Sensitivity of the Mean-Square DDFSD to a Noisy Estimate of the Noise Variance

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 suboptimal technique for reducing the number of states of the Viterbi detector is the delayed decision feedback sequence detector (DDFSD). In order to achieve good performance a mean-square (MS) prefilter is employed before the DDFSD. This paper is devoted to investigate the sensitivity of the MS-DDFSD to a noisy estimate of the variance of the noise. Analytical performance evaluation is addressed. A truncated version of the union bound is used to approximate the bit error rate. The analysis includes a method for determining the terms that dominate the bound.

I. INTRODUCTION

In digital mobile radio channels, time-varying multipath propagation can cause severe performance degradation. For high-speed data transmission, 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], is often too complex. For practical implementation, a suboptimum that allows a favorable trade off between complexity and performance is the delayed decision feedback sequence detector (DDFSD) [2].

In order to achieve good performance, an appropriate design of the prefilter and of the impulse response used for metric calculation in the DDFSD is essential. In [2] the noise whitening prefilter of [1] was considered before the DDFSD (W-DDFSD). In [3] has been shown that performance, especially for highly frequency selective channels, can be improved by adopting the mean-square prefilter (MS-DDFSD). In [4] the prefilter is obtained as an approximation to the closed form solution of the MMSE-DFE feedforward filter. In this approximation, the true variance of the noise is replaced by a free parameter. The choice of this parameter is up to the designer, whose aim is to approximate an all-pass filter. Actually, the use of a different value of the noise variance could also be a consequence of estimation errors. For example, in TDMA mobile communications a training sequence is inserted in each burst in order to have a promptly adaptation of equalizer's coefficients. An adaptive channel estimator can be constructed using either detected data or the training sequence. By using the estimated channel, one can get an estimate of the variance of the additive noise. Then, the prefilter's coefficients are calculated by substituting the estimated channel in the closed form solution.

In this paper the sensitivity of the performance of the MS-DDFSD to a noisy estimate of the noise variance is considered. For mathematical tractability we consider only the effect of errors in the noise variance estimate, and assume that the impulse response of the channel is perfectly known. The paper is organized as follows. In section II the system model is described and the closed form of the prefilter is given. In section III performance evaluation is carried out. In section IV, the accuracy of the approximation is demonstrated by comparing it to simulation results. In section V conclusions are drawn.

II. SYSTEM MODEL

We consider the model of a binary uncoded data sequence transmitted over a baseband linear channel corrupted by additive white Gaussian noise. The receiver consists of the sampled matched filter, the prefilter, and the DDFSD. The block diagram of the system is depicted 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*

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

where that $d(z)$ causal and minimum phase is taken. In order to take into account errors in the estimate of the noise variance, $\hat{\sigma}^2$ has been used in (1), instead of the true variance σ^2. The signal to noise ratio is SNR$= r_0/\sigma^2$.

In the DDFSD with 2^μ states, $\mu \leq \nu$, 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

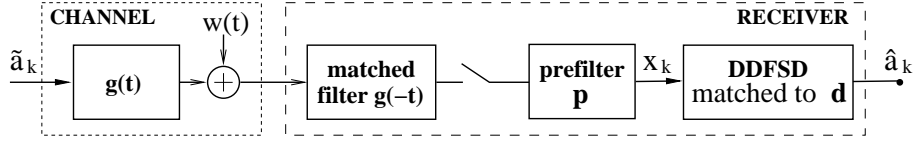$$b_k\,(a_{k-\mu}, \ldots, a_k) = (x_k - \sum_{j=0}^{\mu} d_j a_{k-j}$$

Fig. 1. Channel and receiver block diagram.

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

where x_k is the k-th sample at the output of the prefilter and $\hat{a}_{k-j}(a_{k-\mu}, \ldots, a_{k-1})$ is the estimate of the bit transmitted at time $k-j$, which is present in the survivor that at time $k-1$ merges in the state $(a_{k-\mu}, \ldots, a_{k-1})$.

