POWER-OF-TWO QUANTIZATION IN DECISION FEEDBACK EQUALIZATION

Mansour A. Aldajani

Systems Engineering Department
King Fahd University of Petroleum and Minerals
Dhahran 31261, Saudi Arabia
dajani@ccse.kfupm.edu.sa

ABSTRACT

In this work, we propose and efficient structure for decision feedback equalization (DFE). The proposed structure makes use of the power-of-two quantization concept to completely eliminate the real multiplications in the DFE's update equations. The resulting structure is shown to substantially reduce the complexity of the DFE without any loss of performance. The study includes performance analysis of the proposed method and closed form expressions for the mean square errors.

1. INTRODUCTION

Decision feedback equalization (DFE) is a commonly used practice to mitigate the effect of Inter-symbol interference (ISI) in communication systems [1]. However, one of the main challenges in using DFE in high speed applications is the computation burden it involves.

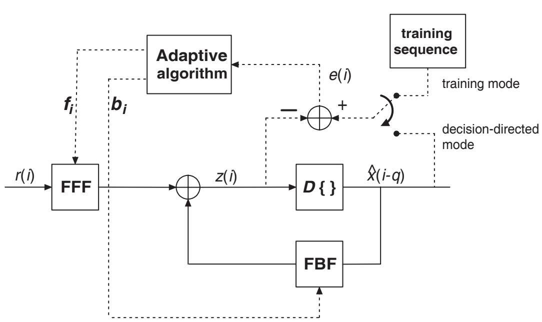

Fig. 1. *Structure of the Decision Feedback Equalizer.*

Fig. 1 shows the structure of the DFE. The main components of the DFE are the feed forward filter (FFF), feedback filter (FBF), and the decision device function $D\{\}$. The coefficients of the DFE are updated according to the recursion

$$\boldsymbol{f}_i = \boldsymbol{f}_{i-1} + \mu \boldsymbol{r}^*_{i+M-1} e(i) \qquad (M_f \times 1) \tag{1}$$

$$\boldsymbol{b}_i = \boldsymbol{b}_{i-1} + \mu \widehat{\boldsymbol{x}}^*_{i-1} e(i) \qquad (M_b \times 1) \tag{2}$$

where

$$z(i) = \boldsymbol{f}_i \boldsymbol{r}_i + \boldsymbol{b}_i \widehat{\boldsymbol{x}}_i \tag{3}$$

$$\widehat{x}(i) = D\{z(i)\} \tag{4}$$

$$e(i) = \widehat{x}(i-q) - z(i) \tag{5}$$

$$\tag{6}$$

The vectors $\boldsymbol{f}_i$ and $\boldsymbol{b}_i$ denote the coefficients of the FFF and FBF respectively. The vectors $\boldsymbol{r}_i$ $(1 \times M_f)$ and $\widehat{\boldsymbol{x}}_i$ $(1 \times M_b)$ contain the latest M_f and M_b values of the received signal $r(i)$ and the decision samples $\widehat{x}(i)$ respectively. μ is the step-size of the adaptation while q is the equalizer's delay and $M = M_b + M_f$. The star $(*)$ indicates the conjugate transpose. If we denote

$$\boldsymbol{u}_i \triangleq [\boldsymbol{r}_{i+M-1} \quad \widehat{\boldsymbol{x}}_{i-1}] \tag{7}$$

$$\boldsymbol{w}_i \triangleq [\boldsymbol{f}_i; \boldsymbol{b}_i] \qquad (M \times 1) \tag{8}$$

$$d(i) \triangleq \widehat{x}(i-1) \tag{9}$$

Then, we can write the DFE equations in the standard LMS form

$$e(i) = d(i) - \boldsymbol{u}_i \boldsymbol{w}_{i-1} \tag{10}$$

$$\boldsymbol{w}_i = \boldsymbol{w}_{i-1} + \mu \boldsymbol{u}^*_i e(i). \tag{11}$$

