UNBIASED ADAPTIVE DECISION FEEDBACK EQUALIZATION

Hyun-Chool Shin and Woo-Jin Song

Division of Electronics and Computer Engineering
Pohang University of Science and Technology (POSTECH)
Pohang, Kyungbuk, 790-784, Republic of Korea
e-mail:wjsong@postech.ac.kr

ABSTRACT

It is well-known that the decision rule in the minimum mean-squares-error decision feedback equalizer (MMSE-DFE) is biased, and therefore suboptimum with respect to error probability. We present a new family of algorithms that solve the bias problem in the adaptive DFE. A novel constraint, called the ***constant-norm*** constraint, is introduced unifying the quadratic constraint and the monic one. A new cost function based on the constant-norm constraint and Lagrange multiplier is defined. Minimizing the cost function gives birth to a new family of unbiased adaptive DFE. The simulation results demonstrate that the proposed method indeed produce unbiased solution in the presence of noise while keeping very simple both in computation and implementation.

1. INTRODUCTION

In wireless communication systems, the multi-path channel is a major obstacle in reliable communication. The limited bandwidth of the communication channel introduces time dispersion effects which result in spreading of the energy of one pulse into neighboring pulse intervals, which is known as intersymbol interference (ISI) [1]–[2]. Numerous works have been done to solve this problem. A method of reducing the degrading effects of the channel that has found in many practical applications is to use decision feedback equalizer (DFE).

Since its introduction by Austin [3], the DFE receiver structure has received considerable attention from many researchers due to its improved performance over the linear equalizer and reduced implementation complexity as compared with the nonlinear maximum-likelihood receiver.

The equalizer coefficients are determined so that the performance of the DFE is optimized according to a chosen performance criterion. The most relevant criterion would be the probability of error. However, this criterion complicates the analysis and does not offer further insight into the problem. So, the mean-square error (MSE) criterion has been widely used. This choice simplifies the analysis and leads to an easy adaptive implementation. However, the DFE with the minimum MSE (MMSE-DFE) is biased due to noise and therefore suboptimum with respect to error probability [4]–[5]. It was shown in [5] that unbiased DFE has the lowest probability error among all possible symbol-by-symbol decisions. So, it is obvious that removing the bias in DFE receiver is one of the most important issues.

In this paper, by imposing the *constant-norm* constraint on filter coefficients, we solve the bias problem in the MMSE-DFE. A new cost function based on the *constant-norm* constraint and the Lagrange multiplier is defined. By minimizing the cost function a new unbiased adaptive DFE algorithm is derived.

The remainder of this paper is organized as follows. Section 2 describes the system model and a new unbiased DFE based on the *constant-norm constraint* is presented. Then Section 3 gives a stochastic gradient algorithm for unbiased adaptive DFE and its stationary points analysis and studies its convergence properties when stochastic gradient search is used. Section 4 evaluates performance of the proposed algorithm by computer simulation. Finally conclusions are presented in Section 5.

2. CONSTRAINED OPTIMIZATION FOR UNBIASED DFE

The discrete-time baseband equivalent transmission system incorporating a decision feedback equalizer is shown in Fig. 1. Let a zero-mean discrete-time signal $x(n)$ be the transmitted data sequence. Assuming that the channel impulse response h_k exists only over the finite time interval $(-K \leq k \leq L)$, then the input-output relation for the discrete-time equivalent channel has the

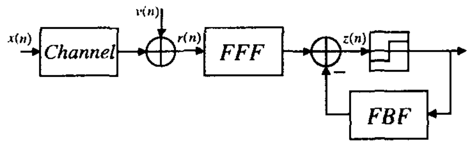

Figure 1: Proposed unbiased DFE

form

$$y(n) = \sum_{k=-K}^{L} h_k x(n-k) \qquad (1)$$

and the observed input to the DFE is given by

$$r(n) = y(n) + v(n), \qquad (2)$$

where $v(n)$ is white measurement noise with variance σ_v^2 and is independent of the data $x(n)$.

As shown in Fig. 1, the DFE consists of two symbol-spaced FIR filters: a feedforward filter (FFF) that has N_f taps, and a feedback filter (FBF) that has N_b strictly causal taps.

