Performance Evaluation of the Mean-Square Prefiltered Delayed Decision Feedback Sequence Detector

M. Magarini, A. Spalvieri, and G. Tartara,
Dipartimento di Elettronica e Informazione, Politecnico di Milano,
Piazza Leonardo da Vinci, 32, I-20133 Milano (Italy).

Abstract— **In signal equalization, a technique that allows reduction of the number of states of the Viterbi detector is the Delayed Decision Feedback Sequence Detector (DDFSD). In order to achieve good performance, it is essential to operate, before the DDFSD, an appropriate prefiltering of the received sequence. This paper is devoted to performance evaluation of the DDFSD when the feedforward filter of a minimum mean square error decision feedback equalizer is adopted as a prefilter. A truncated version of the union bound is used to approximate the bit error rate. The analysis includes a method for determining the error events that dominate the bound.**

I. Introduction

The Delayed Decision Feedback Sequence Detector (DDFSD) is an equalization scheme based on a sampled matched filter, a prefilter, and a Viterbi algorithm where the channel memory is truncated. The performance loss due to memory truncation is mitigated by a per-survivor processing [1], where the past history of each survivor is used in a DFE scheme. In the DDFSD originally proposed in [2], the front-end was the Whitened Matched Filter (WMF) of [3]. In [4] it was proposed to adopt the FIR feedforward filter of a Minimum Mean Square Error Decision Feedback Equalizer (MMSE-DFE). When no restrictions are imposed on the number of taps of the FIR, the front-end turns out to be the Mean Square Whitened Matched Filter (MSWMF) of [5]. In [6] it was shown that, without complexity reduction in the Viterbi algorithm, the MSWMF leads to Maximum Likelihood Sequence Detection (MLSD) with minimum number of states, and that, when the DDFSD is considered, the MSWMF allows to improve over the WMF. In this paper, performance evaluation of the MSWMF-DDFSD is addressed. The outline of the paper is as follows. In section II the system model and the MSWMF-DDFSD are described. Section III is devoted to performance evaluation. In section IV, the accuracy of the approximation is demonstrated by comparing it to simulation results.

II. System Description

We consider the model of a binary uncoded data sequence transmitted over a baseband linear channel corrupted by additive white Gaussian noise. The received signal is passed through the front-end filter and is detected by the DDFSD. The block diagram of the system is reported in Fig. 1.

Let $2\nu + 1$ be the time spanning of the impulse response of the system from the source to the output of the sampled matched filter, that is the sampled autocorrelation of the impulse response $g(t)$ represented in Fig. 1, and let $r(z) = \sum_{i=-\nu}^{\nu} r_i z^{-i}$ be its z-transform (z^{-1} represents the unit delay). The receiver is based on what is called in [5] the *key equation*:

$$d(z)d(z^{-1}) = r(z) + \sigma^2, \tag{1}$$

taking for $d(z)$ that impulse response that is causal and minimum phase. In (1), σ^2 is the two-sided power density spectrum of the white noise $w(t)$. The Signal to Noise Ratio is SNR$= r_0/\sigma^2$. Note that for $\sigma > 0$ the power density spectrum $r(e^{j\omega}) + \sigma^2$ is nonnull everywhere. As a consequence, the existence of $d(z)$ is guaranteed and all its roots are strictly inside the unit circle, thus the existence of $d^{-1}(z)$ is guaranteed as well.

