EUROPEAN COOPERATION
IN THE FIELD OF SCIENTIFIC
AND TECHNICAL RESEARCH

EURO-COST

COST 273 TD(02)032
Guildford, UK
January 17-18, 2002

SOURCE: Dipartimento di Elettronica e Informazione
Politecnico di Milano

Reduced Complexity Sequence Detection

M. Magarini and A. Spalvieri
Dipartimento di Elettronica e Informazione
Politecnico di Milano
Piazza Leonardo da Vinci, 32
I-20133 Milano (Italy)
Phone: +39 02 2399 3450
Fax: +39 02 2399 3413
Email: magarini@elet.polimi.it, spalvier@elet.polimi.it

Reduced Complexity Sequence Detection

Maurizio Magarini and Arnaldo Spalvieri,
Dipartimento di Elettronica e Informazione, Politecnico di Milano,
Piazza Leonardo da Vinci, 32, I-20133 Milano (Italy).

Abstract— **The paper deals with the design of suboptimal receivers for data transmission over frequency selective channels. The complexity of the optimum detector, that is the maximum likelihood sequence detector (MLSD), turns out be exponential in the channel memory. Hence, when dealing with channels with long memory, suboptimal receiver structures must be considered. Among suboptimal methods, a technique that allows reduction of the complexity is the delayed decision feedback sequence detector (DDFSD). This receiver is based on a Viterbi processor where the channel memory is truncated. The memory truncation is compensated by a per-survivor decision feedback equalizer. In order to achieve good performance, it is crucial to operate an appropriate prefiltering of the received sequence before the DDFSD. Our contribution is to extend the principles of MLSD and DDFSD to the case where the prefilter is the feedforward filter of a minimum mean-square error decision feedback equalizer (MMSE-DFE). Moreover performance evaluation of the MMSE prefiltered DDFSD is addressed. The union upper bound is used to evaluate the probability of first-event error. Simulation results show that our proposed design of the MMSE-DDFSD gives substantial benefits when a severe frequency selective channel is considered.**

I. Introduction

In digital mobile radio channels time-varying multipath propagation can cause severe performance degradation. For high-speed data transmissions the effect of multipath is that of introducing intersymbol interference (ISI). Equalization of the received signal is necessary to mitigate the effects of ISI and noise. For channels with large delay spread the optimum equalization algorithm, that is, the maximum likelihood sequence detector (MLSD) [1], has a very high complexity. In the MLSD receiver the number of states of the Viterbi algorithm is exponential in the channel memory. Hence, when dealing with channels with long memory, one is forced to consider suboptimal sequence detectors.

Several architectures of reduced complexity sequence detectors have been proposed and studied in the huge literature of channel equalization [2]. Most of earlier works [3, 4, 5] concentrate on preprocessing techniques to shape the overall impulse response of the channel to a desired one of shorter length. In these solutions, a linear equalizer is employed before the Viterbi algorithm with a prefixed number of states. The idea behind the prefiltered Viterbi detector is to introduce an equalizer, which takes the form of a FIR filter, before the conventional Viterbi detector. The equalizer should be designed in such a way that the overall 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 equalized signal can be colored, and noise coloration is not taken into account in the metric used in the conventional Viterbi algorithm. The criterion that has found widespread use in the design of the equalizer is the minimization of the mean square error (MSE) between the output of the equalizer and the transmitted sequence filtered by the desired impulse response (DIR). In the extreme case where the length of the DIR is 1, this solution coincides with the classical MMSE linear equalizer.

Another possible solution for reducing the complexity of MLSD lies in simplifying the Viterbi algorithm itself. This leads to the delayed decision feedback sequence detector (DDFSD) of [6]. In this scheme, a per-survivor decision feedback is introduced in the branch metric computation for each state. This introduces error propagation whose effects are drastically reduced with respect to decision feedback equalizer (DFE). As we will see, when an appropriate prefiltering is adopted, DDFSD provides the best performance/complexity trade-off.