In this paper, the FIR filters are expressed as vectors. The FFF is denoted by $\mathbf{f}$, where $\mathbf{f}^T = [f_0\ f_1\ \cdots\ f_{(N_f-1)}]$. Concerning the FBF, for analytical convenience the augmented vector $\mathbf{b}$ is used as $\mathbf{b}^T = [1\ b_1\ \cdots\ b_{N_b}]$. The actual FBF coefficients are $\{-b_1, -b_2, \cdots, -b_{N_b}\}$. From Fig. 1, it is straitforward to verify that under the assumption of correct past decisions, the input to the decision device is given by

$$z(n) = \sum_{k=0}^{N_f-1} f_k r(n-k) - \sum_{l=1}^{N_b} b_l x(n-\Delta-l), \qquad (3)$$

where $\Delta (0 \leq \Delta \leq N_f - 1)$ is the delay of the FFF. Then the error signal is expressed as

$$\begin{aligned} e(n) &= x(n-\Delta) - z(n) \\ &= \mathbf{x}^T(n)\mathbf{b} - \mathbf{r}^T(n)\mathbf{f}, \end{aligned} \qquad (4)$$

where $\mathbf{x}^T(n) = [x(n-\Delta)\ x(n-\Delta-1)\ \cdots\ x(n-\Delta-N_b)]$ and $\mathbf{r}^T(n) = [r(n)\ r(n-1)\ \cdots\ r(n-N_f+1)]$.

Let $e_s(n)$ be a noise-free error defined by

$$e_s(n) = \mathbf{x}^T(n)\mathbf{b} - \mathbf{y}^T(n)\mathbf{f}. \qquad (5)$$

Since signal and noise are independent [6], squaring and taking expectation of (5) give the expression of the mean-square error (MSE)

$$E[e^2(n)] = E[e_s^2(n)] + \sigma_v^2 \left(\sum_{k=0}^{N_f-1} f_k^2 \right). \qquad (6)$$

In minimizing the MSE, clearly the unwanted second term on the right hand side of (6) adds a penalty function proportional to $||\mathbf{f}||^2$ where $||f||$ denotes the Euclidean norm of a vector. This introduces an undesirable bias which depends on the noise power.

This bias problem is overcome by using the *constant-norm* constraint

$$\sum_{k=0}^{N_f-1} f_k^2 = C. \qquad (7)$$

With this constraint, the MSE in (6) becomes

$$E[e^2(n)] = E[e_s^2(n)] + C\sigma_v^2. \qquad (8)$$

Since $C\sigma_v^2$ is a constant, minimizing (8) is equivalent to minimizing the MSE under the noise-free condition. This in turn produces a unbiased solution.

The optimum solution is obtained by minimizing the MSE subject to (7). The augmented cost function, with the *constant-norm* constraint, using Lagrange multiplier λ, is given by

$$J = \frac{E\left[e^2(n)\right] + \lambda\left(C - \mathbf{f}^T\mathbf{f}\right)}{2}. \qquad (9)$$

Taking derivatives of J with respect to $\mathbf{f}$, $\mathbf{b}$, and λ results in

$$\frac{\partial J}{\partial \mathbf{f}} = -R_{rr}\mathbf{f} + R_{rx}\mathbf{b} - \lambda\mathbf{f}, \qquad (10)$$

$$\frac{\partial J}{\partial \mathbf{b}} = R_{xr}\mathbf{f} - R_{xx}\mathbf{b}, \qquad (11)$$

and

$$\frac{\partial J}{\partial \lambda} = C - \mathbf{f}^T\mathbf{f}, \qquad (12)$$

where $R_{rr} = E\left[\mathbf{r}(n)\mathbf{r}^T(n)\right]$, $R_{rx} = E\left[\mathbf{r}(n)\mathbf{x}^T(n)\right]$, $R_{xr} = E\left[\mathbf{x}(n)\mathbf{r}^T(n)\right]$, and $R_{xx} = E\left[\mathbf{x}(n)\mathbf{x}^T(n)\right]$.