One way to reduce the complexity of DFE is by avoiding the multiplication $\boldsymbol{u}^*_i e(i)$ that appears in the equalizer recursion (11). This can be achieved by rounding the element $e(i)$ to its nearest power-of-two value. In this case, the binary representation of the quantized values will have only one bit that is "1" and the rest are zeros. This means that the multiplication can be attained by, at most, simple shift operations of the vector $\boldsymbol{u}^*_i$. In this way, real multiplication can be completely avoided in the recursion. This technique is

This work was supported by King Fahd University of Petroleum and Minerals.

known as the power-of-two quantization (PTQ). This technique was successfully adopted in reducing the complexity of *linear* equalizers [2, 3, 4]. In these articles, the PTQ was shown to substantially reduce the complexity of the equalizer without major loss in performance.

In this work, we consider the implementation of the PTQ technique to reduce the complexity of the DFE. The objective is to eliminate the expensive multiplication operations from the recursion of the equalizer without compromising its performance.

2. POWER-OF-TOW QUANTIZATION

The power-of-two quantization for a given signal x can be mathematically expressed as [5]

$$Q_{PTQ}(x) = 2^{\lfloor \log_2 |x| \rfloor} \text{sign}(x) \tag{12}$$

where $\lfloor . \rfloor$ denotes the *floor* operation. Fig. 2 shows the block diagram representation of this quantizer. The signum function carries the sign information of the input while the magnitude information is passed through log-2, linear quantization, and exponential blocks sequentially. The linear quantizer produces the nearest integer value less than its input sample.

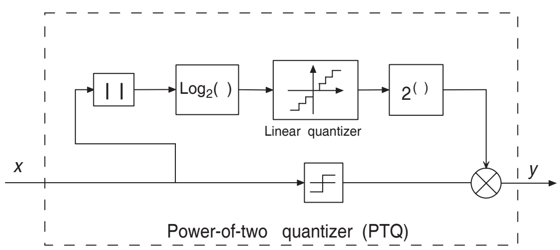

Fig. 2. *Block diagram representation of the power-of-two quantizer.*

The PTQ in (12) assumes infinite number of bits for the linear quantizer. Another finite bit version is given by [2]

$$Q_{PTQ}(x) = \begin{cases} \text{sign}(x) & |x| \geq 1 \\ 2^{\lfloor \log_2 |x| \rfloor} & 2^{-B+1} \leq |x| < 1 \\ 0 & |x| < 2^{-B+1} \end{cases} \tag{13}$$

where B is the number of bits assigned to the quantization. In [3] and [4], the third line in this expression is changed to

$$Q_{PTQ}(x) = 2^{-B+1} \text{sign}(x) \qquad |x| < 2^{-B+1} \tag{14}$$

to avoid the adaptation dead-zone when $|x| \leq 2^{-B+1}$.

3. STRUCTURE OF THE PROPOSED PTQ-DFE

In this work we implement the PTQ concept to the conventional structure of the DFE shown in Fig. 1 and represented by the recursions (1) and (2). For ease of analysis, the linear quantizer in this study uses rounding to nearest integer instead of the floor operation.

The error signal $e(i)$ is quantized using the PTQ and the update equations of the DFE coefficients become

$$\boldsymbol{f}_i = \boldsymbol{f}_{i-1} + \mu \boldsymbol{r}^*_{i+M-1} Q_{PTQ}\{e(i)\} \quad (M_f \times 1) \tag{15}$$
$$\boldsymbol{b}_i = \boldsymbol{b}_{i-1} + \mu \widehat{\boldsymbol{x}}^*_{i-1} Q_{PTQ}\{e(i)\} \quad (M_b \times 1) \tag{16}$$

The introduction of the PTQ converts the multiplication operation between the error signal $e(i)$ and the regressor vectors $\boldsymbol{r}$ and $\widehat{\boldsymbol{x}}$ into simple shift operations. If μ is chosen as power-of-two, then the recursions can be digitally implemented without any multiplications. In the LMS standard format, we can write this recursion as (utilizing the transformation (7-9))

$$e(i) = d(i) - \boldsymbol{u}_i \boldsymbol{w}_{i-1} \tag{17}$$
$$\boldsymbol{w}_i = \boldsymbol{w}_{i-1} + \mu \boldsymbol{u}^*_i Q_{PTQ}\{e(i)\} \quad (M \times 1). \tag{18}$$

