Proceedings of the 39th IEEE
Conference on Decision and Control
Sydney, Australia • December, 2000

On Achieving Reduced Error Propagation Sensitivity in DFE Design via Convex Optimization

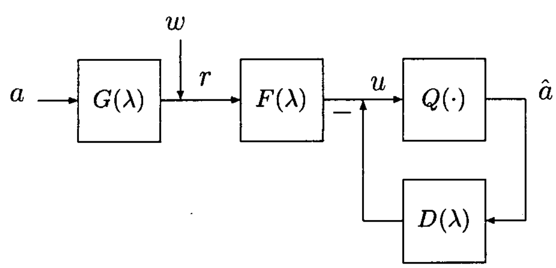
Robert L. Kosut[1] Wonzoo Chung and C. Richard Johnson, Jr.[2] Stephen P. Boyd[3]

1 Introduction

Decision Feedback Equalization (DFE) is expected in digital TV receivers and other high error rate environments, *e.g.*, [1]. Error propagation usually occurs in infrequent bursts, [4, App.10-A],[2, sec.7.5.4] It is argued here and in [1] that the minimum mean-square-error (MMSE) adaptation mechanism in the presence of error propagation will find a better answer than the solution computed in the absence of decision errors. This paper attempts to formalize this benefit during the design phase, by considering other (convex) performance measures than MSE assuming perfect decisions. After all, any such modified objective is just a proxy for determining the optimal error rate. As discussed in [3], error propagation is "enhanced" by large gains in the decision portion of the DFE portion. We consider a method to penalize these gains, but not in the unconstrained (perfect decision) MSE sense.

2 Standard DFE System

Consider the DFE communications system:



The system eqations are:

$$\begin{aligned} \hat{a}_t &= Q(u_t) && (1)\\ u_t &= F(\lambda)r_t - D(\lambda)\hat{a}_t && (2)\\ r_t &= w_t + G(\lambda)a_t && (3) \end{aligned}$$

where λ denotes the backward shift operator, *i.e.*, for integer times t, $\lambda x_t = x_{t-1}$, $G(\lambda), F(\lambda), D(\lambda)$ are linear-time-invariant systems, and $Q(\cdot)$ is the quantizer.

[1]SC Solutions, Inc., Santa Clara, CA, kosut@scsolutions.com
[2]Cornell University, Ithaca, NY, wonzoo@ee.cornell.edu, johnson@ee.cornell.edu
[3]Stanford University, Stanford, CA, boyd@isl.stanford.edu

The source sequence, a, consists of elements in the ℓ-ary alphabet (ℓ is even),

$$\mathbf{A} = \{\pm 1, \pm 3, \ldots, \pm(\ell - 1)\} \quad (4)$$

The received signal sequence, r, is a version of the source sequence, a, distorted by the channel dynamics, $G(\lambda)$, and corrupted by an additive noise sequence, w. The sequence $\hat{a}$ is an estimate of the source sequence and is the output of the quantizer,

$$Q(u_t) = \arg\min_{\alpha \in \mathbf{A}} |\alpha - u_t| \quad (5)$$

In the binary case ($\ell = 2$), $\mathbf{A} = \{\pm 1\}$ and the quantizer reduces to the sign function. The filters $F(\lambda)$ and $D(\lambda)$ are FIR, *i.e.*,

$$\begin{aligned} F(\lambda) &= f_0 + f_1\lambda + \cdots + f_n\lambda^n \\ D(\lambda) &= d_1\lambda + \cdots + d_m\lambda^m \end{aligned} \quad (6)$$

The channel, $G(\lambda)$, is stable and possibly IIR, that is,

$$G(\lambda) = \sum_{i=0}^{\infty} g_i \lambda^i \quad (7)$$

with $g_i \to 0$ as $i \to \infty$, exponentially. The sequences $f = \{f_0, \ldots, f_n\}$, $d = \{d_1, \ldots, d_m\}$, and $g = \{g_0, g_1, \ldots\}$ are the impulse responses of $F(\lambda), D(\lambda), G(\lambda)$, respectively.