The paper is organized as follows. In section II the system model is described. In section III a class of MLSDs is introduced. In section IV the MMSE-DDFSD is described. We also considered the application of the MMSE-DDFSD to the generalized Viterbi algorithm. Performance evaluation of the MMSE-DDFSD is addressed in section V: the union upper bound on the probability of first-event error of the DDFSD is presented. In the section of experimental results the accuracy of the approximation is demonstrated by comparing it to simu-

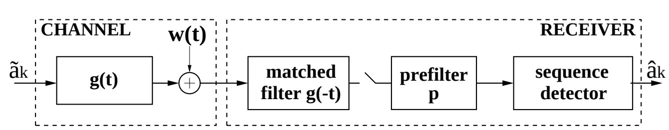

Fig. 1. Channel and receiver block diagram.

lation results. Finally, conclusions are drawn.

II. System Model

We consider the model of a binary uncoded data sequence transmitted over a baseband linear channel corrupted by zero-mean additive white Gaussian noise. The receiver consists of the sampled matched filter, a prefilter, and a sequence detector. The block diagram of the system is reported in Fig. 1. With reference to the figure, we assume in the following that $\tilde{a}_k \in \{-1, +1\}$, and that the two-sided power spectral density of the AWGN is σ^2. The impulse response of the system from the source to the output of the sampled matched filter is represented by the z-transform $r(z) = \sum_{k=-\nu}^{\nu} r_k z^{-k}$, where z^{-1} represents the unit delay.

III. Maximum Likelihood Sequence Detection with Minimum Number of States

The class of sequence detectors that we consider is based on the spectral factorization

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

where $s(z)$ is up to the designer. It must be such that $r(z) + s(z)$ is the z-transform of an autocorrelation function, hence $r(e^{j\omega}) + s(e^{j\omega}) \geq 0$ everywhere. The prefilter is

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

The existence of the prefilter is guaranteed if $r(e^{j\omega}) + s(e^{j\omega}) > 0$ everywhere. In this case, $d(z)$ has no roots on the unit circle, hence $d(z^{-1})$ is invertible. Note that the existence of the cascade of sampled matched filter and prefilter is subject to broader conditions. For example, if $s(z) = 0$ and $r(e^{j\omega}) = 0$ in a null-measure interval, the prefilter, that is the noise whitening filter, does not exist, while the whitened matched filter exists ([1]).

Detection is based on a Viterbi algorithm with $2^{max\{\nu,\nu'\}}$ states and $2^{max\{\nu+1,\nu'+1\}}$ transitions, where $2\nu' + 1$ is the time spanning of $s(z)$. Since we are interested in MLSD with minimum number of states, we impose the condition $\nu' \leq \nu$. The metric of the transition that diverges from state $(a_{k-\nu}, \ldots, a_{k-1})$ at time $k-1$ and merges in state $(a_{k-\nu+1}, \ldots, a_k)$ at time k is

$$b_k(a_{k-\nu}, \ldots, a_k) = (y_k - \sum_{j=0}^{\nu} d_j a_{k-j})^2 - a_k(s_0 a_k + 2\sum_{j=1}^{\nu} s_j a_{k-j}), \tag{3}$$

where y_k is the kth sample at the output of the prefilter and, without loosing generality, $d_k = 0$, $k = \ldots, -2, -1, \nu+1, \nu+2, \ldots$. is assumed. Members of this class are:

- Forney' s detector [1], where $s(z) = 0$,
- Ungerboeck's detector [7], where $s(z) = 1 - r(z)$,
- the MMSE-MLSD proposed in [8], where $s(z) = \sigma^2$.

Note that Barbosa' s detector [9], which is based on $p(z) = r^{-1}(z)$ and has minimum number of states, does not belong to the class. The main result of this section is the following :

Theorem. A Viterbi detector based on (1), (2), and (3), performs MLSD with minimum number of states.

The proof of the theorem is given in appendix of [8].

Hereafter, and throughout the paper, among the possible choices of $s(z)$ in (1), we consider only $s(z) = 0$ and $s(z) = \sigma^2$ that respectively correspond to the whitened matched filter (WMF) and to the mean-square whitened matched filter (MSWMF) [11] models for the receiver. In the WMF model $n(z)$ is a zero mean white Gaussian noise of variance σ^2, while in the MSWMF mean and variance are the same, but now the distortion is given by the sum of colored Gaussian noise and residual ISI.

As we can see in (3), in order to perform branch metric computation it is fundamental to know the DIR $d(z)$. Therefore, an important issue is the estimation of the channel. This latter estimate is used by the receiver in order to realize the matched filter and to calculate the spectral factorization (1). In this work we focus on receivers structure, hence in what follows we assume that channel is estimated through some types of adaptive algorithm.