4. PERFORMANCE ANALYSIS OF THE PTQ-DFE

Consider a power-of-two quantizer in the form shown in Fig.2. It was proven in [6, 7] that the PTQ can be represented by the following input-output expression

$$\boxed{y = Kx} \tag{19}$$

where K is a random scalar gain defined as

$$K \triangleq 2^{e_Q}. \tag{20}$$

and e_Q is the quantization noise of the quantizer inside the PTQ. This noise is assumed uniform within the interval $[-\Delta, \Delta]$, where $\Delta \left(= \frac{1}{2^{B+1}}\right)$ is half the quantizer's step-size. Consequently, the first and second moments of the gain K can be readily computed as

$$E_K \triangleq E\{K\} = \frac{1}{2\Delta \ln(2)} \left(2^{\Delta} - 2^{-\Delta}\right) \tag{21}$$
$$E_{K^2} \triangleq E\{K^2\} = \frac{1}{4\Delta \ln(2)} \left(2^{2\Delta} - 2^{-2\Delta}\right) \tag{22}$$

Using this result, we can rewrite the PTQ-DFE update equation (18) as

$$\boldsymbol{w}_i = \boldsymbol{w}_{i-1} + \mu \boldsymbol{u}^*_i K(i) \left[\boldsymbol{u}_i \boldsymbol{w}^{opt} + v(i) - \boldsymbol{u}_i \boldsymbol{w}_{i-1}\right]. \tag{23}$$

where $\boldsymbol{w}^{opt}$ is an $(M \times 1)$ optimal weights vector and $v(i)$ is the measurement noise with variance σ_v^2. Subtracting $\boldsymbol{w}^{opt}$ from both sides leads to

$$\widetilde{\boldsymbol{w}}_i = [I - \mu \boldsymbol{u}^*_i \boldsymbol{u}_i K(i)]\, \widetilde{\boldsymbol{w}}_{i-1} + \mu \boldsymbol{u}^*_i K(i) v(i) \tag{24}$$

where the weight error vector $\widetilde{\boldsymbol{w}}_i$ is defined as

$$\widetilde{\boldsymbol{w}}_i \triangleq \boldsymbol{w}_i - \boldsymbol{w}^{opt}. \qquad (25)$$

This structure is similar to that of traditional LMS except with the introduction of the gain K. Therefore, we can borrow the same methodologies of the LMS analysis to investigate the convergence of the proposed scheme (see, for example, [8]). In the analysis, the following assumptions are used

1. The quantization noise $e_Q(i)$ is assumed independent of the quantizer input $e(i)$. In this case, the random gain $K(i)$ can also be assumed independent of $e(i)$. This assumption is reasonable when B is sufficiently large (say $B \geq 4$).
2. The additive noise $v(i)$ is zero-mean gaussian noise.
3. The input $\boldsymbol{u}_i$ is stationary with autocorrelation matrix R.

Due to the limitation on the paper size, we state directly the following results

Lemma 1 (Convergence in the mean) The recursion (15) and (16) converges asymptotically in the mean if

$$0 < \mu < \frac{2}{\lambda_{max} E_K}. \qquad (26)$$

where λ_{max} is the maximum eigenvalue of R.

