

# Decision-Feedback Equalization Using Multiple-Hyperplane Partitioning for Detecting ISI-Corrupted $M$ -ary PAM Signals

S. Chen, L. Hanzo, and B. Mulgrew

**Abstract**—A decision-feedback equalizer scheme is derived based on multiple-hyperplane partitioning of signal space for detecting  $M$ -ary pulse amplitude modulation symbols transmitted through a noisy intersymbol interference channel. The proposed scheme is based on the fact that the optimal Bayesian decision boundary separating two neighboring signal classes is asymptotically piecewise linear and consists of several hyperplanes, when the signal-to-noise ratio tends to infinity. An algorithm is developed to determine these hyperplanes, which are then used to partition the observation signal space. The resulting detector can closely approximate the optimal Bayesian detector, at an advantage of considerably reduced detector complexity.

**Index Terms**—Asymptotic decision boundary, Bayesian decision-feedback equalizer, multiple-hyperplane detector, signal space partitioning.

## I. INTRODUCTION

THE decision-feedback equalizer (DFE) is a powerful technique for combating distortion and interference in communication links and high-density data storage systems. For the class of DFE structures that employs symbol-by-symbol detection, the optimal solution is well understood to be the maximum *a posteriori* probability or Bayesian detector [1]–[5]. The complexity of the optimal Bayesian DFE is determined by the factor of  $M^{n_a}$ , where  $M$  being the size of the symbol constellation and  $n_a$  the channel-impulse response (CIR) length. As the complexity of the optimal detector increases exponentially with the size of symbol set  $M$ , the conventional or linear-combiner DFE [6]–[10] is often used in practice to provide a tradeoff between performance and detector complexity.

For the binary case, the difference in performance between the conventional and optimal Bayesian DFEs can be explained geometrically: a linear-combiner DFE can only partition the observation space with a hyperplane while the optimal Bayesian detector can do so with a hypersurface [4]. Asymptotically, the optimal Bayesian hypersurface becomes piecewise linear and is made up of a set of hyperplanes [11]. In practice, at large signal-to-noise ratio (SNR) (usually 10–20 dB), the Bayesian decision hypersurface can closely be approximated by a multiple-hyperplane form. Thus, a detector that partitions the observation space with multiple hyperplanes can approximate the

optimal Bayesian detector more closely than the conventional DFE.

Signal space partitioning techniques for binary intersymbol interference (ISI) channels have been developed. Notably, Moon and Jeon [12] developed a detector using the Voronoi diagram. Kim and Moon [13], [14] improved on the design and developed a new signal space partitioning technique. Their technique determines a set of hyperplanes that separate clusters of noiseless channel states. The convex regions associated with individual states are constructed by intersecting hyperplanes. The overall decision region is then formed from these convex regions. The decision complexity and performance of the detector is controlled during design by a specified minimum separating distance. The main drawback of their design is that a combinatorial search and optimization process is carried out to find those hyperplanes, which requires extensive computational efforts. In our previous work [15], we have shown that this design complexity can be avoided and we have proposed a much simpler alternative design to explicitly realize the asymptotic Bayesian decision boundary for binary ISI channels.

In this letter, we extend this multiple-hyperplane detector design to  $M$ -ary pulse amplitude modulation (PAM) channels. Based on a geometric translation property for the  $M$  sets of noiseless channel states, the asymptotic Bayesian boundary for separating any two neighboring signal classes can be deduced, and this allows us to extend the algorithm for the binary case [15] to the general  $M$ -ary PAM case. Similar to the binary case, the design of our multiple-hyperplane detector for  $M$ -ary PAM channels is straightforward, and guarantees to realize the asymptotic Bayesian DFE detector. Furthermore, the reduction in detector complexity with signal space partitioning approach is more significant for  $M > 2$ . A simulation example is included to compare the performance and detection complexity of the resulting multiple-hyperplane detector with those of the optimal Bayesian DFE.

