UNBIASED MMSE DECISION-FEEDBACK EQUALIZATION FOR PACKET TRANSMISSION†

Dirk T.M. Slock and Elisabeth de Carvalho
Institut EURECOM, 2229 route des Crêtes, B.P. 193,
06904 Sophia Antipolis Cedex, FRANCE
Tel: +33 93002606 Fax: +33 93002627
email:{slock,carvalho}@eurecom.fr

ABSTRACT

We derive expressions for the different linear and decision feedback equalizers in burst mode in the multichannel case. Among them we derive the class of unbiased minimum mean squared error equalizers. Optimal burst mode filters are found to be time-varying. Performance comparisons between these equalizers are done in terms of SNR and probability of error: these measures depend on the position in the burst. We study furthermore the performance when symbols are known or not at the edges of the burst and compare it to the continuous processing level. Finally we show that (time-invariant) continuous processing applied to burst mode can be organized to give sufficiently good performance, so that optimal (time-varying) burst processing implementation can be avoided.

1 INTRODUCTION

We consider here a FIR multichannel model. The multiple FIR channels are due to oversampling of a single received signal and/or the availability of multiple received signals from an array of antennas (in the context of mobile digital communications). To further develop the case of oversampling, consider linear digital modulation over a linear channel with additive noise so that the cyclostationary received signal can be written as

$$\mathrm{y}(t) \;=\; \sum_k \mathrm{h}(t-kT)a(k) \;+\; \mathrm{v}(t) \qquad (1)$$

where the $a(k)$ are the transmitted symbols, T is the symbol period and $\mathrm{h}(t)$ is the channel impulse response. The channel is assumed to be FIR with duration NT (approximately). If the received signal is oversampled at the rate $\frac{m}{T}$ (or if m different received signals are captured by m sensors every T seconds, or a combination of both), the discrete input-output relationship can be written as:

$$\boldsymbol{y}(k) \;=\; \sum_{i=0}^{N-1} \boldsymbol{h}(i)a(k-i)+\boldsymbol{v}(k) \;=\; \boldsymbol{H}A_N(k)+\boldsymbol{v}_k$$

$$\boldsymbol{y}(k)=\begin{bmatrix} y_1(k)\\ \vdots \\ y_m(k)\end{bmatrix},\boldsymbol{v}(k)=\begin{bmatrix} v_1(k)\\ \vdots \\ v_m(k)\end{bmatrix},\boldsymbol{h}(k)=\begin{bmatrix} h_1(k)\\ \vdots \\ h_m(k)\end{bmatrix}$$

$$\boldsymbol{H}{=}[\boldsymbol{h}(N{-}1)\cdots\boldsymbol{h}(0)],\, A_N(k){=}\left[a^H(k{-}N{+}1)\cdots a^H(k)\right]^H$$

where the subscript i denotes the i^{th} channel and superscript ^{H} denotes Hermitian transpose. In the case of oversampling, $y_i(k), i=1,\ldots,m$ represent the m phases of the polyphase representation of the oversampled signal: $y_i(k)=\mathrm{y}(t_0+(k+\frac{i}{m})T)$. In this representation, we get a discrete-time circuit in which the sampling rate is the symbol rate. Its output is a vector signal corresponding to a SIMO (Single Input Multiple Output) or vector channel consisting of m SISO discrete-time channels where m is the sum of the oversampling factors used for the possibly multiple antenna signals. Let $\mathbf{H}(z)=\sum_{i=0}^{N-1}\boldsymbol{h}(i)z^{-i}=[\mathrm{H}_1^H(z)\cdots\mathrm{H}_m^H(z)]^H$ be the SIMO channel transfer function. Consider additive independent white Gaussian noise $\boldsymbol{v}_k$ with $r_{\boldsymbol{vv}}(k{-}i)=\mathrm{E}\,\boldsymbol{v}(k)\boldsymbol{v}^H(i)=\sigma_v^2 I_m\,\delta_{ki}$. Assume we receive M samples:

$$\boldsymbol{Y}_M(k) \;=\; \mathcal{T}_M(\boldsymbol{H})\,A_{M+N-1}(k)+\boldsymbol{V}_M(k) \qquad (2)$$

where $\boldsymbol{Y}_M(k){=}[\boldsymbol{y}^H(k{-}M{+}1)\cdots\boldsymbol{y}^H(k)]^H$ and similarly for $\boldsymbol{V}_M(k)$, and $\mathcal{T}_M(\boldsymbol{H})$ is a block Toepliz matrix with M block rows and $[\boldsymbol{H}\;\;0_{m\times(M-1)}]$ as first block row.