According to [3], the prefilter is

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

Note that, in practical systems, $\hat{\sigma} > 0$, hence $d(z)$ is always invertible. Therefore the existence of $p(z)$ is guaranteed. It is worth noting that, for $\mu = \nu$ the receiver is the MLSD (see the appendix of [3] for the proof), while for $\hat{\sigma}^2 = \sigma^2$ (2) gives the mean-square prefilter of [3]. In this latter case, for $\mu = 0$ one writes the equations of the MMSE-DFE [1] as

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

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

where the notation $[f(z)]_0 = f_0$ is adopted. The distortion sequence at the output of the prefilter is

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

where $\tilde{a}_k \in \{-1, +1\}$ is the k-th element of the transmitted sequence.

III. PERFORMANCE EVALUATION

Performance evaluation is carried out by adapting the analysis developed in [6, 7] to the case where $\hat{\sigma}^2$ is used in the spectral factorization (1). Let $\tilde{a}(z)$ and $\hat{a}(z)$ be respectively the z-transform of the transmitted and of the detected sequence. The input error event is defined as

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

where $e_0 \neq 0$, $e_{l-1} \neq 0$, and there are no more than $\mu - 1$ consecutive zeros between 0 and $l-1$ (the notation $[f(z)]_i^{i+j} = \sum_{k=i}^{i+j} f_k z^{-k}$ is adopted). Moreover, we assume

$$[\hat{a}(z) - \tilde{a}(z)]_{-\nu}^{-1} = 0, \qquad (4)$$

that is, error propagation is neglected. The bit error rate (BER) is approximated as [6]

$$BER \approx \sum_{e(z) \in \mathcal{E}_M} w_e 2^{-w_e} P(\tilde{a}(z) \mapsto \tilde{a}(z) + e(z)), \qquad (5)$$

where $w_e = [e(z)e(z^{-1})]_0/4$ is the Hamming weight of $e(z)$, $\mathcal{E}_M$ is the set of the M input error polynomials that dominate the union bound, and $P(\tilde{a}(z) \mapsto \tilde{a}(z) + e(z))$ is the pairwise error probability in the binary test between $\tilde{a}(z)$ and $\tilde{a}(z) + e(z)$. Note that the binary test may take place only if the two events (4) and

$$[\hat{a}(z) - \tilde{a}(z)]_l^{l+\mu-1} = 0 \qquad (6)$$

occur. Condition (6) corresponds to the first merging in the trellis of the DDFSD between the competing sequences. In what follows the computation of $P(\tilde{a}(z) \mapsto \tilde{a}(z) + e(z))$ and the search of the M error events that form the set $\mathcal{E}_M$ will be discussed.

As far as the computation of the pairwise error probability is concerned, it should be noted that both ISI and noise affect the distortion sequence [5]. From a 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

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

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 in the decision space is

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

which also represents the energy of the output error. Let

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

be the z-transform of the distortion sequence (3). The error occurs when the projection of the distortion along the output error, defined as

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

exceeds half the Euclidean distance between the competing sequences. The pairwise error probability $P(\tilde{a}(z) \mapsto \tilde{a}(z) + e(z))$ is

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

where $f_{\phi_e}(x)$ is the probability density function of ϕ_e. In order to compute $f_{\phi_e}(x)$, we will take into account that the distortion is the sum of ISI and noise. The projection of the noise along the output error is

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

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

$$m_{\zeta_e} = 0,$$

and variance

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

The projection of the ISI along the output error is

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

where $c(z)$ is the polynomial of the coefficients of the ISI. The probability density function $f_{\psi_e}(x)$ can be computed from the coefficients of the ISI. Since ISI and noise are independent random variables, the probability density function of ϕ_e is

$$f_{\phi_e}(x) = f_{\psi_e}(x) \otimes f_{\zeta_e}(x),$$

where $\otimes$ denotes the convolution. In the section devoted to the experimental results, we adopt the method proposed in [8] to compute $f_{\psi_e}(x)$.