IV. Delayed Decision Feedback Sequence Detection

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). The DDFSD is a Viterbi algorithm with 2^μ states, $0 \leq \mu \leq \nu$ where the performance loss due to memory truncation is mitigated by a per-survivor processing [10]. The metric of each survivor is calculated using a DFE with $\nu - \mu$ taps. Unfortunately, deriving the best DDFSD from the broad class of MLSD described previously is a hard task, because no general guidelines are given about the choice of $s(z)$. The DDFSD was originally proposed in [6] for the WMF, that is $s(z) = 0$. In consideration of the success that the MMSE-DFE has had since the classical papers of Monsen [12] and Salz [13], we feel that it calls for the MMSE-DDFSD.

A. *The MMSE-DDFSD*

The MMSE-DDFSD is based on complexity reduction of the specific MLSD with $s(z) = \sigma^2$, which yields the spectral factorization

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

where that $d(z)$ that is minimum phase is taken. Note that we now require that $d(z)$ be minimum phase, while in MLSD this is not required. Actually, the minimum phase property guarantees that the energy is concentrated in the first taps of the impulse response, which is a desirable property in a DFE scheme. The prefilter is

$$p(z) = \frac{d(z)}{r(z) + \sigma^2} = d^{-1}(z^{-1}). \quad (5)$$

Of course, $\sigma > 0$ guarantees the existence of both $d(z)$ and $p(z)$. 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

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

where $\hat{a}_{k-j}(a_{k-\mu}, \ldots, a_{k-1})$ is the estimate of the $(k-j)$-th bit which is present in the survivor that at time $k-1$ merges in the state $(a_{k-\mu}, \ldots, a_{k-1})$. Note that, for binary transmission, the term $-\sigma^2 a_k^2$ is common to all the metrics and can be omitted. Let

$$u(z) = v(z)p(z) - d(z)\tilde{a}(z),$$

be the z-transform of the distortion sequence. It is shown in [4] that the prefilter given in (5) minimizes the expected value of u_k^2, for any given $d(z)$. After straightforward manipulation, for the z-transform of the autocorrelation of the distortion sequence one finds

$$E\left\{u(z)u(z^{-1})\right\} = \sigma^2. \quad (6)$$

How we can observe, the distortion sequence is white. For this reason, the cascade of matched filter and mean-square prefilter is called in [11] mean-square whitened matched filter. We emphasize that the two extreme cases of the MMSE-DDFSD are MLSD with minimum number of states for $\mu = \nu$, and the MMSE-DFE for $\mu = 0$.

B. *Mean Square Generalized Delayed Decision Feedback Sequence Detector*

A possible way to improve the performance of the MMSE-DDFSD is that of adopting the scheme of the generalized Viterbi algorithm (GVA) [14], leading to what we call mean square generalized DDFSD (MMSE-GDDFSD). The GDDFSD is obtained by allowing M survivors for each state. Hence in the GDDFSD there are 2^μ states and $2M$ transitions diverging from and merging in each state. At each step in the trellis, the metrics of the $2M$ survivors that merge in each state are sorted in ascending order, and the sequences associated to the M lower metrics are selected as survivors. In the section of experimental results we will show that through a suitable choice of M and μ the MMSE-GDDFSD allows to get a further degree of freedom in the trade-off between complexity and performance.

V. Performance Evaluation of the DDFSD

In what follows, it is shown that the first-event error rate (FEER) of the DDFSD can be upperbounded by the union bound. The approximation to the bit error rate (BER) is then obtained as a by-product from the upper bound on the FEER. Roughly speaking, the FEER is a measure of the probability of error burst. More precisely, the FEER is a conditional probability. Without loosing generality, we assume that the event is

$$\hat{a}_0 - \tilde{a}_0 \neq 0,$$

and that the condition is

$$[\hat{a}(z) - \tilde{a}(z)]_{-\nu}^{-1} = 0, \quad (7)$$

where $\hat{a}_k$ is the k-th decision of the sequence detector. The union upper bound on the FEER is [15]:

$$\text{FEER} \leq \sum_{e(z) \in \mathcal{E}} 2^{-w_e} P(e). \quad (8)$$

In (8), $\mathcal{E}$ is the set of error polynomials having the form

$$e(z) = [\hat{a}(z) - \tilde{a}(z)]_0^{l-1}, \quad l = 1, 2, \ldots, \quad (9)$$

where $e_0 \neq 0$, $e_{l-1} \neq 0$, and there are no more than $\mu-1$ consecutive zeros between 0 and $l-1$, and w_e is the Hamming weight of $e(z)$,

$$w_e = \frac{[e(z)e(z^{-1})]_0}{4}.$$

The probability $P(e)$ appearing in (8), which is called pairwise error probability, is the error probability in the binary test between $[\tilde{a}(z)]_0^{l-1}$ and $[\tilde{a}(z)]_0^{l-1} + e(z)$. Note that the binary test may take place only if the two events (7) and

$$[\hat{a}(z) - \tilde{a}(z)]_l^{l+\mu-1} = 0 \quad (10)$$

occur. Actually, in the DDFSD, the binary test between $[\tilde{a}(z)]_0^{l-1}$ and $[\tilde{a}(z)]_0^{l-1} + e(z)$ takes place when the two competitors merge for the first time in the reduced trellis, that is when (10) is fulfilled. It is worth noting that in MLSD the merging condition is $[\hat{a}(z) - \tilde{a}(z)]_l^{l+\nu-1} = 0$, which is exactly the condition (7) for the first-event error at the next step. Conversely, for $\mu < \nu$, when the binary test between $[\tilde{a}(z)]_0^{l-1}$ and $[\tilde{a}(z)]_0^{l-1} + e(z)$ takes place, there is at least one nonzero coefficient, namely $e_{l-1} \neq 0$, and at most $\nu - \mu$ nonzero coefficients in the polynomial $[\hat{a}(z) - \tilde{a}(z)]_{l+\mu-\nu}^{l+\mu-1}$. Hence, when the decision is $[\hat{a}(z)]_0^{l-1} =$

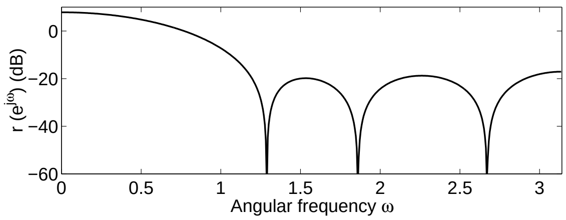

Fig. 2. Spectrum of the channel: $[r(z)]_0^\nu = 0.9978 + 0.9185z^{-1} + 0.7304z^{-2} + 0.4881z^{-3} + 0.2674z^{-4} + 0.1112z^{-5} + 0.031z^{-6}$.

$[\tilde{a}(z)]_0^{l-1} + e(z)$, the condition (7) for the first-event error is not satisfied at the next step after the binary test.

When MLSD is considered, the BER is upperbounded by attaching the Hamming weight of the error polynomial to each term in the sum (8) [1]:

$$\text{BER} \leq \sum_{e(z)\in\mathcal{E}} w_e 2^{-w_e} P(e). \quad (11)$$

Recall that, in MLSD, the condition for the first-event error is satisfied at the first step after a binary test. Roughly speaking, this means that when the sequence detector takes the decision $[\hat{a}(z)]_0^{l-1} = [\tilde{a}(z)]_0^{l-1} + e(z)$, the error burst terminates, enabling the condition for the next burst. In consideration of this fact, the upper bound on the BER is derived from the upper bound on the FEER by counting the number of errors in the burst, that is w_e. Conversely, for $\mu < \nu$, the condition (7) is not satisfied at the next step after the wrong decision, therefore the upper bound (11) on the BER does not hold true. In other words, the error burst may not terminate when the sequence detector takes the wrong decision, hence the number of errors contained in the burst may be not w_e.