2 BURST TRANSMISSION

We consider a transmission by burst in which detection is done burst by burst. We suppose that the channel is time-invariant during the transmission of a burst.

In the input burst, denoted B, some symbols are known: $n1$ at the beginning, grouped in the vector A_1, and $n2$ at the end, grouped in the vector A_2. The total length of the burst is $M{+}n1{+}n2$; we want to detect the M central unknown symbols, grouped in the vector A. For that purpose, we will consider as observation data, Y_{obs}, the channel outputs that contain only symbols of burst B (the symbols to be detected or the known symbols of the burst), and not outputs containing symbols of neighbouring bursts: see fig. 1. The input-ouput relationship (2) between the observation data $\boldsymbol{Y}_{obs}$ and B

[0]The work of Elisabeth de Carvalho was supported by Laboratoires d'Electronique PHILIPS under contract LEP 95FAR008.

is written in simplified notation as $\boldsymbol{Y}_{obs} = \mathcal{T}_B B + \boldsymbol{V}$. More data could be considered also, but this possibiblity will not be explored in this paper.

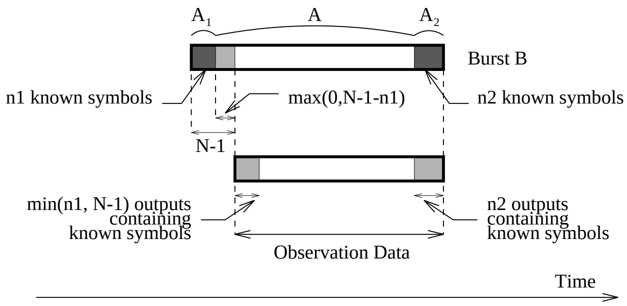

Figure 1: Burst Transmission

3 BURST-MODE EQUALIZERS

In this section, we derive the expressions for the different equalizers in burst mode. Some related work can be found in [1]. Linear Equalizers (LE) and Decision Feedback Equalizers (DFE) are considered for the Minimum Mean Squared Error (MMSE), the MMSE Zero-Forcing (MMSE ZF), the Unbiased MMSE (UMMSE) criteria.

In the following, we consider the decomposition $\boldsymbol{Y}_{obs} = \mathcal{T}_B B + \boldsymbol{V} = \mathcal{T}_1 A_1 + \mathcal{T} A + \mathcal{T}_2 A_2 + \boldsymbol{V}$, where $\mathcal{T}_i A_i$ represents the contribution of the symbols in A_i. We will denote: $\boldsymbol{Y} = \mathcal{T} A + V = \boldsymbol{Y}_{obs} - \mathcal{T}_1 A_1 - \mathcal{T}_2 A_2$, the processing data.

The different equalizers are linear estimators of the input symbols. Linear equalizers give linear estimates given $\boldsymbol{Y}_{obs}$, A_1 and A_2. DFEs give linear estimates given $\boldsymbol{Y}_{obs}$, A_1 and A_2, as well as the decisions on the previous input symbols. We shall assume those previous decisions to be error-free.

3.1 Linear Equalizers

3.1.1 The MMSE Linear Equalizer

Given the observations $\boldsymbol{Y}'^H = \left[\begin{array}{ccc} \boldsymbol{Y}^H_{obs} & A_1^H & A_2^H \end{array}\right]^H$, the linear MMSE estimator of A is:

$$\widehat{A} = R_{A\boldsymbol{Y}'} R^{-1}_{\boldsymbol{Y}'\boldsymbol{Y}'} \boldsymbol{Y}' = R_{A\boldsymbol{Y}} R^{-1}_{\boldsymbol{Y}\boldsymbol{Y}} \boldsymbol{Y} \tag{3}$$

The last equality the proof of which is omitted shows that linear estimation in terms of $\boldsymbol{Y}'$ is the same as in terms of $\boldsymbol{Y}$: the optimal processing can be seen as eliminating first the contributions of known symbols from the observation data $\boldsymbol{Y}_{obs}$ to get $\boldsymbol{Y}$ and then applying the MMSE equalizer determined on the basis of $\boldsymbol{Y}$. For the other equalizers, the previous result is also true but will not be restated. From equation (3):

$$\widehat{A} = \sigma_a^2 \mathcal{T}^H (\sigma_a^2 \mathcal{T}\mathcal{T}^H + \sigma_v^2 I)^{-1} \boldsymbol{Y} = \left(\mathcal{T}^H\mathcal{T} + \frac{\sigma_v^2}{\sigma_a^2} I\right)^{-1} \mathcal{T}^H \boldsymbol{Y} \tag{4}$$

The last equality is obtained via the matrix inversion lemma. We will denote $R = \mathcal{T}^H\mathcal{T} + \frac{\sigma_v^2}{\sigma_a^2} I$.