In order to evaluate the performance, we should determine the M error events to be considered in (5). To an efficient selection, a sensible figure of merit is the signal to distortion ratio relevant to $e(z)$ [7], which is defined as

$$\text{SDR}_e = \frac{(\delta_e - 2m_{\phi_e})^2}{4\sigma_{\phi_e}^2}. \tag{7}$$

In (7), the mean and the variance of the distortion are used. Since the distortion is the sum of ISI and noise, we have to compute the mean and the variance of the ISI. Following the procedure developed in appendix of [7], one finds

$$m_{\psi_e} = \frac{2\hat{\sigma}^2 w_e}{\delta_e},$$

and

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

Taking into account that the mean value of the Gaussian noise is zero, for m_{ϕ_e} one gets

$$m_{\phi_e} = m_{\psi_e} = \frac{2\hat{\sigma}^2 w_e}{\delta_e}.$$

Using $p(z)(r(z) + \hat{\sigma}^2)p(z^{-1}) = 1$, for $\sigma_{\phi_e}^2$ one gets

$$\begin{aligned}
\sigma_{\phi_e}^2 &= \sigma_{\zeta_e}^2 + \sigma_{\psi_e}^2 \\
&= \hat{\sigma}^2 - 4\frac{\hat{\sigma}^4}{\delta_e^2}w_e + \frac{\sigma^2 - \hat{\sigma}^2}{\delta_e^2}[e_o(z^{-1})p(z)r(z)p(z^{-1})e_o(z)]_0.
\end{aligned}$$

At intermediate-to-high SNR, the sum (5) will be dominated by the terms corresponding to the error events at lower SDR_e. Hence, one can produce a large list of error polynomials ordered by increasing SDR_e, and then compute $2^{-w_e} P(\tilde{a}(z) \mapsto \tilde{a}(z) + e(z))$ for each element of the list to produce a re-ordered list, which can be used in (5). The problem in calculating (7) is that in general the denominator depends on the particular error event. Hence, algorithms which look for the error events at minimum distance cannot be used. By applying the method described in [7], we can look for the list of error polynomials at minimum distance only for $\hat{\sigma}^2 = 0$ and $\hat{\sigma}^2 = \sigma^2$. For other values of $\hat{\sigma}^2$ a new list of error polynomials is created by merging the two above lists. By using (7), the error polynomials of the new list are ordered by increasing SDR_e. We adopt this procedure in order to determine the first M error polynomials that contribute to the sum (5) for a generic $\hat{\sigma}^2$. This procedure has been derived after observing that the M dominant error polynomials for $\hat{\sigma}^2 = 0$ ($\hat{\sigma}^2 = \sigma^2$) could be extracted from the list of the error polynomials found for $\hat{\sigma}^2 = \sigma^2$ ($\hat{\sigma}^2 = 0$). Of course, the list from which error polynomials are extracted must be larger than M. We consider a list of $3M$ error polynomials both for the W-DDFSD and the MS-DDFSD.

IV. EXPERIMENTAL RESULTS

Computer simulations were carried out to evaluate the accuracy of the approximation (5). In the simulations we focused on the channel with $\nu = 6$ studied in [9], whose impulse response at the output of the matched filter is $r(z) = 0.031 + 0.1112z^{-1} + 0.2674z^{-2} + 0.4881z^{-3} + 0.7304z^{-4} + 0.9185z^{-5} + 0.9978z^{-6} + 0.9185z^{-7} + 0.7304z^{-8} + 0.4881z^{-9} + 0.2674z^{-10} + 0.1112z^{-11} + 0.031z^{-12}$. This channel is characterized by three deep notches in its spectrum. Fig. 2 shows the BER versus $\zeta = \hat{\sigma}^2/\sigma^2$, for $\mu = 4$ and $\sigma^2 = 10^{-2}$ (SNR=20 dB). In the simulations, the BER is measured by a random sequence of 10^6 data. From the figure one observes that the measured BER is a smooth function of ζ.

