Performance Evaluation of Bayesian Decision Feedback Equalizer with M-PAM Symbols Using Importance Sampling Simulation

S. Chen
Department of Electronics and Computer Science
University of Southampton, Southampton SO17 1BJ, U.K.

Abstract— **An importance sampling (IS) simulation method is presented for evaluating the lower-bound symbol error rate (SER) of the Bayesian decision feedback equalizer (DFE) with M-PAM symbols, under the assumption of correct decision feedback. By exploiting an asymptotic property of the Bayesian DFE, a design procedure is developed, which chooses appropriate bias vectors for the simulation density to ensure asymptotic efficiency (AE) of the IS simulation.**

I. Introduction

As the complexity of the Bayesian DFE [1]-[3] increases exponentially with the channel impulse response (CIR) length n_h and the symbol size M, SER evaluation under high signal-to-noise-ratio (SNR) conditions becomes impossible using a conventional Monte Carlo simulation. This paper considers SER evaluation of the Bayesian DFE using an IS simulation method. The idea of IS is that certain values of the input variables have more impact on the error rate than others and, by sampling these "important" values more frequently, the estimator variance can be reduced [4]. The issue is then how to choose a biased distribution to encourage the important regions of the input variables. One of the most effective IS techniques is the mean translation approach [5]–[7], where the distribution is moved toward the error region.

For binary symbols, Iltis [8] developed a randomized bias technique for the IS simulation of Bayesian equalizers. This IS simulation method was extended to evaluate the lower-bound (assuming correct decision feedback) bit error rate of the Bayesian DFE with binary symbols [9],[10]. For the M-PAM case, the asymptotic Bayesian decision boundary for separating any two neighbouring signal classes can be deduced [11]. By exploiting a symmetric distribution within each signal subset, the SER of the Bayesian DFE for the M-PAM case can be shown to be a scaled error rate of the equivalent "binary" Bayesian DFE evaluated on two neighbouring signal subsets. These two properties enable an extension of the IS simulation technique to the M-PAM case.

II. The Bayesian decision feedback equalizer

Consider the real-valued channel generates the received signal samples of:

$$y(k) = \sum_{i=0}^{n_h-1} h_i s(k-i) + n(k) \tag{1}$$

where h_i are the CIR taps, the Gaussian white noise $n(k)$ has zero mean and variance σ_n^2, and $s(k)$ takes the value from the symbol set $\mathcal{S} = \{s_i = 2i - M - 1,\ 1 \le i \le M\}$. The DFE uses the observed vector $\mathbf{y}(k) = [y(k) \cdots y(k-m+1)]^T$ and the past detected symbol vector $\hat{\mathbf{s}}_b(k) = [\hat{s}(k-d-1) \cdots \hat{s}(k-d-n_b)]^T$ to produce an estimate $\hat{s}(k-d)$ of $s(k-d)$, where d, m and n_b are the decision delay, feedforward and feedback orders, respectively. The choice of $d = n_h - 1$, $m = n_h$ and $n_b = n_h - 1$ will be used, as this choice guarantees a linear separability for different signal classes [12]. Let $\mathbf{s}_f(k) = [s(k) \cdots s(k-d)]^T$ and $\mathbf{s}_b(k) = [s(k-d-1) \cdots s(k-d-n_b)]^T$. Express $\mathbf{y}(k)$ as

$$\mathbf{y}(k) = \mathbf{H}_1\ \mathbf{s}_f(k) + \mathbf{H}_2\ \mathbf{s}_b(k) + \mathbf{n}(k) \tag{2}$$

where

$$\mathbf{H}_1 = \begin{bmatrix} h_0 & h_1 & \cdots & h_{n_h-1} \\ 0 & h_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & h_1 \\ 0 & \cdots & 0 & h_0 \end{bmatrix} \tag{3}$$

$$\mathbf{H}_2 = \begin{bmatrix} 0 & \cdots & 0 \\ h_{n_h-1} & \ddots & \vdots \\ \vdots & \ddots & 0 \\ h_1 & \cdots & h_{n_h-1} \end{bmatrix} \tag{4}$$

are the $m \times (d+1)$ and $m \times n_b$ CIR matrices, respectively. Assuming correct past decisions, we have

$$\mathbf{y}(k) = \mathbf{H}_1\ \mathbf{s}_f(k) + \mathbf{H}_2\ \hat{\mathbf{s}}_b(k) + \mathbf{n}(k)\,. \tag{5}$$