In the continuous processing case, the MMSE equalizer gives the output:

$$\hat{a}_k = \left(H^+(q)H(q) + \frac{\sigma_v^2}{\sigma_a^2} I\right)^{-1} H^+(q) y_k \tag{5}$$

where $H^+(z) = H^H(1/z^*)$ and $q^{-1}y_k = y_{k-1}$. By analogy with the continuous processing case, we can find interpretations for the expression (4) in filtering terms:

- $\mathcal{T}^H$ represents the multichannel matched filter. It is toeplitz, banded and upper triangular, which implies that the filtering is time-invariant, FIR and anticausal.
- R^{-1} is the FIR denominator of an IIR filter, it is non-causal.

The filters are in general time-varying.

We define the MSE of the ith symbol as:

$$MSE_i = \left(E(\widehat{A} - A)(\widehat{A} - A)^H\right)_{ii} \tag{6}$$

and the Signal to Noise Ratio (SNR) of the ith symbol:

$$SNR_i = \frac{\sigma_a^2}{MSE_i} \tag{7}$$

For the burst mode MMSE LE:

$$SNR_i = \frac{\sigma_a^2}{\sigma_v^2 (R^{-1})_{ii}} \tag{8}$$

It has to be noted that the SNR depends on the position of the symbol in the burst. This remark will also valid for the other equalizers.

3.1.2 The MMSE-ZF Linear Equalizer

The MMSE ZF LE corresponds to the Best Linear Unbiased Estimator (BLUE). Given the linear model: $\boldsymbol{Y} = \mathcal{T} A + \boldsymbol{V}$, the BLUE is given by:

$$\begin{aligned} \widehat{A}_{BLUE} &= (\mathcal{T}^H R_{VV}^{-1} \mathcal{T})^{-1} \mathcal{T}^H R_{VV}^{-1} \boldsymbol{Y} \\ &= (\mathcal{T}^H R_{\boldsymbol{Y}\boldsymbol{Y}}^{-1} \mathcal{T})^{-1} \mathcal{T}^H R_{\boldsymbol{Y}\boldsymbol{Y}}^{-1} \boldsymbol{Y} \end{aligned} \tag{9}$$

So the MMSE ZF is:

$$\widehat{A} = (\mathcal{T}^H\mathcal{T})^{-1}\mathcal{T}^H\boldsymbol{Y} \tag{10}$$

The output burst mode SNR is:

$$SNR_i = \frac{\sigma_a^2}{\sigma_v^2((\mathcal{T}^H\mathcal{T})^{-1})_{ii}} \tag{11}$$

In the continuous processing case, the MMSE ZF LE output is:

$$\hat{a}_k = \left(H^+(q)H(q)\right)^{-1} H^+(q) y_k \tag{12}$$

3.1.3 The Unbiased MMSE Linear Equalizer

A MMSE equalizer produces a biased estimate of the symbol a_k. This bias increases the probability of error [2]. The Unbiased MMSE LE is the element-wise BLUE and is given by:

$$\left(I - \frac{\sigma_v^2}{\sigma_a^2}\mathrm{diag}(R^{-1})\right)^{-1} R^{-1}\mathcal{T}^H \tag{13}$$

It is simply a scaled version of the MMSE LE.

The SNR of the UMMSE LE is related to the SNR of the MMSE LE:

$$SNR_i(\text{UMMSE LE}) = SNR_i(\text{MMSE LE}) - 1 \quad (14)$$

The advantage of the unbiased LE is that its probability of error is the lowest of all the (linear) equalizers. Indeed for any unbiased equalizer, the probability of error is an increasing function of MSE (with a gaussian approximation for residual ISI).

In the continuous processing case, the output of Unbiased MMSE LE has for expression:

$$\left(1 - \frac{\sigma_v^2}{\sigma_a^2} \oint \frac{dz}{z} \left(H^+(z)H(z) + \frac{\sigma_a^2}{\sigma_v^2}\right)\right)^{-1} \hat{a}_{k,MMSE-LE} \quad (15)$$

3.2 Decision Feedback Equalizers

We will not derive the expression of the DFEs but will expose a way to get their expression from the LEs.