Let $\mathbf{f}_o$ and $\mathbf{b}_o$ be solutions for which the derivatives of J are zero. Expressing $\mathbf{b}_o$ in terms of $\mathbf{f}_o$ in (11) and substituting it into (10) yield

$$\left(-R_{rr} + R_{rx}R_{xx}^{-1}R_{xr}\right)\mathbf{f}_o = \lambda\mathbf{f}_o. \qquad (13)$$

It follows from (13) that $\mathbf{f}_o$ is an eigenvector of the matrix $\left(R_{rr} - R_{rx}R_{xx}^{-1}R_{xr}\right)$ with the corresponding eigenvalue λ [7]. It is obvious that the equation in (13) holds even if we multiply the vector $\mathbf{f}_o$ to both sides of (13). Then λ can be obtained as

$$\lambda = \frac{\mathbf{f}_o^T\left(R_{rr} - R_{rx}R_{xx}^{-1}R_{xr}\right)\mathbf{f}_o}{\mathbf{f}_o^T\mathbf{f}_o} \qquad (14)$$

since $\mathbf{f}_o^T\mathbf{f}_o \neq 0$.

3. STOCASTIC GRADIENT ALGORITHM FOR UNBIASED ADAPTIVE DFE

To adapt the FIR filters $\mathbf{f}$ and $\mathbf{b}$, we shall consider here the least mean square (LMS) algorithm. Global convergence is guaranteed because there is only a single global minimum in the constrained MSE surface [6][8].

Replacing R_{rr} and R_{rx} by their respective values $E[\mathbf{r}(n)\mathbf{r}^T(n)]$ and $E[\mathbf{r}(n)\mathbf{x}^T(n)]$ in (10) and (11) produces

$$\frac{\partial J}{\partial \mathbf{f}(n)} = E[\mathbf{r}(n)e(n)] - \lambda \mathbf{f}(n) \quad (15)$$

and

$$\frac{\partial J}{\partial \mathbf{b}(n)} = -E[\mathbf{x}(n)e(n)], \quad (16)$$

since $e(n) = \mathbf{x}^T(n)\mathbf{b}(n) - \mathbf{r}^T(n)\mathbf{f}(n)$. Also, since $\mathbf{b}_o = R_{xx}^{-1}R_{xr}\mathbf{f}_o$, we can rewrite λ in (14) as

$$\lambda = \frac{\mathbf{f}_o^T(-R_{rr}\mathbf{f}_o + R_{rx}\mathbf{b}_o)}{\mathbf{f}_o^T\mathbf{f}_o}. \quad (17)$$

Since the exact value of λ is unknown a priori, we use an estimated eigenvalue for coefficients update.

Replacing $\mathbf{f}_o$ and $\mathbf{b}_o$ with $\mathbf{f}(n)$ and $\mathbf{b}(n)$, respectively, then an estimated eigenvalue is given by

$$\hat{\lambda}(n) = \frac{\mathbf{f}^T(n)(-R_{rr}\mathbf{f}(n) + R_{rx}\mathbf{b}(n))}{\mathbf{f}^T(n)\mathbf{f}(n)}. \quad (18)$$

Since $-R_{rr}\mathbf{f}(n) + R_{rx}\mathbf{b}(n) = E[\mathbf{r}(n)e(n)]$, the equation (18) can be rewritten as

$$\hat{\lambda}(n) = \frac{\mathbf{f}^T(n)E[\mathbf{r}(n)e(n)]}{\mathbf{f}^T(n)\mathbf{f}(n)}. \quad (19)$$

Using $\hat{\lambda}(n)$ in (19) instead of λ in $\partial J/\partial \mathbf{f}$, an estimated gradient is formulated as

$$\frac{\hat{\partial J}}{\partial \mathbf{f}(n)} = E[\mathbf{r}(n)e(n)] - \frac{\mathbf{f}^T(n)E[\mathbf{r}(n)e(n)]}{\mathbf{f}^T(n)\mathbf{f}(n)}\mathbf{f}(n). \quad (20)$$

Replacing the expected values by respective instantaneous values in (20) and (16), we have the following stochastic gradient based update equations:

$$\begin{cases} \mathbf{f}(n+1) = \mathbf{f}(n) - \mu e(n)\left(\mathbf{r}(n) - \frac{\mathbf{f}^T(n)\mathbf{r}(n)}{\mathbf{f}^T(n)\mathbf{f}(n)}\mathbf{f}(n)\right) \\ \mathbf{b}(n+1) = \mathbf{b}(n) + \mu e(n)\mathbf{x}(n), \end{cases} \quad (21)$$

where μ is the step size that controls convergence of the adaptation process. Since

$$\mathbf{f}^T(n+1)\mathbf{f}(n+1) = \mathbf{f}^T(n)\mathbf{f}(n) + O(\mu^2),$$

where the symbol $O(\mu^2)$ denotes the terms which are in the order of μ^2 or higher, computation of $\mathbf{f}^T(n)\mathbf{f}(n)$ is only needed after a fixed number of iterations to suppress the accumulated error due to $O(\mu^2)$.

