-1	2.2E-1
15	1.5E-1	1.1E-1	1.3E-1	8.9E-2	1.3E-1	8.8E-2	1.1E-1	8.3E-2
20	4.9E-2	2.7E-2	3.2E-2	1.1E-2	2.8E-2	9.6E-3	2.4E-2	8.5E-3
25	1.2E-2	3.7E-3	6.6E-3	1.1E-3	2.8E-3	6.5E-4	3.0E-3	4.3E-4
30	2.6E-3	6.6E-4	1.1E-3	1.5E-4	4.6E-4	9.2E-5	5.1E-4	6.2E-5

Table 2: BER of LEE and MSE for fixed DIR length $\mu = 4$.

filter length $n = 6$, 10, 14, 20, and DIR length $\mu = 4$. This table shows that our method allows to reduce up to ten times the BER of the MSE receiver. It also shows that, at SNR$\geq$ 25dB, the BER of the MSE receiver with $n = 20$ is worse than the BER of the MSE receiver with $n = 14$, and that this pathological behaviour does not affect our receiver. Regarding the optimization procedure, two aspects merit to be considered: the decreasing rule of γ and Δ, and the halting condition. After repeated trials, we found that, in the initial part of the simulation it is convenient to keep these parameters constant. After, we adopt for $\Delta^{(j)}$ and $\gamma^{(j)}$ the following rules:

$$\Delta^{(j)} = \Delta^{(0)} n^{-0.2}, \quad \gamma^{(j)} = \gamma^{(0)} n^{-0.5} \qquad (12)$$

where n represents the number of nonnull updatings at time j. In our simulation we choose $\Delta^{(0)} = 0.5$ and $\gamma^{(0)} = \Delta^{(0)}/500$. To switch from constant to decreasing parameters and to halt the optimization procedure, we adopted the following method. During the adaptation, the BER is measured on successive windows of suitable duration. We denote BER$_{obs}$ the BER measured on the last window, and BER$_{min}$, the minimum BER measured up to the last window, and apply the procedure described in the flow graph in figure 4. The switching and the halting conditions are satisfied only when the measured BER does not decrease for K successive measures. To work

out the results reported in the tables, $K = 5$ has been used, while the duration of the window depends on the BER. Convergence examples of the algorithm for window duration $50 \cdot 10^3$, prefilter length $n = 14$, DIR memory $\mu = 4$, and SNR= 20 dB are shown in figure 5.

4 Conclusion

A new design method for the prefiltered Viterbi detector has been introduced, and its advantage over the minimum-MSE receiver of [9] has been demonstrated by computer simulation. Our results show that the BER achieved by the proposed receiver outperforms the BER achieved by the classical MSE method, and that, when a severely distorted channel is considered, the BER can be improved by our method of a factor ranging from 2 to 10 at moderate to high SNR. One important issue, that is not pursued in the present paper, is globability of the minimum. In [4] sufficient conditions are stated for globality of the minimum for the memoryless detector. Unfortunately, the detector with memory complicates the analysis of the cost function. To our experience, the method seems to be robust, in the sense that different initial guesses lead to the same optimal parameters. Another important issue is the application of the method to real receivers. To this purpose, we recall that the
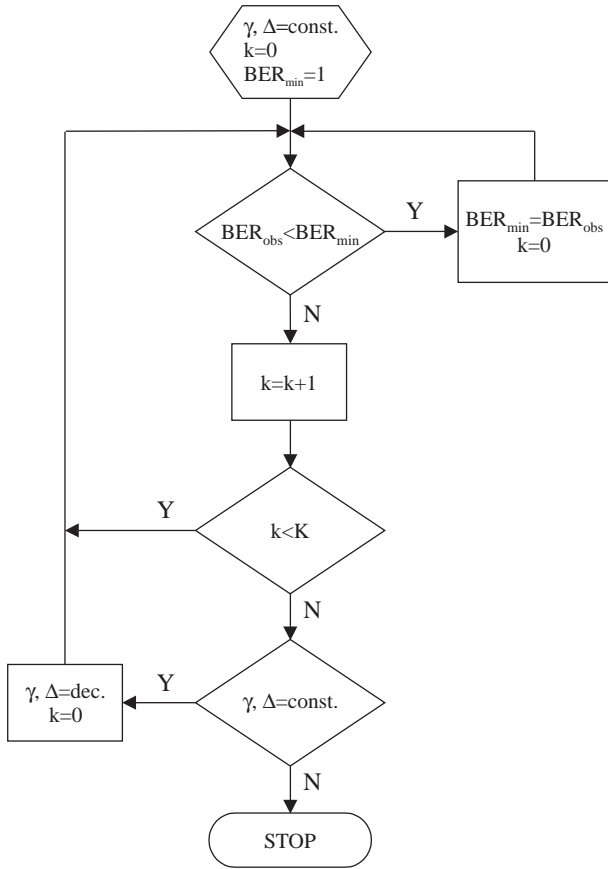
Figure 4: Flow chart for the switching and the halting conditions.