Thus the decision feedback translate $\mathbf{y}(k)$ into a new space:

$$\mathbf{r}(k) = \mathbf{y}(k) - \mathbf{H}_2\ \hat{\mathbf{s}}_b(k)\,. \tag{6}$$

Let the $N_f = M^{d+1}$ combinations of $\mathbf{s}_f(k)$ be $\mathbf{s}_{f,j}$, $1 \le j \le N_f$. The set of the noiseless channel states, namely, $R = \{\mathbf{r}_j = \mathbf{H}_1\ \mathbf{s}_{f,j},\ 1 \le j \le N_f\}$ can be partitioned into M subsets conditioned on $s(k-d)$:

$$R^{(i)} = \{\mathbf{r}_j \in R :\ s(k-d) = s_i\},\ 1 \le i \le M\,. \tag{7}$$

The optimal Bayesian DFE [3] involves computing the M decision variables for $1 \le i \le M$

$$\rho_i(\mathbf{r}(k)) = \sum_{\mathbf{r}_j \in R^{(i)}} \exp\left(-\frac{\|\mathbf{r}(k) - \mathbf{r}_j\|^2}{2\sigma_n^2}\right) \tag{8}$$

and making the decision according to

$$\hat{s}(k-d) = s_{i^*} \ \text{ with } \ i^* = \arg \max_{1 \le i \le M} \{\rho_i(\mathbf{r}(k))\} \, . \qquad (9)$$

A geometric translation property [11] is re-iterated here. For $1 \le i \le M-1$, $R^{(i+1)}$ is a translation of $R^{(i)}$ by the amount $2\mathbf{h}_{\text{rev}}$: $R^{(i+1)} = R^{(i)} + 2\mathbf{h}_{\text{rev}}$, where $\mathbf{h}_{\text{rev}} = [h_{n_h-1} \cdots h_1 \ h_0]^T$. This shifting property implies that, asymptotically when the SNR tends to infinity, the decision boundary $\mathcal{B}_{i+1}$ for separating $R^{(i+1)}$ and $R^{(i+2)}$ is a shift of $\mathcal{B}_i$ by an amount $2\mathbf{h}_{\text{rev}}$. Without the loss of generality, consider the two neighbouring subsets $R^{(M/2)}$ and $R^{((M/2)+1)}$. A pair of opposite-class channel states ($\mathbf{r}^{(+)} \in R^{((M/2)+1)}, \mathbf{r}^{(-)} \in R^{(M/2)}$ is said to be *Gabriel neighbours* if $\forall \mathbf{r}_j \in R^{(M/2)} \bigcup R^{((M/2)+1)}$, $\mathbf{r}_j \ne \mathbf{r}^{(+)}$ and $\mathbf{r}_j \ne \mathbf{r}^{(-)}$:

$$\|\mathbf{r}_j - \mathbf{r}_0\|^2 > \|\mathbf{r}^{(+)} - \mathbf{r}_0\|^2 \qquad (10)$$

where $\mathbf{r}_0 = (\mathbf{r}^{(+)} + \mathbf{r}^{(-)})/2$. The following lemma [11] describes the asymptotic decision boundary $\mathcal{B}_{M/2}$.