The DFE system can also be expressed in terms of the impulse response sequences g, f, d, that is,

$$\begin{aligned} \hat{a}_t &= Q(u_t) && (8)\\ u_t &= (f * r)_t - (d * \hat{a})_t && (9)\\ r_t &= w_t + (g * a)_t && (10) \end{aligned}$$

where $*$ denotes convolution. The DFE design variables are the FIR filter coefficients $f = \{f_0, \ldots, f_n\}$ and $d = \{d_1, \ldots, d_m\}$.

3 Symbol-Error-Rate (SER)

There are two basic DFE design problems: (1) design with known channel, and (2) design with unknown channel, *i.e.*, adaptation. Here we consider only the former problem.

0-7803-6638-7/00$10.00 © 2000 IEEE

To specify the design issues more precisely, define the symbol error sequence,

$$e_t = \hat{a}_t - a_{t-\delta} \tag{11}$$

for some integer $\delta \geq 0$. The symbol error is often referred to as the *hard error.* The design goal is to select the filter impulse response coefficients f, d, and the delay δ, to minimize the *symbol error rate*,

$$\rho = \mathbf{P}\{ e_t \neq 0 \} = 1 - \mathbf{P}\{ e_t = 0 \} \tag{12}$$

where the probability measure, $\mathbf{P}\{\cdot\}$, is over the distributions of the source sequence a and the noise sequence w.

We also define the *soft error*,

$$\varepsilon_t = u_t - h_\delta\, a_{t-\delta} \tag{13}$$

where h_δ is the δ-th impulse response coefficient of

$$H(\lambda) = F(\lambda)G(\lambda) = \sum_{i=0}^{\infty} h_i\, \lambda^i \tag{14}$$

After some algebra, the soft error can be expressed as,

$$\begin{aligned} \varepsilon_t &= F(\lambda)w_t + \tilde{H}(\lambda)a_t - D(\lambda)e_t \\ \tilde{H}(\lambda) &= F(\lambda)G(\lambda) - \lambda^\delta(D(\lambda) + h_\delta) \end{aligned} \tag{15}$$

Observe that $\tilde{H}(\lambda)$ is FIR only if $G(\lambda)$ is FIR. Also, the term $\tilde{H}(\lambda)a_t$ does not depend on $a_{t-\delta}$ because $h_\delta\, a_{t-\delta}$ is subtracted from $F(\lambda)G(\lambda)a_t$. The soft error can be written in terms of impulse responses as,

$$\begin{aligned} \varepsilon_t &= (f * w)_t + (\tilde{h} * a)_t - (d * e)_t \\ \tilde{h} &= g * f - \lambda^\delta * (h_\delta + d) \end{aligned} \tag{16}$$

where $\tilde{h}$ is the impulse response of $\tilde{H}(\lambda)$. The quantizer output can then be expressed as,

$$\hat{a}_t = Q(h_\delta a_{t-\delta} + \varepsilon_t) \tag{17}$$

4 Design Heuristic with Known Channel

The obvious design goal is to minimize the symbol error rate,

$$\min_{f,d,\delta} \rho \tag{18}$$

Unfortunately there is no computationally feasible solution to this problem for the DFE configuration. Here we present heuristic solutions which modify the classical minimum-mean-square-error (MMSE) method.

Binary alphabet

In the binary case, $\mathbf{A} = \{\pm 1\}$, and hence, $Q(x) = \text{sgn}(x)$. If $h_\delta > 0$, then,

$$\hat{a}_t = \text{sgn}(h_\delta a_{t-\delta} + \varepsilon_t) = \text{sgn}(a_{t-\delta} + \varepsilon'_t) \tag{19}$$

where prime denotes division by h_δ, *i.e.*, $(\cdot)' = (\cdot)/h_\delta$. Thus,

$$\begin{aligned} \varepsilon'_t &= \varepsilon_t/h_\delta = (f' * w)_t + (\tilde{h}' * a)_t - (d' * e)_t \\ \tilde{h}' &= g * f' - \lambda^\delta * (1 + d') \end{aligned} \tag{20}$$

Since only the sign of h_δ matters in the binary case,[1] it follows that the design goal is to select (f, d, δ) to make ε' small in some sense.