## II. PROBLEM FORMULATION

We will assume that the real-valued channel generates the received signal samples of

$$y(k) = \sum_{i=0}^{n_a-1} a_i s(k-i) + e(k) \quad (1)$$

where  $a_i$  are the CIR taps, the Gaussian white noise  $\{e(k)\}$  has zero mean and variance  $\sigma_e^2$ , and the  $M$ -ary PAM symbol  $s(k)$  takes the value from the symbol set  $\mathcal{S} \triangleq \{s_i = 2i - M - 1, 1 \leq i \leq M\}$ . The SNR of the system is defined as  $\text{SNR} \triangleq (\sum_{i=0}^{n_a-1} a_i^2) \sigma_s^2 / \sigma_e^2$ , where  $\sigma_s^2$  is the symbol variance. The generic DFE studied in this letter uses the informa-

Paper approved by S. L. Miller, the Editor for Spread Spectrum of the IEEE Communications Society. Manuscript received June 16, 2000; revised August 10, 2000.

S. Chen and L. Hanzo are with Department of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: sqc@ecs.soton.ac.uk).

B. Mulgrew is with Department of Electronics and Electrical Engineering, University of Edinburgh, Edinburgh EH9 3JL, U.K.

Publisher Item Identifier S 0090-6778(01)04075-2.

tion present in the noisy observation vector  $\mathbf{y}(k) = [y(k) \ y(k-1) \ \dots \ y(k-m+1)]^T$  and the past detected symbol vector  $\hat{\mathbf{s}}_b(k) = [\hat{s}(k-d-1) \ \dots \ \hat{s}(k-d-n)]^T$  to produce an estimate  $\hat{s}(k-d)$  of  $s(k-d)$ , where  $d, m$ , and  $n$  are the decision delay, the feedforward and feedback orders, respectively. We will choose  $d = n_a - 1, m = n_a$ , and  $n = n_a - 1$ , as this choice is sufficient to guarantee a desired linear separability for different signal classes [8], [9].

The observation vector  $\mathbf{y}(k)$  can be expressed as [5], [8]

$$\mathbf{y}(k) = F_1 \mathbf{s}_f(k) + F_2 \mathbf{s}_b(k) + \mathbf{e}(k) \quad (2)$$

where  $\mathbf{s}_f(k) = [s(k) \ \dots \ s(k-d)]^T, \mathbf{s}_b(k) = [s(k-d-1) \ \dots \ s(k-d-n)]^T, \mathbf{e}(k) = [e(k) \ \dots \ e(k-m+1)]^T$ , and

$$F_1 = \begin{bmatrix} a_0 & a_1 & \dots & a_{n_a-1} \\ 0 & a_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_1 \\ 0 & \dots & 0 & a_0 \end{bmatrix} \quad (3)$$

$$F_2 = \begin{bmatrix} 0 & 0 & \dots & 0 \\ a_{n_a-1} & 0 & \ddots & \vdots \\ a_{n_a-2} & a_{n_a-1} & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ a_1 & \dots & a_{n_a-2} & a_{n_a-1} \end{bmatrix} \quad (4)$$

are the  $m \times (d+1)$  and  $m \times n$  CIR matrices, respectively. Assuming correct past decisions, we have  $\hat{\mathbf{s}}_b(k) = \mathbf{s}_b(k)$  and

$$\mathbf{y}(k) = F_1 \mathbf{s}_f(k) + F_2 \hat{\mathbf{s}}_b(k) + \mathbf{e}(k). \quad (5)$$

Thus, the decision feedback translates the original space  $\mathbf{y}(k)$  into a new space  $\mathbf{r}(k)$

$$\mathbf{r}(k) \triangleq \mathbf{y}(k) - F_2 \hat{\mathbf{s}}_b(k). \quad (6)$$

Let the  $N_f = M^{d+1}$  possible sequences of  $\mathbf{s}_f(k)$  be  $\mathbf{s}_{fj}, 1 \leq j \leq N_f$ . The set of the noiseless channel states in the translated signal space, namely

$$R \triangleq \{\mathbf{r}_j = F_1 \mathbf{s}_{fj}, 1 \leq j \leq N_f\} \quad (7)$$

can be partitioned into  $M$  subsets conditioned on  $s(k-d)$

$$R^{(i)} \triangleq \{\mathbf{r}_j \in R : s(k-d) = s_i\}, \quad 1 \leq i \leq M. \quad (8)$$