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

Consequently, a necessary condition for $\mathbf{r}_B \in \mathcal{B}_{M/2}$ is

$$\mathbf{r}_B = \frac{\mathbf{r}^{(+)} + \mathbf{r}^{(-)}}{2} + \left[\frac{\mathbf{r}^{(+)} - \mathbf{r}^{(-)}}{2}\right]^{\perp} \qquad (11)$$

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

$$\|\mathbf{r}_B - \mathbf{r}^{(+)}\|^2 < \|\mathbf{r}_B - \mathbf{r}_l\|^2, \ \forall \mathbf{r}_l \in R^{((M/2)+1)}, \ \mathbf{r}_l \ne \mathbf{r}^{(+)} \, , \qquad (12)$$

$$\|\mathbf{r}_B - \mathbf{r}^{(-)}\|^2 < \|\mathbf{r}_B - \mathbf{r}_j\|^2, \ \forall \mathbf{r}_j \in R^{(M/2)}, \ \mathbf{r}_j \ne \mathbf{r}^{(-)} \, , \qquad (13)$$

$$\|\mathbf{r}_B - \mathbf{r}^{(+)}\|^2 = \|\mathbf{r}_B - \mathbf{r}^{(-)}\|^2 \, . \qquad (14)$$

A simple algorithm can be used to select the set of all L pairs of Gabriel neighbours $\{\mathbf{r}_l^{(+)}, \mathbf{r}_l^{(-)}\}_{l=1}^{L}$ [8],[11].

Due to the symmetric distribution of $\mathcal{S}$, the states of $R^{(i)}$ are distributed *symmetrically* around the mass center of $R^{(i)}$. In particular, if a point $\mathbf{r}_j \in R^{(i)}$ has a distance x to the decision boundary $\mathcal{B}_{i-1}$, then there is another point $\mathbf{r}_l \in R^{(i)}$ with the same distance to the other decision boundary $\mathcal{B}_i$. Now consider the lower-bound SER for the Bayesian DFE

$$P_E = \text{Prob}\{\hat{s}(k-d) \ne s(k-d)\} \qquad (15)$$

First create a "binary" Bayesian DFE defined on $R^{(M/2)}$ and $R^{((M/2)+1)}$ with the decision function given by

$$f_b(\mathbf{r}(k)) = \sum_{\mathbf{r}_j \in R^{((M/2)+1)}} \exp\left(-\frac{\|\mathbf{r}(k) - \mathbf{r}_j\|^2}{2\sigma_n^2}\right) - \sum_{\mathbf{r}_l \in R^{(M/2)}} \exp\left(-\frac{\|\mathbf{r}(k) - \mathbf{r}_l\|^2}{2\sigma_n^2}\right) \qquad (16)$$

and the decision rule defined by

$$\hat{s}(k-d) = \begin{cases} 1, & \text{sgn}(f_b(\mathbf{r}(k))) \ge 0 \, , \\ -1, & \text{sgn}(f_b(\mathbf{r}(k))) < 0 \, . \end{cases} \qquad (17)$$

Denote the error probability of this "binary" Bayesian DFE as P_e. Taking into account of the shifting and symmetric properties discussed previously, it is straightforward to verify that $P_E = \gamma P_e$, with $\gamma = 2(M-1)/M$.

III. IS SIMULATION FOR THE M-PAM CASE

To evaluate the SER, P_E, of the Bayesian DFE with M-PAM symbols, it is only needed to evaluate the error probability, P_e, of the equivalent binary Bayesian DFE defined on $R^{(M/2)}$ and $R^{((M/2)+1)}$. The IS simulation technique [9],[10] can readily be used to evaluate P_e as follows:

$$\hat{P}_e = \frac{1}{N_s} \frac{1}{N_k} \sum_{j=1}^{N_s} \sum_{k=1}^{N_k} I_E(\mathbf{r}_j(k)) \frac{p(\mathbf{r}_j(k)|\mathbf{r}_j)}{p^*(\mathbf{r}_j(k)|\mathbf{r}_j)} \qquad (18)$$

where $I_E(\mathbf{r}(k)) = 1$ if $\mathbf{r}(k)$ causes an error, and $I_E(\mathbf{r}(k)) = 0$ otherwise; $p(\mathbf{r}_j(k)|\mathbf{r}_j)$ is the true conditional density given $\mathbf{r}_j \in R^{((M/2)+1)}$, and $N_s = M^d = N_f/M$ is the number of states in $R^{((M/2)+1)}$; the sample $\mathbf{r}_j(k)$ is generated using the simulation density $p^*(\mathbf{r}_j(k)|\mathbf{r}_j)$ chosen to be

$$p^*(\mathbf{r}_j(k)|\mathbf{r}_j) = \sum_{l=1}^{L_j} \frac{p_{l,j}}{(2\pi\sigma_n^2)^{\frac{m}{2}}} \exp\left(-\frac{\|\mathbf{r}_j(k) - \mathbf{v}_{l,j}\|^2}{2\sigma_n^2}\right) \, . \qquad (19)$$