When he symbol-error-rate, ρ, is typically *very* small, there can be very few errors over the tap length of the decision filter $D(\lambda)$. For example, suppose there is only one non-zero error in any time window $\{t-1, \ldots, t-m\}$. Then,

$$\begin{aligned} |\varepsilon'_t| &= |(f' * w)_t + (\tilde{h}' * a)_t - (d' * e)_t| \\ &\leq |(f' * w)_t + (\tilde{h}' * a)_t| + |(d' * e)_t| \\ &\leq |(f' * w)_t + (\tilde{h}' * a)_t| + 2\|d'\|_\infty \end{aligned} \tag{21}$$

The classical MMSE design approach is to select (f', d') to minimize the mean-square-error, under the assumption of perfect decisions ($e = 0$),

$$\text{MSE} = \mathbf{E}\|(f' * w)_t + (\tilde{h}' * a)_t\|_2^2 \tag{22}$$

Under the assumption that w_t is gaussian IID with variance σ_w^2 and a_t is gaussian IID with unity variance, the MMSE design is obtained from:

$$\min_{f',d'} \; \sigma_w^2 \|f'\|_2^2 + \|g * f' - \lambda^\delta * (1 + d')\|_2^2 \tag{23}$$

Since any choice of $h_\delta > 0$ does not effect the SER, the DFE taps are re-scaled so that $h_\delta = 1$. Thus, from the optimum solution (f', d'), set $f = f'/h'_\delta$, $d = d'/h'_\delta$ with $h'_\delta = (g * f')_\delta$.

To incorporate the possibility of one error over the tap length of the decision filter $D(\lambda)$, following (21) suggests the design optimization:

$$\min_{\substack{f',d' \\ \|d'\|_\infty \leq \gamma}} \; \sigma_w^2 \|f'\|_2^2 + \|g * f' - \lambda^\delta * (1 + d')\|_2^2 \tag{24}$$

This approach penalizes the largest decision filter coefficient by introducing the d-tap constraint γ. If $\gamma = \infty$ then we return to the MMSE design. As γ decreases we sweep out new designs. Simulations of these designs (section 5) show modest robustness gains to certain types of error propagation environments.

[1] If it turns out that $h_\delta < 0$, then replace the quantizer, in the binary case only, with $Q(x) = -\text{sgn}(x)$.

Another interpretation is that this approach models error propagation *as if it were an exogenous noise*, *e.g.*, e is a random sequence with an infrequent single error, at most one error possibly every m samples. (Recall m is the decision tap length). Proceeding in this way we can also consider other models of e. For example, if we assumed that e was white Gaussian with variance σ_e^2, then DFE designs could be obtained from,

$$\min_{f',d'} \quad \sigma_w^2\|f'\|_2^2 + \|g * f' - \lambda^\delta * (1+d')\|_2^2 + \sigma_e^2\|d'\|_2^2 \quad (25)$$

In this case σ_e is a design variable which weighs the effect of constraining the d'-taps under the two-norm. In a manner similar to the previous formulation (24), if $\sigma_e = 0$ then we return to the MMSE design and as σ_e increases we sweep out new designs. Simulations with this approach are also examined in section 5.

ℓ-ary alphabet
In the ℓ-ary case, $\mathbf{A} = \{\pm 1, \pm 3, \ldots, \pm(\ell - 1)\}$, and hence,

$$\hat{a}_t = \arg\min_{\alpha \in \mathbf{A}} |h_\delta a_{t-\delta} + \varepsilon_t - \alpha| \quad (26)$$

The choice of h_δ is now not arbitrary in magnitude. Since the quantizer will return the correct symbol if and only if

$$|(h_\delta - 1)a_{t-\delta} + \varepsilon_t| \leq 1 \quad (27)$$

it follows that $h_\delta = 1$ is required, otherwise even when $\varepsilon_t = 0$ there would be an error. But setting $h_\delta = 1$ is equivalent to the procedure proposed for the binary case.