The optimal Bayesian DFE [5] can now be summarized. The  $M$  decision variables are given by

$$\rho_i(\mathbf{r}(k)) = \sum_{\mathbf{r}_j \in R^{(i)}} \exp\left(-\frac{\|\mathbf{r}(k) - \mathbf{r}_j\|^2}{2\sigma_e^2}\right), \quad 1 \leq i \leq M \quad (9)$$

and the minimum-error-probability decision is defined by

$$\hat{s}(k-d) = s_{i^*}, \quad \text{with } i^* = \arg \max_{1 \leq i \leq M} \{\rho_i(\mathbf{r}(k))\}. \quad (10)$$

The complexity of this optimal detector is given in Table I.

TABLE I  
COMPARISON OF DECISION COMPLEXITY FOR THE FULL BAYESIAN AND MULTIPLE-HYPERPLANE DETECTORS

	Full Bayesian detector	Multiple-hyperplane detector
Additions	$2n_a M^{n_a} - M$	$(n_a + M - 2)L$
Multiplications	$(n_a + 1)M^{n_a}$	$n_a L$
exp( $\cdot$ )	$M^{n_a}$	—

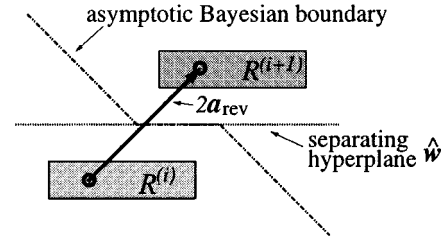


Fig. 1. Illustration of shift property for subsets of channel states.

### III. MULTIPLE-HYPERPLANE DETECTOR

We first establish a geometric translation property for any two neighboring subsets of channel states.

*Lemma 1:* For  $1 \leq i \leq M-1$ , the subset  $R^{(i+1)}$  is a translation of  $R^{(i)}$  by the amount  $2\mathbf{a}_{\text{rev}}$

$$R^{(i+1)} = R^{(i)} + 2\mathbf{a}_{\text{rev}} \quad (11)$$

where  $\mathbf{a}_{\text{rev}} = [a_{n_a-1} \ \dots \ a_1 \ a_0]^T$ . Furthermore,  $R^{(i)}$  and  $R^{(i+1)}$  are linearly separable.

*Proof:* From the definitions of  $R^{(i)}$  and  $F_1$ , for any  $\mathbf{r}_l \in R^{(i)}$ , there exists a  $\mathbf{r}_j \in R^{(i+1)}$  such that  $\mathbf{r}_j = \mathbf{r}_l + (s_{i+1} - s_i)\mathbf{a}_{\text{rev}} = \mathbf{r}_l + 2\mathbf{a}_{\text{rev}}$ , which implies (11). To prove the linear separability, consider the hyperplane

$$H(\mathbf{r} + \mathbf{c}_i) \triangleq \hat{\mathbf{w}}^T \left( \mathbf{r} + 2 \left( \frac{M}{2} - i \right) \mathbf{a}_{\text{rev}} \right) = 0 \quad (12)$$

with  $\hat{\mathbf{w}} = [0 \ 0 \ \dots \ 0 \ (1/a_0)]^T$ . For any  $\mathbf{r}_l \in R^{(i)}$  and any  $\mathbf{r}_j \in R^{(i+1)}$ , we have  $H(\mathbf{r}_l + \mathbf{c}_i) = -1 < 0$  and  $H(\mathbf{r}_j + \mathbf{c}_i) = 1 > 0$ . Fig. 1 illustrates this lemma graphically.  $\square$

#### A. Asymptotic Optimal Boundary for Two Neighboring Classes

Although it is always possible to construct a hyperplane to correctly separate  $R^{(i)}$  from  $R^{(i+1)}$ , the optimal decision boundary  $\mathcal{D}_i$  that separates  $R^{(i)}$  from  $R^{(i+1)}$  cannot generally be approximated by a single hyperplane. Without the loss of generality, consider  $i = (M/2)$ , the optimal decision boundary  $\mathcal{D}_{M/2}$  for separating  $R^{(M/2)}$  and  $R^{((M/2)+1)}$ . Because of Lemma 1, when  $\text{SNR} \rightarrow \infty$  (or  $\sigma_e^2 \rightarrow 0$ ), the influence from all the other  $R^{(i)}$  for  $i \neq M/2$  and  $i \neq (M/2) + 1$  vanishes much more quickly, and it effectively becomes a two-class problem. We have the following definition, similar to the one given in [15].