Lemma 2 (Convergence in the mean square) Consider the PTQ-LMS recursion (15) and (16). If the step-size of the recursion is chosen such that

$$0 < \mu < \frac{2E_K}{\lambda_{max} E_{K^2}}$$

then the recursion converges in the mean square to the quantity

$$C_\infty = \frac{\alpha \sigma_v^2}{1-c} \left[2I - \alpha R\right]^{-1} \qquad (27)$$

where $c = \sum_{j=1}^{M} \frac{\alpha \lambda_j}{2 - \alpha \lambda_j}$, and $\{\lambda_j\}$ are the eigne values of R, and

$$\alpha \triangleq \mu \frac{E_{K^2}}{E_K}$$

. Furthermore, the steady state MSE is given by

$$MSE_{ss} = \sigma_v^2 + \frac{c\sigma_v^2}{1-c} \qquad (28)$$

while the expression for the excess MSE is

$$MSE_{ex} = \frac{c\sigma_v^2}{1-c} \qquad (29)$$

and therefore, the MSE misadjustment is given by

$$\Omega = \frac{c}{1-c}. \qquad (30)$$

5. SIMULATION

In this section, the performance of the proposed PTQ-DFE is investigated via simulation and results are compared with the analytical findings. In simulation, a 4-QAM complex signal is transmitted through a 4-tap channel with impulse response

$$h = [0.5 \quad 1.2 \quad 1.5 \quad -1] \qquad (31)$$

The additive noise is complex Gaussian noise with zero mean. The number of taps M_f and M_b for the DFE are chosen as 20 and 2 respectively. The equalizer's delay q is set to 10 samples and the adaptation step size μ is chosen as 2^{-10}. In the training mode, 200 samples are used while 10000 samples are used in the decision-directed mode.

Fig. 3 shows typical learning curves for the proposed PTQ-DFE implementing the three PTQ types given by (12), (13) and (14), averaged over 100 runs and with noise variance of -30dB. The conventional (non-quantized) LMS-based DFE is also included for comparison purpose. We notice that the four curves are almost coinciding with each other indicating a similar learning performance between the PTQ-DFE systems and the conventional DFE. In Fig. 4, typical constellation setup obtained using the proposed PTQ-DFE with SNR=10dB is shown, including source, transmitted, received, and equalized signals.

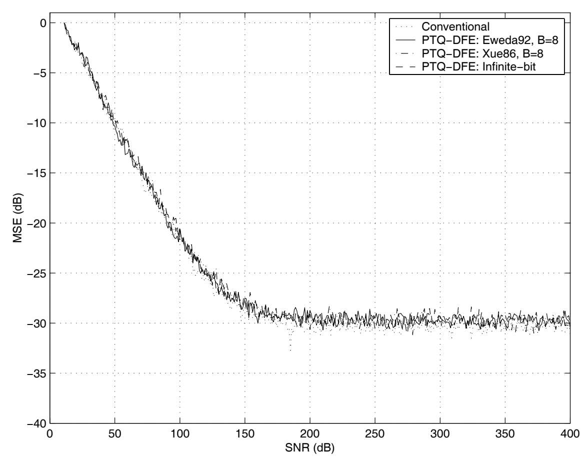

Fig. 3. *Learning curve performance for the PTQ-DFE techniques compared to that of conventional DFE (results are averaged over 100 runs).*

The bit error rate (BER) is also measured under varying SNR for both proposed PTQ-DFE and conventional DFE. Results are shown in Fig. 5. We notice again the similarity between the BER performance of the PTQ-DFE and that of the conventional DFE. In Fig. 6, the simulated steady state MSE is plotted as a function of SNR and results are compared with the analytical findings from (28). In this case, both simulated and analytical results are matching.

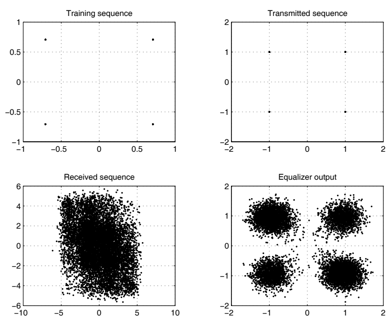

Fig. 4. *Typical constellations of a 4-QAM signal going through a PTQ-DFE with SNR=10dB.*