3.2.1 MMSE DFE and MMSE ZF DFE

For the MMSE LE, we consider the UDL factorization of $R = L^H DL$. For the MMSE ZF LE, we consider the UDL factorization of $\mathcal{T}^H\mathcal{T} = L^H DL$. The ouput of these two equalizers can then be written as:

$$\widehat{A} = L^{-1}D^{-1}L^{-H}\mathcal{T}^H\mathbf{Y} = D^{-1}L^{-H}\mathcal{T}^H\mathbf{Y} - (L-I)\widehat{A} \quad (16)$$

The DFE operation consists in taking $(L - I)\text{dec}(\widehat{A})$ instead of $(L-I)\widehat{A}$, where dec is the decision operation.

The symbol estimator is then:

$$\widehat{A} = D^{-1}L^{-H}\mathcal{T}^H\mathbf{Y} - (L-I)\text{dec}(\widehat{A}) \quad (17)$$

The forward filter consists in the cascade of the multichannel matched filter and an anticausal filter $D^{-1}L^{-H}$. $L - I$ is a strictly causal filter, so that the feedback operation involves only past decisions.

If we suppose past decisions to be good, the SNR is:

$$SNR_i = \frac{\sigma_a^2}{\sigma_v^2(D^{-1})_{ii}} \quad (18)$$

In the continuous processing case:

$$\hat{a}_k = \frac{H^+(q)}{dG^+(q)}y_k - (G(q)-1)\text{dec}(\hat{a}_k) \quad (19)$$

where $H^+(q)H(q) + \frac{\sigma_v^2}{\sigma_a^2} = G^+(q)dG(q)$, $G(q)$ is causal and $G(\infty) = 1$.

3.2.2 Unbiased MMSE DFE

The output of the Unbiased MMSE DFE is:

$$\widehat{A} = (I - \frac{\sigma_v^2}{\sigma_a^2}D^{-1})^{-1}\widehat{A}_{MMSE-DFE} \quad (20)$$

The burst output SNR is decreased by 1 with respect to the MMSE DFE.

The continuous processing equalizer output is:

$$\left(1 - \frac{\sigma_v^2}{\sigma_a^2} exp\left(-\oint \frac{dz}{z} ln(H^+(z)H(z) + \frac{\sigma_a^2}{\sigma_v^2})\right)\right)^{-1} \hat{a}_{k,MMSE-DFE} \quad (21)$$

4 PERFORMANCE COMPARISONS

In this section, we discuss the performance of the equalizers in terms of SNR and probability of error.

4.1 Case of no known symbols: n1=n2=0

In fig. 2 (left), the SNR curves are drawn for a channel $\boldsymbol{H}$ of length 7 with 3 subchannels and which coefficients where randomly chosen. The SNR per channel is 10dB.

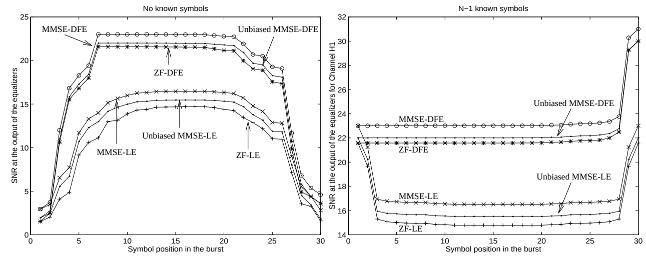

Figure 2: SNRs at the output of the different equalizers when no symbols are known (left) and when N-1 symbols at each end of the burst are known (right)

We notice that degradations appear at the ends of the burst. The middle symbols appear in N outputs. When no symbols are known, the first and last unknown symbols of the burst appear in strictly less than N outputs, so that there is less information about those symbols in the observations.

The SNR in the middle of the burst converges to the continuous processing level as the burst length increases: burst processing is inferior to continuous processing in this case.

4.2 Case of N-1 known symbols at each end

We suppose now that n1=n2=N-1. We drawn the SNR curves in fig. 2 (right). This time, burst processing performs better than continuous processing. The middle observations contains N symbols. After eliminating the contributions of the known symbols the outputs at the edges contain strictly less than N symbols, so that there is more information on those symbols, which are then better estimated.

4.3 Equalizers Comparisons

4.3.1 In terms of SNR

- Burst DFEs have a better average performance than the corresponding Burst LEs.
- MMSE Burst DFEs have a better average performance than ZF Burst DFEs.
- MMSE Burst LEs have a better average performance than ZF Burst LEs.