*Definition 1:* A pair of opposite-class channel states ( $\mathbf{r}^{(+)} \in R^{((M/2)+1)}, \mathbf{r}^{(-)} \in R^{(M/2)}$ ) is said to be *dominant* if  $\forall \mathbf{r}_j \in R^{(M/2)} \cup R^{((M/2)+1)}, \mathbf{r}_j \neq \mathbf{r}^{(+)}$  and  $\mathbf{r}_j \neq \mathbf{r}^{(-)}$

$$\|\mathbf{r}_j - \mathbf{r}_0\|^2 > \|\mathbf{r}^{(+)} - \mathbf{r}_0\|^2 \quad (13)$$

where  $\cup$  denotes the union operator and

$$\mathbf{r}_0 = \frac{\mathbf{r}^{(+)} + \mathbf{r}^{(-)}}{2}. \quad (14)$$

The following properties of  $\mathcal{D}_{M/2}$  are useful in the derivation of a multiple-hyperplane detector (see [11]). A necessary condition for a point  $\mathbf{r}_B \in \mathcal{D}_{M/2}$  is

$$\mathbf{r}_B = \frac{\mathbf{r}^{(+)} + \mathbf{r}^{(-)}}{2} + \left[ \frac{\mathbf{r}^{(+)} - \mathbf{r}^{(-)}}{2} \right]^\perp \quad (15)$$

where  $\mathbf{x}^\perp$  denotes an arbitrary vector in the subspace orthogonal to  $\mathbf{x}$ ,  $\mathbf{r}^{(+)}$  and  $\mathbf{r}^{(-)}$  are a pair of dominant states, and the sufficient conditions for  $\mathbf{r}_B \in \mathcal{D}_{M/2}$  are

$$\left\| \mathbf{r}_B - \mathbf{r}^{(+)} \right\|^2 < \left\| \mathbf{r}_B - \mathbf{r}_l \right\|^2 \quad \forall \mathbf{r}_l \in R^{(\frac{M}{2}+1)}, \mathbf{r}_l \neq \mathbf{r}^{(+)} \quad (16)$$

$$\left\| \mathbf{r}_B - \mathbf{r}^{(-)} \right\|^2 < \left\| \mathbf{r}_B - \mathbf{r}_j \right\|^2 \quad \forall \mathbf{r}_j \in R^{(\frac{M}{2})}, \mathbf{r}_j \neq \mathbf{r}^{(-)} \quad (17)$$

$$\left\| \mathbf{r}_B - \mathbf{r}^{(+)} \right\|^2 = \left\| \mathbf{r}_B - \mathbf{r}^{(-)} \right\|^2. \quad (18)$$

The following lemma describing  $\mathcal{D}_{M/2}$  in the asymptotic case of  $\sigma_e^2 \rightarrow 0$  is a direct consequence of the necessary and sufficient conditions (15)–(18).

*Lemma 2:* Asymptotically, the optimal decision boundary  $\mathcal{D}_{M/2}$  separating  $R^{(M/2)}$  and  $R^{((M/2)+1)}$  is piecewise linear and made up of a set of  $L$  hyperplanes. Each of these hyperplanes is defined by a pair of dominant states, the hyperplane is orthogonal to the line connecting the pair of dominant states and passes through the midpoint of the line.

### B. Multiple-Hyperplane Detector for Two Neighboring Classes

According to Lemma 2, a multiple-hyperplane detector can be constructed to partition the signal space into the two regions of  $\hat{s}(k-d) \leq -1$  and  $\hat{s}(k-d) \geq 1$ , respectively. The detector will consist of  $L$  linear discriminant functions and a many-to-one Boolean mapper, similar to the binary case given in [15]. For completeness, the following design procedure for this multiple-hyperplane detector is produced here with the necessary modifications.