The DDFSD is a Viterbi algorithm with 2^μ states, $\mu \leq \nu$, where, in each state, the branch metric is calculated using a DFE with $\nu - \mu$ taps [2]. Specifically, the metric of the transition that diverges at time $k-1$ from state $(a_{k-\mu}, \ldots, a_{k-1})$ and merges at time k in state $(a_{k-\mu+1}, \ldots, a_k)$ is

$$b_k(a_{k-\mu}, \ldots, a_k) = (x_k - \sum_{j=0}^{\mu} d_j a_{k-j} - \sum_{j=\mu+1}^{\nu} d_j \hat{a}_{k-j}(a_{k-\mu}, \ldots, a_{k-1}))^2,$$

where x_k is the k-th sample at the output of the prefilter, $a_k \in \{-1, +1\}$, and $\hat{a}_{k-j}(a_{k-\mu}, \ldots, a_{k-1})$ is the estimate of the bit transmitted at time $k-j$, present in the survivor that at time $k-1$ merges in the state $(a_{k-\mu}, \ldots, a_{k-1})$. The DDFSD was introduced in [2] using the WMF as a front-end. Here, the front-end is the MSWMF, which consists of the sampled matched filter and of the mean-square prefilter, that is the filter that minimizes the MSE, defined as

$$MSE = E\{u_k^2\}. \tag{2}$$

In (2), $E\{\cdot\}$ denotes the expected value, and

$$u_k = x_k - \sum_{j=0}^{\nu} d_j \tilde{a}_{k-j}, \tag{3}$$

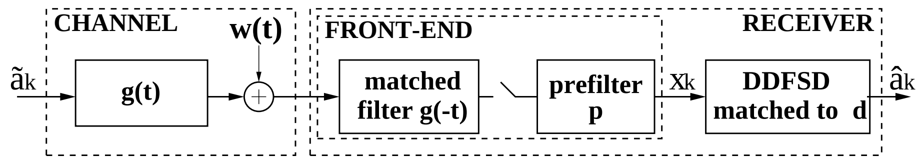

Fig. 1. Channel and receiver block diagram.

is the distortion sequence, where $\tilde{a}_k$ is the k-th element of the transmitted sequence. Minimization of (2) yields [7]

$$p(z) = \frac{d(z)}{r(z)+\sigma^2} = d^{-1}(z^{-1}). \tag{4}$$

Note that, when $\sigma > 0$, $d(z)$ is invertible, hence the existence of $p(z)$ is guaranteed. It is worth noting that, for $\mu = \nu$ the receiver is the MLSD [6], while for $\mu = 0$ one writes the equations of the MMSE-DFE [5] as

$$d_{MMSE}(z) = \frac{d(z)}{d_0}, \tag{5}$$

$$p_{MMSE}(z) = \frac{d_{MMSE}(z)}{r(z)+\sigma^2} = \frac{p(z)}{d_0}, \tag{6}$$

where the notation $[f(z)]_0 = f_0$ has been adopted.

III. Performance Evaluation

Performance evaluation is carried out by using the analysis developed in [8], where, neglecting error propagation, the BER (Bit Error Rate) is approximated as:

$$BER \approx \sum_{e(z)\in\mathcal{E}_M} w_e 2^{-w_e} P(\tilde{a}(z) \mapsto \tilde{a}(z)+e(z)). \tag{7}$$

In (7) $e(z)$ is the input error event, $\mathcal{E}$ is the set of error events having the form $(\ldots, \nu - zeros, e_0, \ldots, e_{l-1}, \mu - zeros, \ldots)$ ($\mathcal{E}_M$ is the subset of $\mathcal{E}$ that contains the M events that dominate the sum) and $w_e = [e(z^{-1})e(z)]_0/4$ is the Hamming weight of the error event. In the trellis of the DDFSD, the error event diverges at time -1 ($e_0 \neq 0$), and merges for the first time at time $l+\mu-1$ ($e_{l-1} \neq 0$), hence the error polynomial has no more than $\mu-1$ consecutive zeros between 0 and $l-1$. In (7) $P(\tilde{a}(z) \mapsto \tilde{a}(z)+e(z))$ is the error probability in the binary test between $\tilde{a}(z)$ and $\tilde{a}(z)+e(z)$, which is hereafter called *pairwise* error probability. Hereafter two issues are considered: computation of $P(\tilde{a}(z) \mapsto \tilde{a}(z)+e(z))$ and search for the error events that form the subset $\mathcal{E}_M$.