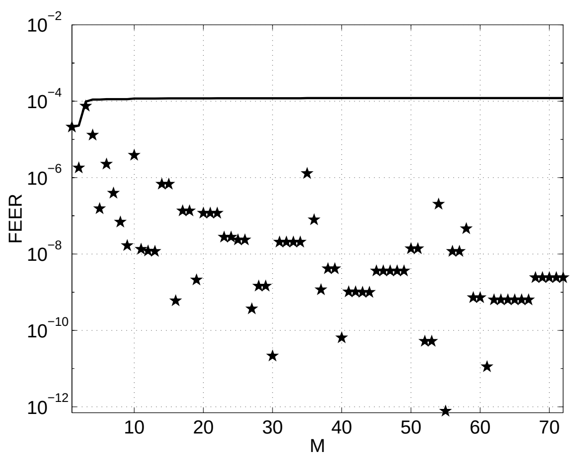

Fig. 3. The solid line is the union bound truncated to the first $2M$ error polynomials versus M, ordered by SDR. The computer simulation gives FEER $= 1.07\ 10^{-4}$. The star is the contribution of the M-th pair of polynomials $e(z)$, $-e(z)$, to the sum.

To overcome this difficulty, in [16] it is considered an ideal DDFSD where error propagation is neglected.

Computation of the pairwise error probability and truncation of the sum (8) are treated in [17].

VI. 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 [18]. The coefficients of $r(z)$ and the spectrum $r(e^{j\omega})$ are depicted in figure 2 versus angular frequency ω. The feature of this channel is that it has the lowest minimum distance for the fixed ν. Figure 3 reports the convergence of the union bound on the FEER for the MMSE-DDFSD with $\mu = 4$ at SNR $= 20dB$, where SNR $= r_0/\sigma^2$. In a brute force approach to performance evaluation, one should compute all the error polynomials up to a length such that convergence of the sum (8) is attained. A more sensible approach is to select those error polynomials that dominate the sum, and to compute

$$\text{FEER} \approx \sum_{e(z)\in\mathcal{E}_M} 2^{-w_e} P(e), \quad (12)$$

where $\mathcal{E}_M$ is the subset of $\mathcal{E}$ that contains the M polynomials that dominate the sum. In principle, one should produce a list by ordering in descending order the terms that appear in the sum (8), and then should truncate the list to the first terms. However, at moderate-to-high SNR the coefficient 2^{-w_e} is dominated by $P(e)$, and ISI is dominated by the Gaussian noise [11], yielding the close approximation

$$P(e) \cong Q(\sqrt{\text{SDR}_e}), \quad (13)$$

where

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{u^2}{2}} du,$$

and SDR_e is the *unbiased* Signal-to-Distortion Ratio relevant to the specific $e(z)$ [17]. The 144 error polynomials at lower SDR have been found by the algorithm described in [17], and the contribution $2^{-w_e}P(e)$ of each pair $e(z)$, $-e(z)$ of error polynomials is reported in the figure. The 18 error polynomials at lower SDR are listed in table I. In this specific example, it happens that the first 18 polynomials are the same for the W-DDFSD and for the MMSE-DDFSD. In table I the pairwise error probability obtained by the method proposed in [17] and the pairwise error probability obtained by the Gaussian approximation are also reported. The agreement between the true probability and its approximation is apparent from the table. Figure 4 reports the FEER versus SNR for MLSD and for the MMSE-DDFSD and the W-DDFSD with $\mu = 0$ and $\mu = 4$. In the simulations, the FEER is measured by collecting at least 100 events. In the computation of

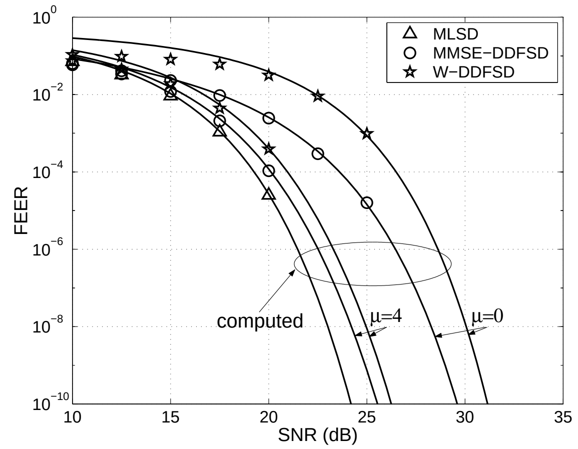

Fig. 4. FEER versus SNR. The solid line is the union bound truncated to the first 18 terms.