- Step 1) Select all the  $L$  pairs of dominant channel states from the two subsets  $R^{(M/2)}$  and  $R^{((M/2)+1)}$ . For each pair, compute a hyperplane that separates these two opposite-class states.
- Step 2) A Boolean logic function is obtained to make a decision based on the location of the observation vector  $\mathbf{r}(k)$  relative to each hyperplane. This is achieved by first defining a convex region associated with each state in a given class, e.g., the class  $R^{((M/2)+1)}$ , and then forming a union of these regions.

From (15)–(18), it is easily seen that pairs of dominant states that define the asymptotic boundary can be selected using the following algorithm.

```

L = 0;
FOR  $\mathbf{r}_q^{(+)} \in R^{((M/2)+1)}$ 
  FOR  $\mathbf{r}_j^{(-)} \in R^{(M/2)}$ 
     $\mathbf{x} = (\mathbf{r}_q^{(+)} + \mathbf{r}_j^{(-)})/2; \eta = \|\mathbf{r}_q^{(+)} - \mathbf{x}\|^2;$ 
    IF  $(\|\mathbf{r}_l^{(+)} - \mathbf{x}\|^2 > \eta, \forall \mathbf{r}_l^{(+)} \in R^{((M/2)+1)}, l \neq q)$  AND
       $(\|\mathbf{r}_l^{(-)} - \mathbf{x}\|^2 > \eta, \forall \mathbf{r}_l^{(-)} \in R^{(M/2)}, l \neq j)$ 
      L += 1;
       $R_{\text{Asym}} \leftarrow (\mathbf{r}_L^{(+)}, \mathbf{r}_L^{(-)}) \hat{=} (\mathbf{r}_q^{(+)}, \mathbf{r}_j^{(-)});$ 
    END IF
  NEXT  $\mathbf{r}_j^{(-)}$ 
NEXT  $\mathbf{r}_q^{(+)}$ 

```

Each pair  $(\mathbf{r}_l^{(+)}, \mathbf{r}_l^{(-)}) \in R_{\text{Asym}}$  determines a hyperplane

$$H_l(\mathbf{r}) = \mathbf{w}_l^T \mathbf{r} + b_l = 0 \quad (19)$$

that is a part of the asymptotic optimal decision boundary. The weight vector  $\mathbf{w}_l$  and bias  $b_l$  of the hyperplane can be computed straightforwardly as

$$\mathbf{w}_l = \frac{2(\mathbf{r}_l^{(+)} - \mathbf{r}_l^{(-)})}{\left\| \mathbf{r}_l^{(+)} - \mathbf{r}_l^{(-)} \right\|^2} \quad (20)$$

and

$$b_l = -\frac{(\mathbf{r}_l^{(+)} - \mathbf{r}_l^{(-)})^T (\mathbf{r}_l^{(+)} + \mathbf{r}_l^{(-)})}{\left\| \mathbf{r}_l^{(+)} - \mathbf{r}_l^{(-)} \right\|^2}. \quad (21)$$

The hyperplane defined by (20) and (21) is a *canonical* hyperplane with  $(\mathbf{r}_l^{(+)}, \mathbf{r}_l^{(-)})$  as its two support vectors [16], and has the property that  $H_l(\mathbf{r}_l^{(+)}) = 1$  and  $H_l(\mathbf{r}_l^{(-)}) = -1$ . The following definition is useful in the optimal multiple-hyperplane partitioning.

*Definition 2:* A state  $\mathbf{r}_j \in R^{(M/2)} \cup R^{((M/2)+1)}$  is said to be *sufficiently separable* by the hyperplane  $H_l$  if  $H_l$  can separate  $\mathbf{r}_j$  correctly with a “canonical distance”  $|\mathbf{w}_l^T \mathbf{r}_j + b_l| \geq 1$ .