A. Computation of the Pairwise Error Probability

The pairwise error probability is the probability of error in the binary test

$$\sum_{k=0}^{l+\mu-1}(x_k - \sum_{j=0}^{\nu} d_j \tilde{a}_{k-j})^2 \gtrless \sum_{k=0}^{l+\mu-1}(x_k - \sum_{j=0}^{\nu} d_j(\tilde{a}_{k-j}+e_{k-j}))^2.$$

From the geometrical perspective, the decision boundary is a hyperplane between the two points that represent $\tilde{a}(z)$ and its competitor. The direction that joins the two mentioned points is hereafter called *output* error, and is represented by the polynomial

$$e_o(z) = [e(z)d(z)]_0^{l+\mu-1}, \tag{8}$$

where the notation $[x(z)]_i^{i+j} = \sum_{k=i}^{k=i+j} x_k z^{-k}$ is adopted. Note that, in contrast to MLSD, here the time spanning of the output error is reduced from $l+\nu$ to $l+\mu$. The squared Euclidean distance between the competitors is

$$\delta_e^2 = [e_o(z)e_o(z^{-1})]_0. \tag{9}$$

Using the z-transform, the binary test takes the form

$$[u(z)u(z^{-1})]_0 \gtrless [(u(z)-e_o(z))(u(z^{-1})-e_o(z^{-1}))]_0, \tag{10}$$

where $u(z)$ is the z-transform of the distortion sequence (3) in the decision space:

$$u(z) = [x(z) - \tilde{a}(z)d(z)]_0^{l+\mu-1}.$$

The binary test (10) is rewritten as

$$\frac{[e_o(z^{-1})u(z)]_0}{\delta_e} \gtrless \frac{\delta_e}{2}. \tag{11}$$

The LHS of (11) is the projection of the distortion along the output error. The error occurs when such a projection, which is called

$$\phi = \frac{[e_o(z^{-1})u(z)]_0}{\delta_e},$$

exceeds half the Euclidean distance between the competitor sequences. The pairwise error probability is

$$P(\tilde{a}(z) \mapsto \tilde{a}(z)+e(z)) = \int_{\delta_e/2}^{\infty} f_\phi(x)dx, \tag{12}$$

where $f_\phi(x)$ is the probability density function of ϕ. The calculation of $f_\phi(x)$ proceeds by considering the distortion as the sum of InterSymbol Interference (ISI) and noise. Specifically, the projection of the noise along the output error is

$$\zeta = \frac{[e_o(z^{-1})n(z)p(z)]_0}{\delta_e},$$

where $n(z)$ is zero mean Gaussian noise with autocorrelation $\sigma^2 r(z)$. The probability density function $f_\zeta(x)$ is Gaussian, with mean

$$m_\zeta = 0, \tag{13}$$

and variance

$$\sigma_\zeta^2 = \frac{\sigma^2}{\delta_e^2}[e_o(z^{-1})p(z)r(z)p(z^{-1})e_o(z)]_0. \quad (14)$$

The projection of the ISI along the output error is

$$\begin{aligned}\psi &= \frac{1}{\delta_e}[e_o(z^{-1})(\tilde{a}(z)(r(z)p(z)-d(z)))]_0 \\ &= -\frac{\sigma^2}{\delta_e}[e_o(z^{-1})p(z)\tilde{a}(z)]_0 = [c(z)\tilde{a}(z)]_0, \quad (15)\end{aligned}$$

where $c(z)$ is the polynomial of the coefficients of the ISI. The probability density function $f_\psi(x)$ can be computed from the coefficients of the ISI. In the section devoted to the experimental results, we adopt the method [9]. Since ISI and noise are independent random variables, the probability density function of ϕ is

$$f_\phi(x) = f_\psi(x) \otimes f_\zeta(x), \quad (16)$$

where $\otimes$ denotes the convolution.