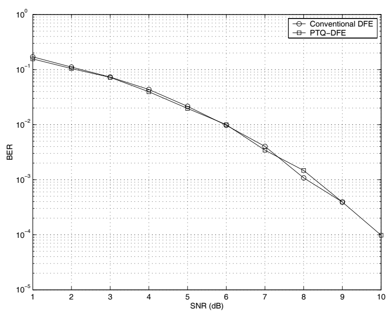

Fig. 5. *The BER versus SNR performance for the PTQ-DFE compared to the conventional DFE.*

6. COMPUTATION COMPLEXITY

As mentioned earlier, the main advantage of the PTQ concept is that it completely eliminates the need for real multiplications in digital implementation of the DFE update equations (15) and (16) by quantizing the error signal $e(i)$ into a power-of-two value. The number of real multiplications that will be saved per iteration is $M_f + M_b$ for real and $4M_f + 4M_b$ for complex signaling respectively. For example, let $M_f = 20$ and $M_b = 2$ then the number of real multiplications saved for 10^4 iterations will be 22×10^4 and 88×10^4 for real and complex signaling respectively.

7. CONCLUSION

In this paper, an efficient structure for the decision feedback equalizer is proposed. The complexity reduction is attained via the utilization of power-of-two quantization techniques. These techniques completely eliminate the need for real multiplication in the adaptation algorithm of the equalizer. It is shown through both analysis and simulation that this reduction in complexity is attained with almost no degradation in the equalization performance. The convergence in both mean and mean square of the proposed algorithm was discussed and expressions for the mean square error were derived.

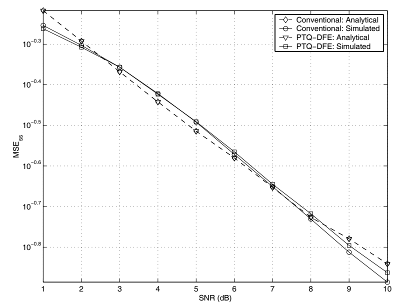

Fig. 6. *Steady state mean square error for both PTQ-DFE and conventional DFE.*

8. ACKNOWLEDGMENT

The author would like to acknowledge King Fahad University of Petroleum and Minerals and the British Council for support of this work. Moreover, the help of Prof. Bayan Sharif in reviewing this work is highly valued.

9. REFERENCES

[1] J. Treichler, C. Johnson, Jr., and M. Larimore, "Theory and Design of Adaptive Filters," Printice Hall, Inc., NJ, 2001.

[2] P. Xue and B. Liu, "Adaptive equalizer using finite-bit power-of-two quantizer," *IEEE Trans. Acoustics, Speech, and Signal Processing,* Vol. 34, No. 6, pp. 1603-11, Dec 1986.

[3] E. Eweda, "Convergence analysis and design of an adaptive filter with finite-bit power-of-two quantized error," *IEEE Trans. Circuits and Systems II: Analog and Digital Signal Processing,* Vol. 39, No. 2, pp. 113-115, Feb. 1992.

[4] E. Eweda, "Comment on: Convergence analysis and design of an adaptive filter with finite-bit power-of-two quantized error," *IEEE Trans. Circuits and Systems II: Analog and Digital Signal Processing,* Vol. 42, No. 7, pp. 498-500, Jul. 1995.

[5] D. L. Duttweiler, "Adaptive filter performance with nonlinearities in the correlation multiplier," *IEEE Trans. Acoust., Speech, Signal Proc.,* Vol. ASSP-30, pp. 578-586, Aug. 1982.

[6] M. A. Aldajani and A. H. Sayed, "Stability and performance analysis of an adaptive sigma-delta modulator," *IEEE Trans. on Circuits and Systems II: Analog and Digital Signal Processing,* vol. 48, no. 3, pp. 233-244, Mar. 2001.

[7] M. A. Aldajani and A. H. Sayed, "A stable adaptive structure for delta modulation with improved performance," *Proc. ICASSP,* Salt Lake City, Utah, Vol. 4, pp. 2621–24, May 2001.

[8] A. H. Sayed, "Fundamentals of Adaptive Filtering," Wiley, NY, 2003.