Notice that  $\mathbf{r}_j \in R^{((M/2)+1)}$  is sufficiently separable by  $H_l$  if and only if  $\mathbf{w}_l^T \mathbf{r}_j + b_l \geq 1$ . Similarly,  $\mathbf{r}_j \in R^{(M/2)}$  is sufficiently separable by  $H_l$  if and only if  $\mathbf{w}_l^T \mathbf{r}_j + b_l \leq -1$ . Number the states in  $R^{(M/2)}$  as  $\mathbf{r}_1^{(-)}$  to  $\mathbf{r}_{N_s}^{(-)}$  and those in  $R^{((M/2)+1)}$  as  $\mathbf{r}_1^{(+)}$  to  $\mathbf{r}_{N_s}^{(+)}$ , where  $N_s = N_f/M$ . All the states in  $R^{(M/2)} \cup R^{((M/2)+1)}$  are tested to see if they can be separated sufficiently by  $H_l, 1 \leq l \leq L$ . This generates the following “separability” matrix:

	$\mathbf{r}_1^{(-)}$	$\mathbf{r}_2^{(-)}$	$\cdots$	$\mathbf{r}_{N_s}^{(-)}$	$\mathbf{r}_1^{(+)}$	$\cdots$	$\mathbf{r}_{N_s}^{(+)}$
$H_1$	$h_{1,1}$	$h_{1,2}$	$\cdots$	$h_{1,N_s}$	$h_{1,N_s+1}$	$\cdots$	$h_{1,2N_s}$
$H_2$	$h_{2,1}$	$h_{2,2}$	$\cdots$	$h_{2,N_s}$	$h_{2,N_s+1}$	$\cdots$	$h_{2,2N_s}$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
$H_L$	$h_{L,1}$	$h_{L,2}$	$\cdots$	$h_{L,N_s}$	$h_{L,N_s+1}$	$\cdots$	$h_{L,2N_s}$

where  $h_{l,j} \in \{0, 1\}$ . The rule in generating this matrix is as follows: if a state can *sufficiently* be separated by  $H_l$ , the corresponding binary index  $h_{l,j} = 1$ ; otherwise  $h_{l,j} = 0$ .

Notice that in every row there are at least two nonzero elements (associated with a dominant pair), and in every column there is at least one nonzero element.



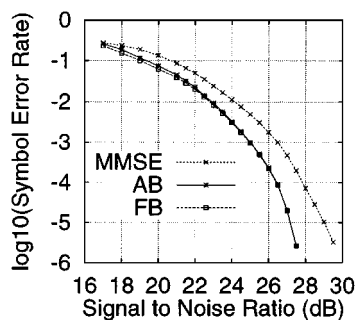


Fig. 2. Performance comparison of the classical MMSE DFE (MMSE), the multiple-hyperplane detector (AB), and the full Bayesian DFE (FB) with detected symbols being fed back.

logic function can be obtained. For instance, the states  $\mathbf{r}_1^{(+,2)}$  to  $\mathbf{r}_6^{(+,2)}$  in  $R^{(3)}$  require the two hyperplanes  $H_1$  and  $H_2$  for separation from  $R^{(2)}$ ;  $\mathbf{r}_7^{(+,2)}$  and  $\mathbf{r}_8^{(+,2)}$  are separated from  $R^{(2)}$  by  $H_3$  and  $H_4$ ; finally,  $H_5$  separates  $\mathbf{r}_9^{(+,2)}$  to  $\mathbf{r}_{16}^{(+,2)}$  from the opposite-class  $R^{(2)}$ . Thus, given the five values of linear discriminant functions  $\bar{H}_l(k) + \bar{b}_{l,i}$ ,  $1 \leq l \leq 5$ , the detector requires two logic AND and two logic OR operations to complete the Boolean logic value  $\alpha_i(k)$ .

The symbol-error rate (SER) performance of this multiple-hyperplane detector is compared with those of the full Bayesian and conventional minimum mean-square-error (MMSE) DFEs in Fig. 2, where it can be seen that there is hardly any SER performance difference between the multiple-hyperplane and full Bayesian detectors. For this example, the full Bayesian DFE requires 380 additions, 256 multiplications, and 64  $\exp(\cdot)$  function evaluations to detect a symbol. The multiple-hyperplane detector, however, needs only 25 additions and 15 multiplications to make a decision, which is less than 6% of the complexity required by the full Bayesian DFE. The linear MMSE DFE requires only five additions and three multiplications to make a decision, as it contains a single hyperplane, but its performance is much inferior.