By using $R_{rr} = R_{yy} + \sigma_v^2 \mathrm{I}_{L+1}$ and $R_{rx} = R_{yx}$ in (18) where I_{L+1} is the identity matrix of size $L+1$, the equation (18) becomes

$$\begin{aligned} \hat{\lambda}(n) &= \frac{\mathbf{f}^T(n)\sigma_v^2\mathrm{I}_{L+1}\mathbf{f}(n)}{\mathbf{f}^T(n)\mathbf{f}(n)} + \frac{\mathbf{f}^T(R_{yy}\mathbf{f}(n) - R_{yx}\mathbf{b}(n))}{\mathbf{f}^T(n)\mathbf{f}(n)} \\ &= \sigma_v^2 + \frac{\mathbf{f}(n)^T(R_{yy}\mathbf{f}(n) - R_{yx}\mathbf{b}(n))}{\mathbf{f}^T(n)\mathbf{f}(n)}. \end{aligned} \quad (22)$$

As $E[\mathbf{y}(n)e_s(n)]$ converges to $\mathbf{0}$, $\hat{\lambda}(n)$ becomes σ_v^2. So, $\hat{\lambda}(n)$ can be interpreted as an estimate of noise power, i.e., $\hat{\lambda}(n) = \hat{\sigma}_v^2(n)$.

In the estimated gradient of (20), the second term on the right hand side, which is the coefficient vector $\mathbf{f}(n)$ weighted by the estimated eigenvalue $\hat{\lambda}(n)$, i.e., the estimated *noise* power $\hat{\sigma}_v^2(n)$, is newly introduced, compared with the conventional equation error algorithm. The first term on the right hand side in (20) which is the gradient in the conventional equation error algorithm can be represented as

$$E[\mathbf{r}(n)e(n)] = E[\mathbf{y}(n)e_s(n)] + \sigma_v^2\mathbf{f}(n). \quad (23)$$

As can be seen in (23), the bias-induced term is the $\sigma_v^2\mathbf{f}(n)$ which corrupts the coefficient update. In the proposed method the bias-induced term $\sigma_v^2\mathbf{f}(n)$ is compensated by the novel correction term $\hat{\sigma}_v^2(n)\mathbf{f}(n)$ and thus the unbiased solution can be obtained.

The constraint in (7) does not mean that C is constant for the whole adaptation period. The value $C(n) = \mathbf{f}^T(n)\mathbf{f}(n)$ converges to $\mathbf{f}_o^T\mathbf{f}_o$ as the iteration goes since $\mathbf{f}(n)$ converges to $\mathbf{f}_o$.

4. SIMULATION RESULTS

We evaluate the performance of the unbiased adaptive equalizer by computer simulation in comparison with the conventional LMS DFE. The discrete-time channel model used for the simulation is given by

$$\begin{aligned} H(z) &= \sum_{k=-5}^{13} h_k z^{-k} \\ &= .3z^5 + 1 + .3z^{-1} + .1z^{-3} + .2z^{-13} \end{aligned} \quad (24)$$

and the transmitted data is 8-level symbol with zero-mean and variance 21, i.e.,

$$x(n) \in \{-7, -5, -3, -1, 1, 3, 5, 7\}.$$

The measurement noise is white Gaussian with variance σ_v^2. The signal-to-noise ratio (SNR) is calculated by

$$\mathrm{SNR} = 10log\left(\frac{E[y^2(n)]}{E[v^2(n)]}\right).$$

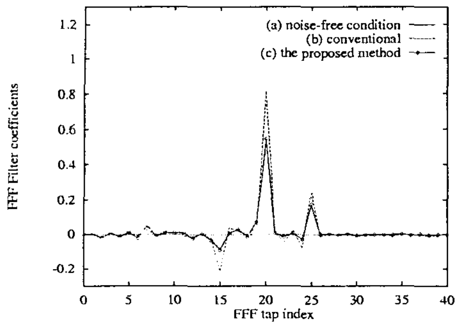

Figure 2: FFF coefficients (SNR=10dB)