In the simulation density (19), L_j is the number of the bias vectors $\mathbf{c}_{l,j} = -\mathbf{r}_j + \mathbf{v}_{l,j}$ for $\mathbf{r}_j \in R^{((M/2)+1)}$, $p_{l,j} \ge 0$ for $1 \le l \le L_j$, and $\sum_{l=1}^{L_j} p_{l,j} = 1$. An estimate of the IS gain for $\hat{P}_e$, which is defined as the ratio of the numbers of trials required for the same estimate variance using the Monte Carlo and IS methods, is given as [6],[8]:

$$\Gamma = \frac{\hat{P}_e(1 - \hat{P}_e)}{\hat{\eta} - \hat{P}_e^2} \qquad (20)$$

with

$$\hat{\eta} = \frac{1}{N_s} \frac{1}{N_k} \sum_{j=1}^{N_s} \sum_{k=1}^{N_k} I_E(\mathbf{r}_j(k)) \left(\frac{p(\mathbf{r}_j(k)|\mathbf{r}_j)}{p^*(\mathbf{r}_j(k)|\mathbf{r}_j)}\right)^2 \, . \qquad (21)$$

The IS simulated P_E is simply $\hat{P}_E = \gamma \hat{P}_e$, and the estimated IS gain for $\hat{P}_e$ will be used as the estimated IS gain for $\hat{P}_E$.

A design procedure is given for constructing the simulation density $p^*(\mathbf{r}_j(k)|\mathbf{r}_j)$ that meets the conditions for AE [6]. Let $\{\mathbf{r}_l^{(+)}, \mathbf{r}_l^{(-)}\}_{l=1}^{L}$ be the set of Gabriel neighbours selected from $R^{(M/2)}$ and $R^{((M/2)+1)}$. Each pairs $(\mathbf{r}_l^{(+)}, \mathbf{r}_l^{(-)})$ defines a hyperplane $H_l(\mathbf{r}) = \mathbf{w}_l^T \mathbf{r} + b_l = 0$ that is part of the asymptotic decision boundary $\mathcal{B}_{M/2}$, with

$$\mathbf{w}_l = \frac{2\left(\mathbf{r}_l^{(+)} - \mathbf{r}_l^{(-)}\right)}{\|\mathbf{r}_l^{(+)} - \mathbf{r}_l^{(-)}\|^2}\,, \tag{22}$$

$$b_l = -\frac{(\mathbf{r}_l^{(+)} - \mathbf{r}_l^{(-)})^T (\mathbf{r}_l^{(+)} + \mathbf{r}_l^{(-)})}{\|\mathbf{r}_l^{(+)} - \mathbf{r}_l^{(-)}\|^2}\,. \tag{23}$$

Note that H_l is a *canonical* hyperplane with $H_l(\mathbf{r}_l^{(+)}) = 1$ and $H_l(\mathbf{r}_l^{(-)}) = -1$. A state $\mathbf{r}_j^{(-)} \in R^{(M/2)}$ is *sufficiently separable* by the hyperplane H_l if $\mathbf{w}_l^T \mathbf{r}_j^{(-)} + b_l \le -1$. Similarly, $\mathbf{r}_j^{(+)} \in R^{((M/2)+1)}$ is sufficiently separable by H_l if $\mathbf{w}_l^T \mathbf{r}_j^{(+)} + b_l \ge 1$. The hyperplane H_l is *reachable* from $\mathbf{r}_j^{(+)} \in R^{((M/2)+1)}$ if the projection of $\mathbf{r}_j^{(+)}$ onto H_l is on $\mathcal{B}_{M/2}$. For each $\mathbf{r}_j^{(-)} \in R^{(M/2)}$, its separability index for H_l is $\alpha_{l,j}^{(-)} = 1$ if $\mathbf{r}_j^{(-)}$ is sufficiently separable by H_l; otherwise $\alpha_{l,j}^{(-)} = 0$. The separability index $\alpha_{l,j}^{(+)}$ for $\mathbf{r}_j^{(+)} \in R^{((M/2)+1)}$ is similarly defined. The reachability of H_l from $\mathbf{r}_j^{(+)} \in R^{((M/2)+1)}$ can be tested by computing