B. Search for the Error Events that Dominate the Union Bound

In performance evaluation, one should select those error events that dominate the sum, and to compute (16) and (12) only for the selected error events. To an efficient selection, one has to establish a sensible figure of merit. One such figure is the SDR (Signal to Distortion Ratio) relevant to $e(z)$, which is defined as

$$SDR_e = \frac{(\delta_e - 2m_\phi)^2}{4\sigma_\phi^2}. \quad (17)$$

Actually, at intermediate-to-high SNR, the sum (7) will be dominated by the terms corresponding to the error events at lower SDR. Our main result is that the SDR of the MSWMF-DDFSD can be written in the form

$$SDR_e = \frac{\delta_e^2 - 4\sigma^2 w_e}{4\sigma^2}. \quad (18)$$

The derivation of (18) is in the appendix. The beauty of (18) is that the denominator is independent of $e(z)$, therefore the algorithms that are used for the search of the error events that dominate the performance of MLSD and of the WMF-DDFSD, where the SDR has the form (19), can be applied to the MSWMF-DDFSD as well. Elaborating upon the specific algorithm is out of the scope of the present paper, hence we consider the popular algorithm reported in [10] for the search of the error event at minimum SDR in MLSD. The algorithm can be applied only to noncatastrophic channels, and works as follows. Let

$$SDR_{min} = \min_{e(z)\in\mathcal{E}}\{\frac{\delta_e^2 - 4\sigma^2 w_e}{4\sigma^2}\}$$

be the minimum SDR. The algorithm is based on the observation that the search for the error sequence that leads to SDR$_{min}$ is the search for the error sequence at minimum squared distance from the all-zeros error, here the squared distance being the numerator of (18). Since the error is ternary, such a search can be operated by a Viterbi algorithm with 3^ν states and three branches diverging from and merging in each state. Applying the numerator of (18) to this trellis, the metric of the branch that diverges at time $k-1$ from state $(e_{k-\nu}, \ldots, e_{k-1})$ and merges at time k in state $(e_{k-\nu+1}, \ldots, e_k)$ is

$$b_k(e_{k-\nu}, \ldots, e_k) = (\sum_{j=0}^{\nu} d_j e_{k-j})^2 - \sigma^2 e_k^2.$$

The transitions that diverge from state $\mathbf{0} = \nu - zeros$ are deleted from the trellis, excepting the first step. Note that, since $e(z)$ and $-e(z)$ have the same squared distance, the search can be limited to the set of error events that begin with $e_0 = -2$, hence only one transition diverges from state **0** at the first step. Then, at each step in the trellis, the metrics of the sequences that merge in state **0** are compared to the minimum metric up to that step, and the lower one is kept as the minimum squared distance found up to that step. The algorithm terminates when **0** is the unique state visited by the survivors.

We have determined the M first terms by extending this algorithm. The extension consists in allowing M survivors per state. At each step in the trellis, three groups each consisting of M parallel transitions merge in each state. The metrics of the $3M$ transitions are sorted and the M sequences with lower metric are kept as survivors. When the algorithm terminates, the metrics of the sequences that merge in state **0** are sorted, and the sequences corresponding to the M lower metrics are kept as the sequences at lower SDR.

In the DDFSD, the error event terminates with a sequence of μ zeros. (Recall that error propagation is neglected.) Hence the algorithm is easily adapted to the DDFSD by deleting from the complete trellis all the transitions that diverge at any time $k > 0$ from each one of the $3^{\nu-\mu}$ states of the type $(e_{k-\nu+1}, \ldots, e_{k-\mu}, \mu - zeros)$. The search for the M lower metrics is now carried out in the $3^{\nu-\mu}$ mentioned states. Note that, due to the memory of the per-survivor DFE, the number of states of the complete trellis cannot be reduced.