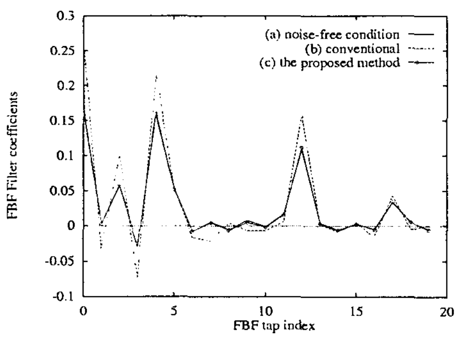

Figure 3: FBF coefficients (SNR=10dB)

The DFE consists of 40 FFF taps and 20 FBF taps and Δ is set to 20. For coefficients update, we use a known training sequence, i.e., training mode adaptation. The step-size μ is chosen as 0.00001 and the SNR is 10dB.

Fig. 2 and Fig. 3 show the filter coefficients of DFE after adaptation. For the comparison purpose, the result under the noise-free condition, i.e., $v(n) = 0$ is plotted. As can be seen in Fig. 2 and Fig. 3, the resulting filter coefficients by the proposed method are almost the same as those under the noise-free condition. These facts demonstrate that the proposed method gives indeed unbiased solution immune to noise. For quantitative description of the debiasing capability of the proposed method, we define the norm squared parameter error as

$$||\mathbf{f}(n) - \mathbf{f}_o||^2 + ||\mathbf{b}(n) - \mathbf{b}_o||^2,$$

where $\mathbf{f}_o$ and $\mathbf{b}_o$ are the MMSE solutions of FFF and FBF under the noise-free condition, respectively. Fig. 4 shows the results.

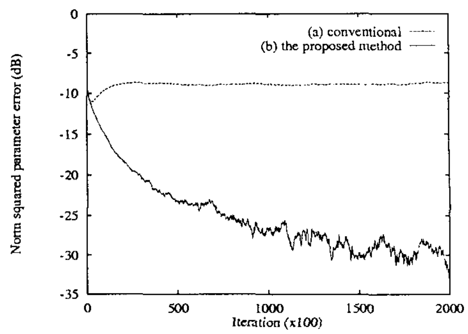

Figure 4: Norm squared parameter error (SNR=10dB)

5. CONCLUSIONS

We have presented an efficient unbiased adaptive DFE based on the *constant-norm* constrained optimization. Although the proposed algorithm is developed on the assumption that the FFF is symbol-spaced, the same algorithm formular is derived for fractionally-spaced DFE receiver.

6. REFERENCES

[1] J. G. Proakis, *Digital Communications*, New York: McGraw-Hill, 1983.

[2] S. U. H. Qureshi, "Adaptive Equalization," *Proc. IEEE*, vol. 73, no. 9, pp. 1349–1389, Sep. 1985.

[3] M. E. Austin, "Decision feedback equalization for digital communication over dispersive channels," *MIT Lincoln Lab., Lexington, MA*, Tech. Rep. 437, Aug., 1967.

[4] N. Al-Dhahir, J. M. Cioffi, "MMSE decision-feedback equalizers: finite-length results," *IEEE Trans. Inform. Theory*, vol. 41, no. 4, July. 1995.

[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, no. 10, Oct. 1995.

[6] S. Haykin, *Adaptive Filter Theory*, Englewood Cliffs, NJ: Prentice Hall, 1991.

[7] P. R. Regalia, "An unbiased equation error identifier and reduced-order approximations," *IEEE Trans. Signal Processing*, vol. 42, pp. 1397–1412, June 1994.

[8] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, Englewood Cliffs, NJ: Prentice Hall, 1985.

[9] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Walton Street, Oxford: Oxford University Press, 1965.

Acknowledgments

This work was supported in part by the Brain Korea (BK) 21 Program funded by the Ministry of Education, Republic of Korea.