the truncated upper bound, the list ordered by SDR is re-ordered by $2^{-w_e}P(e)$, and the first 18 terms are used. From figure 4, one realizes that truncation to the first 18 terms virtually gives the upper bound. Figure 5 reports the BER versus SNR with the same parameters as in figure 4. The solid line is the approximation truncated to the first 18 terms after re-ordering by $w_e 2^{-w_e}P(e)$. The figure shows that the approximation is more accurate for the DDFSD than for the DFE, while for MLSD the first 18 terms virtually give the upper bound. This is an expected result, since the impact of error propagation on the BER diminishes passing from the pure DFE to MLSD. Figure 6 reports the BER of MLSD, MMSE-GDDFSD, and W-GDDFSD, versus SNR. Examining the results reported in Figure 6, one observes that the MMSE-GDDFSD with $\mu = 2$, $M = 2$ outperforms the MMSE-GDDFSD with $\mu = 3$, $M = 1$ (that is the MMSE-DDFSD with 8 states), and that its performance is close to the performance achieved with $\mu = 4$, $M = 2$. This observation suggests that, when severe complexity reduction is

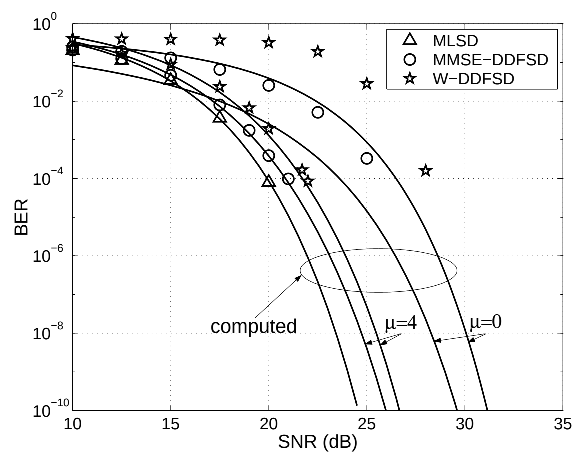

Fig. 5. BER versus SNR. The straight line is the approximation truncated to the first 18 terms.

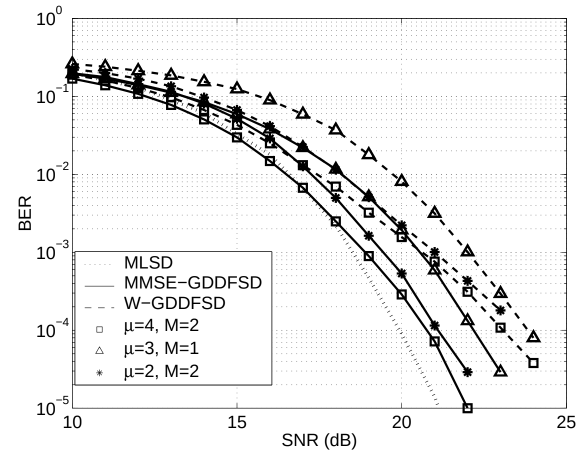

Fig. 6. Bit error rate of MLSD and of the GDDFSD with the noise whitening prefilter and the mean-square prefilter. M is the number of survivors per state, and 2^{μ} is the number of states.

necessary, a well-balanced design of μ and M may offer a good trade-off between performance and complexity.

VII. Summary

The MSWMF is widely known and studied in the theory of DFE [11], but it seems to be less considered in sequence detection. Moving from this observation, we have proved that MLSD with minimum number of states is obtained when the MSWMF is adopted as a front-end. Recently it has been adopted as a prefilter for the DDFSD in [8].

The comparison between the MMSE-DDFSD and the conventional W-DDFSD shows that the MSWMF-DDFSD outperforms its rival when a severe frequency selective channel is considered. This result is intuitive, because the receiver based on the WMF treats spectral nulls as a limiting case. In contrast, the case where the spectrum is null in some interval is not a limiting case for the MSWMF, provided that SNR $\neq \infty$.

The FEER is evaluated by truncating the union bound, where truncation is such that only the terms that dominate the sum (8) are taken into account. Evaluation of the BER is complicated by the error propagation induced by the per-survivor DFE. However, an approximation to the BER can be obtained if error propagation is neglected. As expected, the approximation fairly fits the simulation results only for moderate reduction of complexity.