4.3.2 In terms of probabilities of error

For unbiased equalizers, a higher SNR implies a lower probability of error: MMSE ZF equalizers will then have a higher probability of error than the corresponding Unbiased MMSE equalizers. However, it is not obvious to rank the MMSE equalizers w.r.t. the ZF equalizers because they are biased. In fact people would tend to believe that a MMSE equalizer performs better than the corresponding MMSE-ZF equalizer. In the case of constant modulus modulations, MMSE equalizers have the same performance as the corresponding unbiased MMSE equalizers and so a higher performance than MMSE ZF equalizers. For non constant-modulus constellations, the bias in MMSE equalizers may have a stronger effect than its higher SNR compared to MMSE-ZF equalizers. This is all the more true as the difference in SNRs between the different equalizers tends to be lower as subchannels are added.

In fig. 3, the probabilities of error of the central symbols are plotted for channel $\boldsymbol{H}$ for the different DFEs. The input symbols belong to a 4-PAM constellation. We notice that the MMSE equalizer has poorer performance than the MMSE ZF equalizer for most symbols. Simulations with other channels gave the same result.

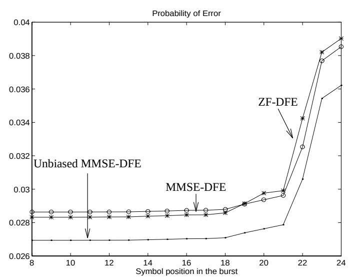

Figure 3: Probability of Error for the ZF-DFE, the MMSE-DFE and the Unbiased MMSE-DFE

5 APPLYING CONTINUOUS PROCESSING EQUALIZERS TO THE BURST CASE

As already mentioned, burst processing involves time-varying filters. We may wonder if it is worth implementing these time-varying filters, and if simply applying the filters corresponding to continuous processing in burst mode could give acceptable performance.

For that purpose we will consider the case of N-1 known symbols at each end of the input burst. We will show that the continuous processing filters also give better SNR at the ends of the burst than at the middle and always give strictly better SNR than in the continuous processing case.

For the LEs, the contributions of known symbols is removed from the observation data. For the DFEs, the initialization is done by putting the N-1 leading known bits in the memory of the feedback filter. Only the trailing known symbols are removed from the processing data. In both cases, we put the channel outputs before and after the data to be processed equal to zero. The only difference with the continuous processing case is that we have a finite input symbol sequence, but also a finite noise sequence.

For the LEs, the different reasonings will be held for zero delay non-causal continuous processing filters. For the DFEs, the forward filter is assumed to be anticausal (zero delay) the feedback filter is causal and FIR (of the same length as the channel). As the channel output is zero outside the time interval of the processing data, these filters will involve only a finite number of data.

In the MMSE ZF case, the MSE contains only the noise contributions. Since the noise is only finite length, the MSE is smaller at the edges. The MSE of MMSE (unbiased or not) equalizers outputs contains residual ISI also. This variance gets also reduced as the input sequence becomes finite length.

5.1 MSE Calculations

The outputs of the different linear equalizers based on the continuous processing filters may be written as:

$$\widehat{A} = FY \tag{22}$$

where F is a structured matrix containing the coefficients of the continuous processing filter.

In general:

$$MSE_i = (\sigma_a^2(F\mathcal{T} - I)(F\mathcal{T} - I)^H + \sigma_v^2 FF^H)_{ii} \tag{23}$$

where $F\mathcal{T} = I$ in the ZF case.

The outputs of the different DFEs be may written as:

$$\widehat{A} = FY - (B - I')A' \tag{24}$$

A' contains A and the leading symbols, $I' = \left[\begin{array}{cc} I & 0 \end{array}\right]$, F contains the coefficients of the continuous processing forward filter. B the coefficients of the continuous processing feedback filter.

In general:

$$MSE_i = (\sigma_a^2(F\mathcal{T} - B)(F\mathcal{T} - B)^H + \sigma_v^2 FF^H)_{ii} \tag{25}$$

where $F\mathcal{T} = B$ in the ZF case.

In fig. 4, we compare performances.

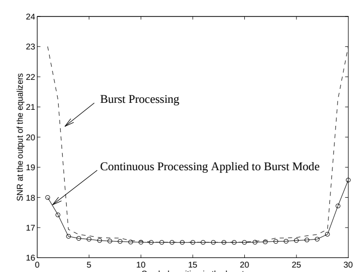

Figure 4: SNR Curves in the case of optimal burst processing compared to continuous processing applied to burst mode for the MMSE Linear Equalizer

References

[1] G. Kawas Kaleh. "Channel Equalization for Block Transmission Systems". *IEEE Journal on Selected Areas in Communications*, 13(1):110–121, Jan. 1995.

[2] 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. Com.*, 43(10):2582–2594, Oct. 1995.