## V. CONCLUSION

We have extended a signal space partitioning technique, originally developed for binary channels, to  $M$ -ary PAM channels. A design scheme is presented to automatically construct a multiple-hyperplane partitioning that is asymptotically optimal. The

resulting detector consists of a set of linear discriminant functions and associated Boolean logic values, and it has much lower decision complexity compared with the optimal Bayesian detector. Although this multiple-hyperplane detector achieves the optimal Bayesian performance only at the asymptotic case of infinite SNR, in practice, it can close approximate the optimal performance under finite SNR conditions.

## REFERENCES

- [1] K. Abend and B. D. Fritchman, "Statistical detection for communication channels with intersymbol interference," *Proc. IEEE*, vol. 58, pp. 779–785, May 1970.
- [2] J. F. Hayes, T. M. Cover, and J. B. Riera, "Optimal sequence detection and optimal symbol-by-symbol detection: Similar algorithms," *IEEE Trans. Commun.*, vol. COM-30, pp. 152–157, Jan. 1982.
- [3] D. Williamson, R. A. Kennedy, and G. W. Pulford, "Block decision feedback equalization," *IEEE Trans. Commun.*, vol. 40, pp. 255–264, Feb. 1992.
- [4] S. Chen, B. Mulgrew, and S. McLaughlin, "Adaptive Bayesian equaliser with decision feedback," *IEEE Trans. Signal Processing*, vol. 41, pp. 2918–2927, Sept. 1993.
- [5] S. Chen, S. McLaughlin, B. Mulgrew, and P. M. Grant, "Bayesian decision feedback equaliser for overcoming co-channel interference," in *Proc. Inst. Elect. Eng.—Commun.*, vol. 143, 1996, pp. 219–225.
- [6] E. Shamash and K. Yao, "On the structure and performance of a linear decision feedback equalizer based on the minimum error probability criterion," in *Proc. ICC'74*, 1974, pp. 25F1–25F5.
- [7] S. U. H. Qureshi, "Adaptive equalization," *Proc. IEEE*, vol. 73, pp. 1349–1387, Sept. 1985.
- [8] S. Chen, B. Mulgrew, E. S. Chng, and G. Gibson, "Space translation properties and the minimum-BER linear-combiner DFE," in *Proc. Inst. Elect. Eng.—Commun.*, vol. 145, 1998, pp. 316–322.
- [9] S. Chen and B. Mulgrew, "The minimum-SER linear-combiner decision feedback equaliser," in *Proc. Inst. Elect. Eng.—Commun.*, vol. 146, 1999, pp. 347–353.
- [10] S. Chen, S. Gunn, and C. J. Harris, "Decision feedback equalizer design using support vector machines," in *Proc. Inst. Elect. Eng.—Vision, Image, Signal Processing*, vol. 147, 2000, pp. 213–219.
- [11] R. A. Iltis, "A randomized bias technique for the importance sampling simulation of Bayesian equalizers," *IEEE Trans. Commun.*, vol. 43, pp. 1107–1115, Feb./Mar./Apr. 1995.
- [12] J. Moon and T. Jeon, "Sequence detection for binary ISI channels using signal-space partitioning," *IEEE Trans. Commun.*, vol. 47, pp. 891–901, July 1998.
- [13] Y. Kim and J. Moon, "Delay-constrained asymptotically optimal detection using signal-space partitioning," in *Proc. ICC'98*, Atlanta, GA, 1998.
- [14] —, "Multi-dimensional signal space partitioning using a minimal set of hyperplanes for detecting ISI-corrupted symbols," *IEEE Trans. Commun.*, vol. 48, pp. 637–647, Apr. 2000.
- [15] S. Chen, B. Mulgrew, and L. Hanzo, "Asymptotic Bayesian decision feedback equalizer using a set of hyperplanes," *IEEE Trans. Signal Processing*, vol. 48, pp. 3493–3500, Dec. 2000.
- [16] V. Vapnik, *The Nature of Statistical Learning Theory*. New York: Springer-Verlag, 1995.