DFE Tutorial

Raúl A. Casas, Philip B. Schniter,
Jaiganesh Balakrishnan, C.R. Johnson, Jr. & C.U. BERG

July 14, 1998

1 Introduction

Due to the continuous increase in transmission rates of digital communication systems, the channel equalization community has recently given much attention to decision feedback equalization. The main reason for this shift in focus from linear equalization to combined linear and (nonlinear) decision feedback equalization (DFE) is due to the fact that DFEs offer intersymbol interference (ISI) cancellation with reduced noise enhancement, and may thus provide significantly lower symbol error rates (SER) than a linear equalizer, which reduces ISI with potentially greater amplification of noise. On the other hand, due to the nonlinear feedback nature of DFEs, noise induced symbol errors may trigger bursts of errors which can lead to poor SER performance. This phenomenon, called error propagation, is not well understood and mystifies researchers and practitioners as to when it is a problem and when it is not.

Behavior of DFEs is quite distinct in two types of operating scenarios. In the first scenario, the channel to be equalized has strong ISI but low noise power, or high signal to noise ratio (SNR), as is the case in high data rates, microwave links. Here, the DFE has the job of providing very low symbol error rates. Noise induced *primary errors* occur infrequently and may result in short bursts of errors (and long bursts, which apparently occur infrequently in practice), but which do not affect significantly the SER. In the second scenario, the DFE operates in a low SNR environment in conjuction with a coding system. The DFE is used to reduce ISI enough for the decoder to make sequence estimates with very low BERs. This equalization scheme has been proposed for high definition television (HDTV) standards.

One must bare in mind that in practice, at startup, the receiver does not have knowledge of the channel dynamics, and that it is also possible that the channel varies in time. Hence, a DFE receiver must be implemented as an adaptive filter which can converge to a solution that satisfactorily reduces SER and can also track time variations in this solution as the channel varies in time. When possible, a training sequence from the transmitter may be used to adapt the receiver. In uncooperative or broadcast communication systems, however, adaptation must be done without *blindly*, without the aid of the transmitter. Again, due to the nonlinear feedback nature of the DFE, blind adaptation of DFEs is extremely difficult: if the DFE is not properly initialized, a blind adaptive DFE algorithm may converge to unacceptable closed-eye parametrizations.

This tutorial covers the basics of DFE including finite and infinite length filter design, joint DFE and coding, error propagation analysis, and blind adaptation with the goal of providing some background and intuition for dealing with the practical DFE problems described above. Each section contains illustrative examples and points out to related papers in the literature. The last section of the tutorial describes THE DFECATOR, a software package written as a tool for DFE research and design.

2 DFE Structure

Let us begin by defining the communications model for studying the DFE. (refer to Figure 1). We will also adopt the following assumptions:

- Source, channel, noise and equalizer are real-valued.

  Extension to complex-valued communications is simple.

- Baud-spaced equalization.

  The conversion to fractionally-spaced equalizationis virtually accomplished by redefining the channel convolution matrix.

- (Linear) time invariant (LTI) channel.

  Time-variations of the channel are avoided for simplicity, but they are a phenomenon of importance which make analysis of DFE behavior very difficult.

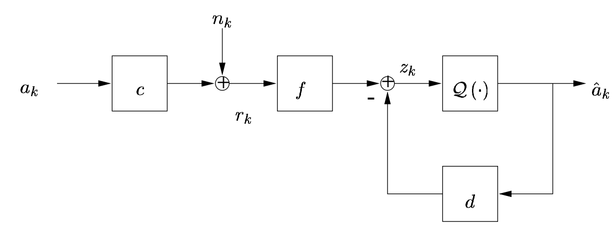

Figure 1: Baud rate DFE communications model.

A sequence $\{a_k\}_{k\in\mathbb{Z}}$ of symbols $a_k \in \mathcal{A}$ from a finite alphabet $\mathcal{A}$ is transmitted every T-seconds. This source signal, with power σ_a^2, is transmitted over a linear, time-invariant (LTI) finite-impulse response (FIR) channel $c = (c_0, c_1, \ldots, c_{N_c})^t$, $c_i \in \mathbb{R}$ with additive white noise $n_k \in \mathbb{R}$ of zero mean and variance σ_n^2. The received signal is then given by

$$r_k = \sum_{i=0}^{N_c} c_i a_{k-i} + n_k.$$

Equalization is achieved via feedforward $f = (f_0, \ldots, f_{N_f})^t$, $f_i \in \mathbb{R}$ and feedback $d = (d_1, \ldots, d_{N_d})^t$, $d_i \in \mathbb{R}$ filters yielding soft estimates

$$z_k = \sum_{i=0}^{N_f} f_i r_{k-i} - \sum_{i=1}^{N_d} d_i \hat{a}_{k-i}$$

which are quantized by a decision device $\mathcal{Q}$ into hard decisions

$$\hat{a}_k = \mathcal{Q}(z_k) = \arg\min_{\alpha\in\mathcal{A}} |\alpha - z_k|.$$