$$\mathbf{c}_{l,j} = -0.5\left(\mathbf{w}_l^T \mathbf{r}_j^{(+)} + b_l\right)\left(\mathbf{r}_l^{(+)} - \mathbf{r}_l^{(-)}\right)\,. \tag{24}$$

If $\mathbf{v}_{l,j} = \mathbf{r}_j^{(+)} + \mathbf{c}_{l,j} \in \mathcal{B}_{M/2}$ (i.e. $f_b(\mathbf{v}_{l,j}) = 0$), H_l is reachable from $\mathbf{r}_j^{(+)}$ ($\mathbf{c}_{l,j}$ is then a bias vector), and the reachability index is $\gamma_{l,j} = 1$; otherwise $\gamma_{l,j} = 0$. The whole process produces the following separability and reachability table:

	$\mathbf{r}_1^{(-)}$	$\cdots$	$\mathbf{r}_{N_s}^{(-)}$	$\mathbf{r}_1^{(+)}$	$\cdots$	$\mathbf{r}_{N_s}^{(+)}$
H_1	$\alpha_{1,1}^{(-)}$	$\cdots$	$\alpha_{1,N_s}^{(-)}$	$\alpha_{1,1}^{(+)}\,(\gamma_{1,1})$	$\cdots$	$\alpha_{1,N_s}^{(+)}\,(\gamma_{1,N_s})$
$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
H_L	$\alpha_{L,1}^{(-)}$	$\cdots$	$\alpha_{L,N_s}^{(-)}$	$\alpha_{L,1}^{(+)}\,(\gamma_{L,1})$	$\cdots$	$\alpha_{L,N_s}^{(+)}\,(\gamma_{L,N_s})$

For each $\mathbf{r}_j^{(+)} \in R^{((M/2)+1)}$, select those hyperplanes that can sufficiently separate $\mathbf{r}_j^{(+)}$ and are reachable from $\mathbf{r}_j^{(+)}$ with the aid of the above table. This yields the integer set:

$$G_j^{(+)} = \{q : \alpha_{q,j}^{(+)} = 1 \text{ and } \gamma_{q,j} = 1\}\,. \tag{25}$$

The convex region $\mathcal{R}_j^{(+)}$ covering $\mathbf{r}_j^{(+)}$ is the intersection of all the half-spaces $\mathcal{H}_q^{(+)} = \{\mathbf{r} : H_q(\mathbf{r}) \ge 0\}$ with $q \in G_j^{(+)}$. In fact, it is not necessary to use every hyperplanes defined in $G_j^{(+)}$ to form $\mathcal{R}_j^{(+)}$. A subset of these hyperplanes will be sufficient, provided that every opposite-class state in $R^{(M/2)}$ can sufficiently be separated by at least one hyperplane in the subset. If this can be done, the error region $\mathcal{E}$ satisfies

$$\mathcal{E} \subset \overline{\mathcal{R}_j^{(+)}} = \bigcup_{q \in G_j^{(+)}} \mathcal{H}_q^{(-)} \tag{26}$$

with the half-spaces $\mathcal{H}_q^{(-)} = \{\mathbf{r} : H_q(\mathbf{r}) < 0\}$. Obviously, all the hyperplanes defined in $G_j^{(+)}$ are reachable from $\mathbf{r}_j^{(+)}$ and at least one of $\{\mathbf{v}_{q,j}\}$ is the minimum rate point (as defined in [6]). If $G_j^{(+)}$ exists for each $\mathbf{r}_j^{(+)} \in R^{((M/2)+1)}$, the simulation density constructed with the bias vectors $\{\mathbf{c}_{q,j}\}$, $q \in G_j^{(+)}$, for all j will guarantee AE.