IV. Experimental Results

To substantiate the results obtained in the previous section, we adopt as a benchmark the time discrete white Gaussian channel with $\nu = 6$ studied in [11]. The spectrum $r(e^{j\omega})$ is depicted in Fig. 2 versus angular frequency ω. The results obtained for the MSWMF are compared to those obtained with the WMF originally proposed in [2] as a front-end. For the WMF, it should be noted that the only contribution to the distortion is the white Gaussian noise, and that ISI is absent. The SDR for the WMF is

$$SDR_e = \frac{\delta_e^2}{4\sigma^2}, \quad (19)$$

where δ_e^2 is computed from (8) and (9) using the spectral factorization $r(z) = d(z)d(z^{-1})$. Fig. 3 reports the ratio between the minimum SDR of the MSWMF, and the minimum SDR of

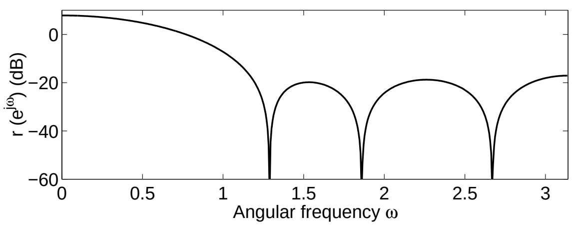

Fig. 2. Spectrum of the channel. $[r(z)]_0^\nu = 0.9978 + 0.9185z^{-1} + 0.7304z^{-2} + 0.4881z^{-3} + 0.2674z^{-4} + 0.1112z^{-5} + 0.031z^{-6}$.

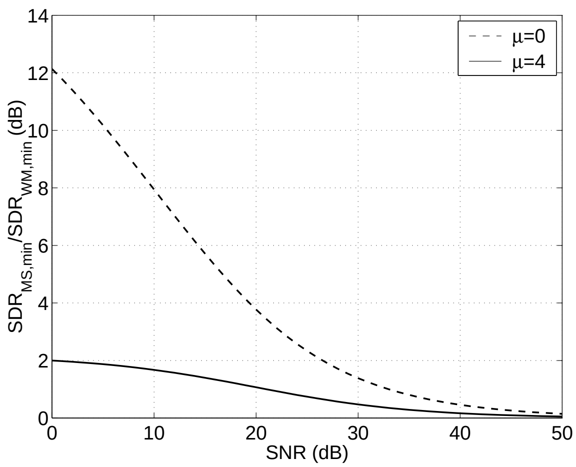

Fig. 3. Ratio between the minimum SDR of the MSWMF and the minimum SDR of the WMF versus SNR for $\mu = 0$ and $\mu = 4$.

the WMF, versus SNR for $\mu = 0$ and $\mu = 4$. The figure shows that the improvement offered by the MSWMF-DDFSD over its competitor is higher at low-to-intermediate SNR. Actually, as $\sigma^2 \to 0$ in (1), the MSWMF tends to the WMF.

Fig. 4 reports the BER versus SNR for $\mu = 4$. In the simulations, the BER is measured by a random sequence of 10^7 data. The figure shows that the approximation is fairly accurate. To fit the simulation results, we find that the first 18 terms contribute to the sum (7). The 18 input error sequences found for SNR$= 20dB$ are listed in table I. The number of error sequences that contribute in the approximation (7) should be determined according to the specific channel and to μ.

V. Conclusions

The main result of the present paper is the performance evaluation of the MSWMF-DDFSD, which is made feasible by equation (18). The method of [10], originally proposed in the context of MLSD, has been adapted to the search for input error events that dominate the union bound on the first error probability. The BER is then approximated by truncating the sum that appears in the union bound, and by attaching their Hamming weights to the error events. Computer simulations show that the accuracy of the proposed approximation is fairly good.