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

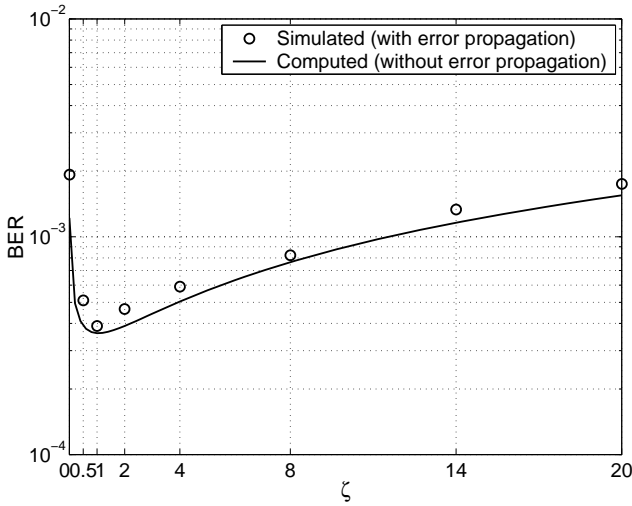| $\text{SDR}_{e,MS}[dB]$	$\text{SDR}_{e,W}$ [dB]	Coefficients of $e(z)$
10.11	9.04	-2 2 2 -2 -2 2
10.89	10.00	-2 2 2 -2 -2 2 2 -2
11.16	10.35	-2 2
11.38	10.61	-2 2 0 0 0 -2 2
11.54	10.80	-2 2 2 -2 -2 2 2 -2 -2 2
11.59	10.86	-2 2 0 0 0 -2 2 0 0 0 -2 2
11.79	11.10	-2 2 0 0 0 -2 2 0 0 0 -2 2 0 0 0 -2 2
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
12.00	11.34	-2 2 2 -2 -2 2 0 2 -2 -2 2 2 -2

Fig. 2. Simulated and estimated BER versus ζ, for $\sigma^2 = 10^{-2}$ and $\mu = 4$.

Fig. 3. Simulated and estimated BER versus ζ, for $\sigma^2 = 3.2 \cdot 10^{-3}$ and $\mu = 0$.