IV. Simulation Examples

Example 1. A 2-tap channel $\mathbf{h} = [0.3\ 1.0]^T$ with 8-PAM symbols was simulated, given $m = 2$, $d = 1$ and $n_b = 1$. The set R had 64 states. Nine pairs of Gabriel neighbours were selected from $R^{(4)}$ and $R^{(5)}$, leading to the separability and reachability table from which an AE simulation density was constructed. The simulation density construction is illustrated in Fig. 1. The bias vectors were selected with uniform probability in the simulation ($p_{l,j} = 1/L_j$). For each SNR, 10^5 iterations were used for each state in $\mathcal{R}^{(5)}$. Thus, the total samples used for a given SNR were 8×10^5. Fig. 2 (a) shows the lower-bound SERs obtained using the IS and conventional sampling (CS) simulations, respectively. It can be seen that the conventional Monte Carlo results for low SNR conditions based directly on the Bayesian DFE of (8) and (9) agreed with those of the IS simulation. The estimated IS gains, depicted in Fig. 2 (b), indicate that exponential IS gains were obtained with increasing SNRs.

Example 2. A 3-tap channel $\mathbf{h} = [0.3\ 1.0\ -0.3]^T$ with 8-PAM symbols was tested, given $m = 3$, $d = 2$ and $n_b = 2$. The set R had 512 states. Nineteen pairs of Gabriel neighbours were found from $R^{(4)}$ and $R^{(5)}$, leading to the separability and reachability table from which an AE simulation density was constructed. Again the bias vectors were selected with uniform probability in the simulation. For each SNR, 10^4 samples were used for each state in $\mathcal{R}^{(5)}$, resulting in a total of 6.4×10^5 samples for a given SNR. Fig. 3 (a) depicts the lower-bound SERs obtained using the IS and CS simulations, respectively. Again, the conventional Monte Carlo results for low SNR conditions agreed with those of the

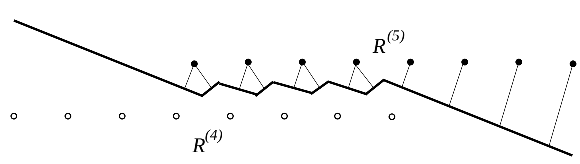

Fig. 1. Illustration of the simulation density construction for the channel $\mathbf{h} = [0.3\ 1.0]^T$ with 8-PAM symbols.

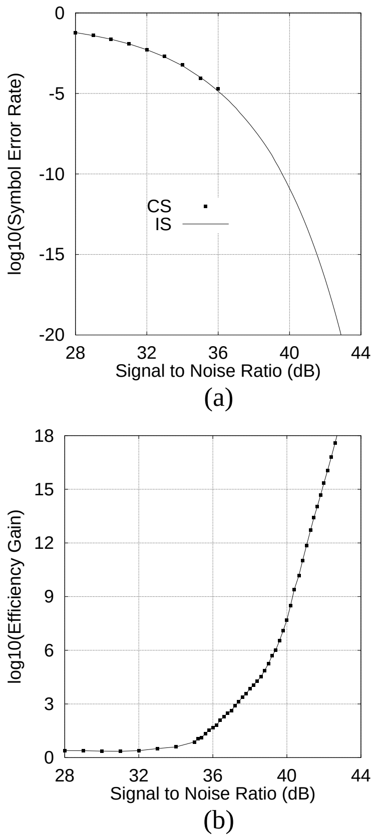

Fig. 2. The lower-bound SERs (a) and the estimated IS gain (b) of the Bayesian DFE for $\mathbf{h} = [0.3\ 1.0]^T$ with 8-PAM symbols.