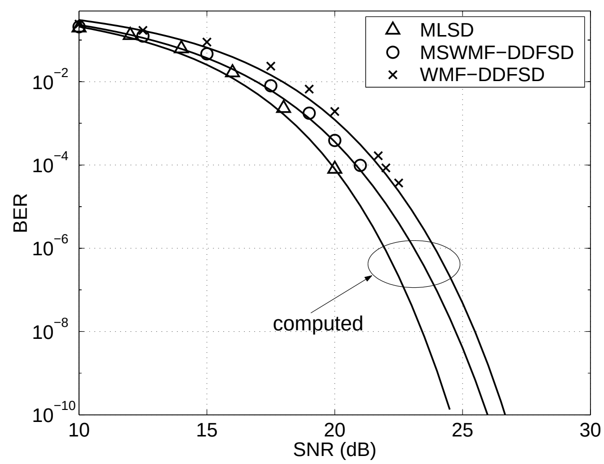

Fig. 4. BER versus SNR for reduction of the Viterbi algorithm to 16 states ($\mu = 4$).

VI. Appendix

The objective of the appendix is to manipulate

$$SDR_e = \frac{(\delta_e - 2m_\phi)^2}{4\sigma_\phi^2}.$$

The mean and the variance of ϕ are:

$$m_\phi = m_\zeta + m_\psi, \tag{20}$$

and

$$\sigma_\phi^2 = \sigma_\zeta^2 + \sigma_\psi^2. \tag{21}$$

The mean and the variance of ζ have been computed in the previous subsection $3.A$. To compute the mean and the variance of ψ, note that the ISI is anticausal and its coefficients can be written in the form

$$c(z) = -\frac{\sigma^2 e_o(z^{-1})p(z)}{\delta_e} = -\frac{\sigma^2(q(z) + e(z^{-1}))}{\delta_e}, \tag{22}$$

where

$$q(z) = \sum_{k=-\infty}^{-l} q_k z^{-k}.$$

The mean value of ψ is computed by taking the expected value of (15) over $\tilde{a}(z) \in \mathcal{A}_e$:

$$m_\psi = E_{\mathcal{A}_e}\{[c(z)\tilde{a}(z)]_0\} = [c(z)E_{\mathcal{A}_e}\{\tilde{a}(z)\}]_0. \tag{23}$$

Since $E_{\mathcal{A}_e}\{\tilde{a}(z)\} = -e(z)/2$, one has

$$m_\psi = -\frac{[c(z)e(z)]_0}{2}. \tag{24}$$

Substituting (22) in (24) one finds

$$m_\psi = \frac{2\sigma^2 w_e}{\delta_e}. \tag{25}$$

TABLE I

FIRST 18 INPUT ERROR EVENTS FOR $\mu = 4$, SNR$= 20dB$. THE ALGORITHM TERMINATES AT THE 47-TH STEP. ONLY THE 9 POLYNOMIALS BEGINNING WITH $e_0 = -2$ ARE LISTED.

SDR (dB)	Coefficients of $e(z)$
10.11	-2 2 2 -2 -2 2
10.89	-2 2 2 -2 -2 2 2 2 -2
11.16	-2 2
11.38	-2 2 0 0 0 -2 2
11.54	-2 2 2 -2 -2 2 2 2 -2 -2 2
11.59	-2 2 0 0 0 -2 2 0 0 0 -2 2
11.79	-2 2 0 0 0 -2 2 0 0 0 -2 2 0 0 0 -2 2
11.98	-2 2 0 0 0 -2 2 0 0 0 -2 2 0 0 0 -2 2 0 0 0 -2 2
12.00	-2 2 2 -2 -2 2 0 2 -2 -2 2 2 2 -2

Substituting (13) and (25) in (20) one has

$$m_\phi = \frac{2\sigma^2 w_e}{\delta_e}.$$

For σ_ψ^2 one writes

$$\sigma_\psi^2 = E_{\mathcal{A}_e}\{(\psi - m_\psi)^2\} = E_{\mathcal{A}_e}\{([c(z)(\tilde{a}(z) + \frac{e(z)}{2})]_0)^2\}. \tag{26}$$