As expected the minimum occurs for $\zeta = 1$. In the figure is also reported the estimated bit error probability. The figure shows that the approximation is fairly accurate. The slightly worse performance of the simulations is due to the error propagation that takes place in the DDFSD. To fit the simulations results, we find that 18 error polynomials contribute to the sum (5). Since $e(z)$ and $-e(z)$ produce the same SDR_e, in table I are listed the coefficients of the 9 error polynomials at lower SDR_e that begin with $e_0 = -2$, for $\hat{\sigma}^2 = 0$ and $\hat{\sigma}^2 = \sigma^2$. As we can observe from the table, in this specific example, it happens that the first 18 error polynomials are the same for the W-DDFSD and for the MS-DDFSD. Fig. 3 shows the BER of the DFE ($\mu = 0$) versus ζ, for $\sigma^2 = 3.2 \cdot 10^{-3}$ (SNR$= 25$ dB). In the figure the simulation results with and without error propagation are reported. In the simulations, the BER is measured by a random sequence of 10^7 data. From the figure we observe that the experi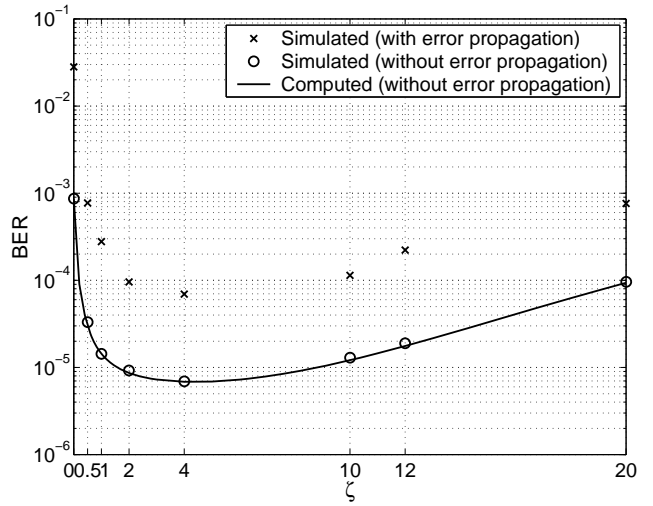mental results, in absence of error propagation, fit very closely the approximation. When error propagation is considered, the approximation (5) turns out to be poor. An analytical estimate of the performance that considers error propagation can be found in [10] and references therein. We also note that the minimum of the BER does not occur for $\zeta = 1$ but for $\zeta = 4$. This behavior is not common to all the channels that we have considered. For less frequency selective channels the minimum occurs for $\zeta = 1$. Hence, for high complexity reduction and for severely frequency selective channels, the use of a larger value for the noise variance in (1) allows to improve performance. For the DFE the idea of using $\zeta > 1$ in (1) has been proposed in [11] in the context of the magnetic recording channel. However, in [11] the prefilter is different from that proposed here. Finally, by comparing Fig. 2 and Fig. 3 we observe that, near the minimum, the curve relative to the pure DFE is more flat than the curve obtained with $\mu = 4$. Hence, we conclude that, for large complexity reduc-

tion, the performance is less sensitive to errors in the estimate of the noise variance.

V. CONCLUSIONS

For severely frequency selective channels, the MS-DDFSD provides a favorable trade off between performance and complexity. In this receiver, the mean-square prefilter that precedes the DDFSD and the impulse response to which the DDFSD is matched are calculated from a spectral factorization, where the autocorrelation of the impulse response of the channel and the noise variance are assumed to be known. Since both the channel and the noise variance should be estimated from the received signal, the performance of the receiver will be influenced by the quality of these estimates. In this paper an approximation to the bit error probability of the MS-DDFSD for a noisy estimate of the noise variance has been derived. From the experimental results the conclusions that we can draw are the following. For moderate complexity reduction, the minimum of the BER is obtained when $\hat{\sigma}^2 = \sigma^2$. For large complexity reduction, the minimum of the BER is obtained when $\hat{\sigma}^2 \geq \sigma^2$.

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

[3] M. Magarini, A. Spalvieri, and G. Tartara, "Comparison between two methods for delayed decision feedback sequence estimation," in *Proc. IEEE Personal, Indoor and Mobile Radio Communications*, London, U.K., 2000, pp. 1300-1304.

[4] W. Gerstacker and J. Huber, "Improved equalization for GSM mobile communications," in *Proc. International Conference on Telecommunications*, Istanbul, Turkey, 1996, pp. 128-131.

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

[7] M. Magarini, A. Spalvieri, and G. Tartara, "Performance evaluation of the mean-squared prefiltered delayed decsion feedback sequence detector ," in *Proc. IEEE Personal, Indoor and Mobile Radio Communications*, 2001, San Diego, CA, in press.

[8] 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, Mar. 1974.

[9] 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.

[10] T. J. Willink, P. H. Wittke and L. L. Campbell, "Evaluation of the effects of intersymbol interference in decision-feedback equalizers," *IEEE Trans. Commun.*, vol. 48, pp. 629 -636, Apr. 2000.

[11] P. S. Bednarz and J. M. Cioffi, "Decision feedback equalization for channels with error correcting capabilities," in *Proc. IEEE International Conference on Communications*, Montreal, Canada, 1997, pp. 1607-1612.