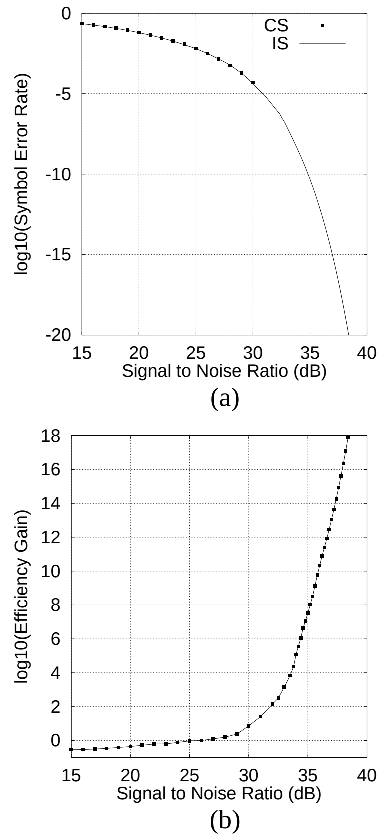

Fig. 3. The lower-bound SERs (a) and the estimated IS gain (b) of the Bayesian DFE for $\mathbf{h} = [0.3\ 1.0\ -0.3]^T$ with 8-PAM symbols.

IS simulation. It can be seen from Fig. 3 (b) that exponential IS gains were obtained with increasing SNRs.

V. Conclusions

An IS simulation has been extended to evaluate the lower-bound SER of the Bayesian DFE with M-PAM symbols. It has been noted that the Bayesian decision boundary separating any two neighbouring signal classes is asymptotically piecewise linear. Furthermore, the SER of the Bayesian DFE for the M-PAM case is a scaled error rate of the equivalent binary Bayesian DFE evaluated on any two neighbouring signal subsets. A design procedure has been presented for constructing the simulation density that meets the asymptotic efficiency conditions.

References

[1] D. Williamson, R.A. Kennedy and G.W. Pulford, "Block decision feedback equalization," *IEEE Trans. Communications*, Vol.40, No.2, pp.255-264, 1992.

[2] S. Chen, B. Mulgrew and S. McLaughlin, "Adaptive Bayesian equaliser with decision feedback," *IEEE Trans. Signal Processing*, Vol.41, No.9, pp.2918–2927, 1993.

[3] S. Chen, S. McLaughlin, B. Mulgrew and P.M. Grant, "Bayesian decision feedback equaliser for overcoming co-channel interference," *IEE Proc. Communications*, Vol.143, No.4, pp.219–225, 1996.

[4] P.J. Smith, M. Shafi and H. Gao, "Quick simulation: a review of importance sampling techniques in communications systems," *IEEE J. Selected Areas in Communications*, Vol.15, No.4, pp.597–613, 1997.

[5] D. Lu and K. Yao,"Improved importance sampling techniques for efficient simulation of digital communication systems," *IEEE J. Selected Areas in Communications*, Vol.6, No.1, pp.67–75, 1988.

[6] J.S. Sadowsky and J.A. Bucklew, "On large deviations theory and asymptotically efficient Monte Carlo estimation," *IEEE Trans. Information Theory*, Vol.36, No.3, pp.579–588, 1990.

[7] H.J. Schlebusch, "On the asymptotic efficiency of importance sampling techniques," *IEEE Trans. Information Theory*, Vol.39, pp.710–715, 1993.

[8] R.A. Iltis, "A randomized bias technique for the importance sampling simulation of Bayesian equalizers," *IEEE Trans. Communications*, Vol.43, No.2/3/4, pp.1107–1115, 1995.

[9] S. Chen,"Importance sampling simulation for evaluating the lower-bound BER of the Bayesian DFE," *IEEE Trans. Communications*, accepted, 2001.

[10] S. Chen and L. Hanzo, "An importance sampling simulation method for Bayesian decision feedback equalizers," in *Proc. 5th Int. Conf. Mathematics in Signal Processing* (University of Warwick, UK), Dec.18-20, 2000, 4 pages.

[11] S. Chen, L. Hanzo and B. Mulgrew, "Decision feedback equalization using multiple-hyperplane partitioning for detecting ISI-corrupted M-ary PAM signals," *IEEE Trans. Communications*, to appear, May 2001.

[12] S. Chen, B. Mulgrew, E.S. Chng and G. Gibson, "Space translation properties and the minimum-BER linear-combiner DFE," *IEE Proc. Communications*, Vol.145, No.5, pp.316–322, 1998.