The computation of the expected value is easily performed by the change of variable

$$a'_k = \tilde{a}_k + \frac{e_k}{2} = \tilde{a}_k(1 - \frac{|e_k|}{2}). \tag{27}$$

Note that $a'_k = 0$ where $e_k \neq 0$, and a'_k is purely random where $e_k = 0$. Substituting (27) in (26), and taking into account (22), one writes

$$\sigma_\psi^2 = \frac{\sigma^4}{\delta_e^2} E_{\mathcal{A}'_e}\{([q(z)a'(z)]_0 + [e(z^{-1})a'(z)]_0)^2\}.$$

Looking at (27) one realizes that $[e(z^{-1})a'(z)]_0 = 0$, hence

$$\sigma_\psi^2 = \frac{\sigma^4}{\delta_e^2} E_{\mathcal{A}'_e}\{([q(z)a'(z)]_0)^2\} = \frac{\sigma^4}{\delta_e^2}[q(z)q(z^{-1})]_0, \tag{28}$$

where the latter equality results from the fact that a'_k is purely random for $k = l, \ldots, \infty$. The variance of ϕ is found by noting that equation (28) can be written as

$$\sigma_\psi^2 = \frac{\sigma^4}{\delta_e^2}([e_o(z^{-1})p(z)p(z^{-1})e_o(z)]_0 - 4w_e). \tag{29}$$

Substituting (14) and (29) in (21) one gets

$$\sigma_\phi^2 = \frac{\sigma^2}{\delta_e^2}([e_o(z^{-1})p(z)(r(z) + \sigma^2)p(z^{-1})e_o(z)]_0 - 4\sigma^2 w_e). \tag{30}$$

From (1) and (4) one realizes that

$$p(z)(r(z) + \sigma^2)p(z^{-1}) = 1. \tag{31}$$

Using (31) in (30) one gets

$$\sigma_\phi^2 = \sigma^2(1 - \frac{4\sigma^2 w_e}{\delta_e^2}). \tag{32}$$

Substituting (32) and (25) in the SDR (17), one finds (18):

$$SDR_e = \frac{\delta_e^2 - 4\sigma^2 w_e}{4\sigma^2}.$$

REFERENCES

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

[2] A. Duel-Hallen and C. Heegard, "Delayed decision-feedback sequence estimation," *IEEE Trans. Commun.*, vol. 37, pp. 428-436, May 1989.

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

[4] W. H. Gerstacker, and J. B. Huber, "Improved equalization for GSM mobile communications," in *Proc. Int. Conf. Telecommun. (ICT 96)*, pp. 128-131, Istanbul, Apr. 1996.

[5] J. M. Cioffi, G. P. Dudevoir, M. V. Eyuboglu, and G. D. Forney Jr., "MMSE decision-feedback equalizers and coding-part I: equalization results, " *IEEE Trans. Commun.*, vol. 43, pp. 2582-2594, Oct. 1995.

[6] M. Magarini, A. Spalvieri, and G. Tartara, "Comparison between two methods for delayed decision feedback sequence estimation," in *Proceedings of PIMRC 2000*, vol. 2, pp. 1300-1304, London, U.K., Sept. 2000.

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

[8] W. Sheen and G. D. Stüber, "Error probability for reduced-state sequence estimation," *IEEE J. Select. Areas Commun.*, vol. SAC-10, pp. 571-578, Apr. 1992.

[9] J. C. Vanelli and N. M. Shehadeh, "Computation of bit error probability using the trapezoidal integration rule," *IEEE Trans. Commun.*, vol. 22, pp. 331-334, March 1974.

[10] E. A. Lee and G. D. Messerschmitt, " Digital Communication," *Kluwer*, pp. 277-278, 1993.

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