Multiple errors
We can go a bit further and account for the possibility of more than one error over the tap length of the decision filter $D(\lambda)$. Say there are $k \leq m$ errors. Then the worst case is,

$$|(d * e)_t| \leq \|e\|_\infty \sum_{i=1}^{k} |d_{(i)}| \quad (28)$$

where $\{d_{(1)}, \ldots, d_{(m)}\}$ are the impulse response coefficients of $D(\lambda)$ ordered by magnitude, that is,

$$|d_{(1)}| \geq |d_{(2)}| \geq \cdots \geq |d_{(m)}| \quad (29)$$

Hence, we can also consider the design obtained from the optimization,

$$\min_{\substack{f',d' \\ \sum_{i=1}^{k} |d_{(i)}| \leq \gamma}} \quad \sigma_w^2\|f'\|_2^2 + \|g * f' - \lambda^\delta * (1+d')\|_2^2 \quad (30)$$

5 Simulation examples

In this section, we evaluate the performance of the following three DFE designs based on the criteria discussed in the previous sections: Two of them are minimizing the MSE cost function (23)

$$J = \sigma_w^2\|f'\| + \|g * f' - \lambda^\delta * (1+d')\|_2^2$$

with the following constraints on the normalized decision taps d':

i) one-tap constraint (24):

$$|d'_i| < d'_{\max}, \text{ for all } i$$

ii) two tap constraint (30):

$$|d'_i| + |d'_j| < d'_{\max}, \text{ for all } i, j,\ i \neq j$$

This is equivalent to (30) with $k = 2$.

For comparison we also consider the cost function (25).

iii) ℓ_2-norm DFE:

$$J' = J + \sigma_e^2\|d'\|_2^2$$

For a moderate channel (Figure 1-a)) and a relatively severe channel (Figure 1-b)), the symbol error rate (SER) of the above three DFE designs are calculated by simulations using 1.5×10^6 of 8-PAM symbols. Table 1 presents SER for each channel and various SNR under different tap-length settings. In each case the choice of the design parameters $d'_{\max}$ or σ_e and δ are chosen to minimize SER from the simulations, *e.g.*, figure 2.

	Channel c_1 ($N_f = 20$, $N_d = 34$)		
SNR	15dB	20dB	25 dB
MMSE	0.3725	0.1197	0.0004
ℓ_∞	0.3512	0.0988	0.0003
2-tap	0.3483	0.0944	0.0003
ℓ_2	0.3656	0.0973	0.0003

a) Moderate channel with properly modeled DFEs

	Channel c_1 ($N_f = 10$, $N_d = 15$)		
SNR	15dB	20dB	25 dB
MMSE	0.4348	0.3211	0.2390
ℓ_∞	0.4180	0.3093	0.2211
2-tap	0.4109	0.2850	0.2067
ℓ_2	0.4095	0.2898	0.2117

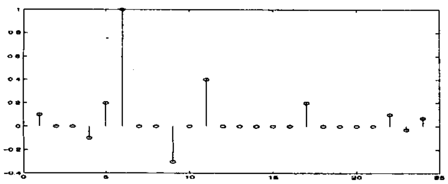
b) Moderate channel with under modeled DFEs

	Channel c_2 ($N_f = 20$, $N_d = 34$)		
SNR	15dB	20dB	25 dB
MMSE	0.5333	0.4360	0.0130
ℓ_∞	0.4828	0.3783	0.0093
2-tap	0.4829	0.3544	0.0118
ℓ_2	0.4791	0.3530	0.0137

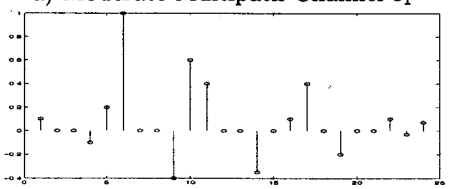
c) Severe channel with properly modeled DFEs

Table 1: SER result of modified DFEs

These results indicate that the modified DFE designs yield better SER than the conventional MMSE DFE



a) Moderate Multipath Channel c_1



b) Severe Multipath Channel c_2

Figure 1: Two Channel Models

design. Specifically, we can find that 2-tap constrained DFE performs always better than ℓ_∞ DFE, while ℓ_2 constrained DFE tends to outperform 2-tap constrained DFE for a severe channel and severe noise environment. Figure 3 shows some examples of DFE taps.