In order to improve the performance of the MMSE-DDFSD the scheme of the GVA has been proposed, leading to what we call MMSE-GDDFSD. The GDDFSD is a detector that consist in a reduced number of states and multiple survivors per state, where a decision feedback equalizer is attached to each survivor. Our results suggest that, in the design of the GDDFSD, a studied balancement between the number of states and the number

TABLE I
FIRST 18 ERROR POLYNOMIALS ORDERED BY SDR FOR THE MMSE-DDFSD AND FOR THE W-DDFSD WITH $\mu = 4$ AT SNR $= 20dB$. ONLY THE 9 POLYNOMIALS BEGINNING WITH $e_0 = -2$ ARE LISTED.

$P_{MMSE}(e)$	$Q(\sqrt{\text{SDR}_{e,MMSE}})$	$\text{SDR}_{e,MMSE}$ [dB]	$\text{SDR}_{e,W}$ [dB]	Coefficients of $e(z)$
$6.77 \cdot 10^{-4}$	$6.81 \cdot 10^{-4}$	10.11	9.04	-2 2 2 -2 -2 2
$2.30 \cdot 10^{-4}$	$2.31 \cdot 10^{-4}$	10.89	10.00	-2 2 2 -2 -2 2 2 2 -2
$1.49 \cdot 10^{-4}$	$1.50 \cdot 10^{-4}$	11.16	10.35	-2 2
$1.04 \cdot 10^{-4}$	$1.04 \cdot 10^{-4}$	11.38	10.61	-2 2 0 0 0 -2 2
$0.79 \cdot 10^{-4}$	$0.79 \cdot 10^{-4}$	11.54	10.80	-2 2 2 -2 -2 2 2 2 -2 -2 2
$0.72 \cdot 10^{-4}$	$0.73 \cdot 10^{-4}$	11.59	10.86	-2 2 0 0 0 -2 2 0 0 0 -2 2
$0.50 \cdot 10^{-4}$	$0.51 \cdot 10^{-4}$	11.79	11.10	-2 2 0 0 0 -2 2 0 0 0 -2 2 0 0 0 -2 2
$0.35 \cdot 10^{-4}$	$0.35 \cdot 10^{-4}$	11.98	11.32	-2 2 0 0 0 -2 2 0 0 0 -2 2 0 0 0 -2 2 0 0 0 -2 2
$0.34 \cdot 10^{-4}$	$0.34 \cdot 10^{-4}$	12.00	11.34	-2 2 2 -2 -2 2 2 0 2 -2 -2 2 2 2 -2

of survivors may offer a good trade-off between performance and complexity.

References

[1] 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.

[2] S. U. H. Qureshi, “Adaptive equalization,” *Proc. IEEE*, vol. 53, pp. 1349-1387, Sept. 1985.

[3] 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.

[4] D. D. Falconer and J. F. R. Magee, “Adaptive channel memory truncation for maximum likelihood sequence estimation,” *Bell Syst. Tech. J.*, vol. 52, pp. 4681-9562, Nov. 1973.

[5] 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.

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

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

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

[9] L. C. Barbosa, “Maximum likelihood sequence estimators: a geometric view,” *IEEE Trans. Inform. Theory*, vol. IT-35, pp. 419-427, Mar. 1989.

[10] 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.

[11] 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.

[12] P. Monsen, “Feedback equalization for fading dispersive channels,” *IEEE Trans. Inform. Theory*, vol. IT-17, pp. 56-64, Jan. 1971.

[13] J. Salz, “Optimum mean-square decision feedback equalization,” *Bell Syst. Tech. J.*, vol. 52, pp. 1341-1373, Dec. 1973.

[14] T. Hashimoto, “A list-type reduced-constraint generalization of the Viterbi algorithm,” *IEEE Trans. Inform. Theory*, vol. 33, pp. 866-876, Nov. 1987.

[15] A. J. Viterbi and J. K. Omura, “ Principles of Digital Communication and Coding,” *McGraw-Hill*, pp. 278-279, 1979.

[16] W. Sheen and G. D. Stüber, “Error probability for reduced-state sequence estimation,” *IEEE J. Select. Areas Commun.*, vol. 10, pp. 571-578, April 1992.

[17] M. Magarini, A. Spalvieri, and G. Tartara, ”Performance evaluation of the MMSE delayed decision feedback sequence detector”, *Proceedings of GLOBECOM '01*, San Antonio, TX, vol. 2, pp. 1356-1360, Nov. 2001.

[18] 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.