method is based on the knowledge of the impulse response of the channel, which can be estimated from the received signal either by blind methods or by a known data sequence. Once the channel has been estimated, one can locally generate a random data sequence and a random noise sequence to optimize receiver's parameters in an off-line manner by the described algorithm. Of course, since the optimization requires a large number of samples, the receiver must be able to perform fast off-line processing to track rapidly varying channels.

5 Acknowledgements

References

[1] G. D. Forney, "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 363-378, May 1972.
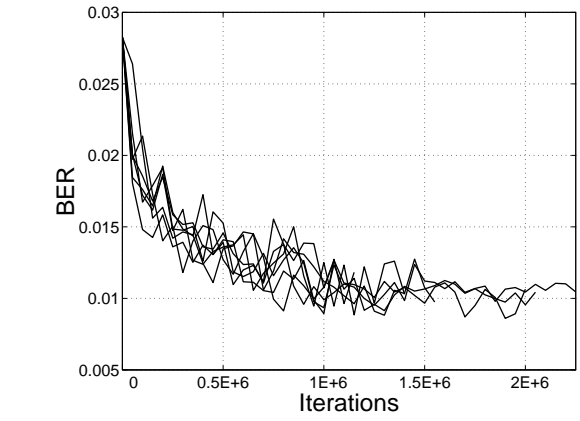

Figure 5: BER versus number of iterations for the LEE algorithm (7 trials).

[2] G. Ungerboeck, "Adaptive maximum-likelihood receiver for carrier modulated data-transmission systems," *IEEE Trans. Commun.*, vol. COM-22, pp. 624-636, May 1974.

[3] C. Yeh and J. R. Barry, "Approximate minimum bit-error rate equalization for binary signaling," *IEEE Int. Conf. Commun.*, June 1997.

[4] C. Yeh and J. R. Barry, "Adaptive minimum bit-error rate equalization for binary signaling," submitted to *IEEE Trans. Commun*, Jan. 1999.

[5] S. H. Qureshi and E. E. Newhall, "An adaptive receiver for data transmission over time-dispersive channels," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 448-457, July 1973.

[6] V. Nedeljkovic, "A novel multilayer neural networks training algorithm that minimizes the probability of classification error," *IEEE Trans. Neural Networks*, vol. NN-4, pp. 650-659, July 1993.

[7] I. Lee and J. M. Cioffi, "Design of equalized maximum-likelihood receiver," *IEEE Commun. Lett.*, vol. 2, pp. 14-16, Jan. 1998.

[8] C. T. Beare, "The choice of the desired impulse response in combined linear-Viterbi algorithm equalizers," *IEEE Trans. Commun.*, vol. COM-26, pp. 1301-1307, Aug. 1978.

[9] D. D. Falconer and J. F. R. Magee, "Adaptive channel memory truncation for maximum likelihood sequence estimation," *Bell Syst. Tech. J.*, vol. 52, pp. 1541-1562, Nov. 1973.

[10] R. R. Anderson and G. J. Foschini, "The minimum distance for MLSE digital data systems of limited complexity," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 544-551, Sept. 1975.

[11] D. G. Messerschmitt, "Design of a finite impulse response for the Viterbi algorithm and decision feedback equalizer," *IEEE Int. Conf. Commun.*, June 1974.