6 Concluding Remarks

Assuming the channel is known, we have shown that somewhat better DFE designs can be obtained by accounting for error propagation. Standard MMSE designs assume perfect decisions. The methods presented here utilize very simple representation of the error propagation. Although properly characterizing error propagation is difficult, the results here motivate a deeper analysis of typical error patterns which can be used to form design constraint on the decision taps.

References

[1] M. Ghosh, "Analysis of the MMSE-DFE with Error Propagation" IEEE Globecom commun. Theory Mini Conf., vol 4. Nov. 1997, pp 85-89.

[2] R.D. Giltin, J.F. Hayes, and S.B. Weinstein, *Data Communication Principles*, Plenum Press 1992

[3] R. A. Kennedy, B. D. O. Anderson, and R. R. Bitmead, "Tight bounds on the error probabilities of decision feedback equalizers," *IEEE Transactions on Communications*, vol. com-35,no. 10, pp.1022-1028, October, 1987.

[4] E. A. Lee and D. G. Messerschmitt, *Digital Communication*, 2nd Edition, Kluwer Academic Press, 1994.

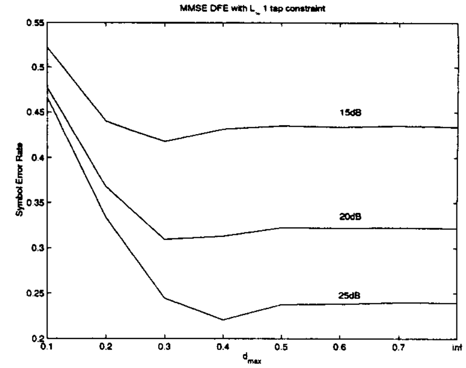
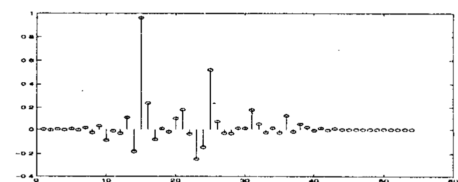
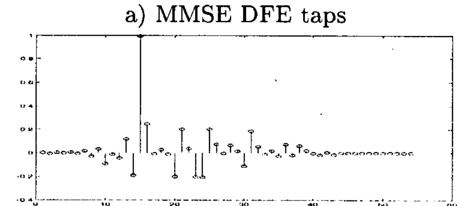


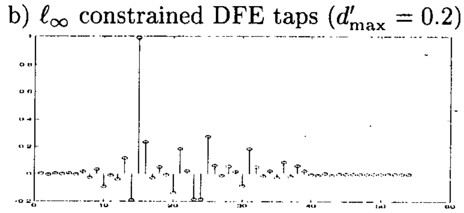
Figure 2: SER vs. $d'_{\max}$ for undermodeled ℓ_∞ constrained DFE for c_2. For each $d'_{\max}$ (and δ) the DFE with optimal (f, d) are simulated and SER computed.



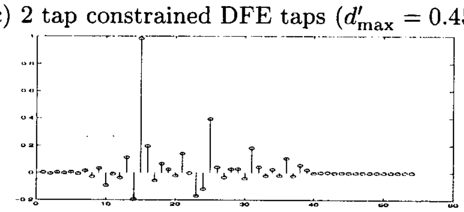
a) MMSE DFE taps



b) ℓ_∞ constrained DFE taps ($d'_{\max} = 0.2$)



c) 2 tap constrained DFE taps ($d'_{\max} = 0.45$)



b) ℓ_2 constrained DFE taps ($\sigma_e = 3.6 \times 10^{-3}$)

Figure 3: DFE taps (FFF+FBF) for channel c_1 at SNR=20dB with N_f=20 and $N_d = 34$