It will prove convenient to define the channel convolution matrix

$$\mathcal{C} = \begin{pmatrix} c_0 & & & \\ c_1 & c_0 & & \\ c_2 & c_1 & & \\ \vdots & c_2 & & c_0 \\ c_{N_c} & \vdots & \ddots & c_1 \\ & c_{N_c} & & c_2 \\ & & & \vdots \\ & & & c_{N_c} \end{pmatrix}.$$

as well as the combined channel and equalizer impulse response

$$h = (h_0, \ldots, h_{N_h})^t = \mathcal{C} f.$$

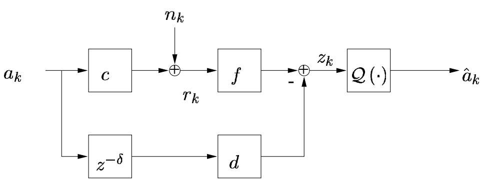

Figure 2: DFE model under key assumption.

We will now try to better understand the decision feedback equalization mechanism by using the standard trick for analyzing DFEs, the *key* assumption of *correct feedback decisions* (refer to Figure 2), *i.e.* $\hat{a}_k = a_{k-\delta}$, for some $\delta \geq 0$. Under this key assumption

$$z_k = \underbrace{h_\delta a_{k-\delta}}_{\substack{\text{information}\\ \text{bearing}\\ \text{cursor}}} + \underbrace{\sum_{i=0}^{\delta-1} h_i a_{k-i}}_{\substack{\text{residual}\\ \text{precursor}\\ \text{ISI}}} + \underbrace{\sum_{i=1}^{N_d} (h_{i+\delta} - d_i) a_{k-\delta-i}}_{\substack{\text{modeled}\\ \text{postcursor}\\ \text{ISI}}} + \underbrace{\sum_{i=\delta+N_d+1}^{N_h} h_i a_{k-i}}_{\substack{\text{residual}\\ \text{postcursor}\\ \text{ISI}}} + \underbrace{\sum_{i=0}^{N_f} f_i n_{k-i}}_{\substack{\text{filtered}\\ \text{noise}}}$$

Therefore, one way to equalize the channel effectively would be to use the feedforward filter to

- shape the combined channel-equalizer impulse response h so that it has a strong (unbiased, meaning unity gain) cursor $h_\delta \approx 1$,
- small residual ISI $h_i \approx 0$ for $0 \leq i < \delta$ and $\delta + N_d < i \leq N_h$,
- while keeping the noise gain, $\sum_i |f_i|^2$, as small as possible,

and the feedback filter to

- exactly cancel the remaining ISI by matching DFE taps $d_i = h_{i+\delta}$, $1 \leq i \leq N_d$.

This is essentially what the optimal design in terms of MSE does, as will become apparent in the next section.

3 MMSE-DFE: finite length filters

Let the mean square error (MSE) in recovering delayed source symbols (given some parametrization (f, d) of the DFE) be

$$\sigma_\varepsilon^2 \equiv \mathrm{E}\{(a_{k-\delta} - z_k)^2\}.$$

We call $\varepsilon_k = a_{k-\delta} - z_k$ the *source recovery error*, or *soft decision error*. Our aim is to compute the pair $(f_\delta^\dagger, d_\delta^\dagger)$ which minimizes σ_ε^2. To do this, we define symbol regressors

$$A_k = (a_k, \ldots, a_{k-N_h})^t, \quad \hat{A}_k = (\hat{a}_{k-1}, \ldots, \hat{a}_{k-N_d})^t, \quad N_k = (n_k, \ldots, n_{k-N_f})^t.$$

Assuming correct past decisions, $\hat{a}_k = a_{k-\delta}$

$$\hat{A}_k = (a_{k-\delta-1}, \ldots, a_{k-\delta-N_d})^t.$$

We then find MSE is a quadratic cost function with respect to equalizer parameters f and d

$$\sigma_\varepsilon^2 = \mathrm{E}\{(\underbrace{a_{k-\delta}}_{\text{desired}} - \underbrace{A_k^t \mathcal{C} f + \hat{A}_k^t d - N_k^t f}_{\text{equalized signal}})^2\}$$

which can be minimized by setting the gradient of σ_ε^2 with respect to f and d to zero. First,

$$\nabla_d \sigma_\varepsilon^2 = \mathrm{E}\left\{-2\hat{A}_k A_k^t \mathcal{C} f + 2\hat{A}_k \hat{A}_k^t d - 2\hat{A}_k N_k^t f + 2a_{k-\delta}\hat{A}_k\right\}.$$

When the input and noise processes are mutually uncorrelated [1] the expression simplifies to

$$\nabla_d \sigma_\varepsilon^2 = -2\sigma_a^2 M \mathcal{C} f + 2\sigma_a^2 d$$

where we use $\mathrm{E}\left\{\hat{A}_k \hat{A}_k^t\right\} = \sigma_a^2 I$, $\mathrm{E}\left\{a_{k-\delta}\hat{A}_k\right\} = 0$, $\mathrm{E}\left\{\hat{A}_k A_k^t\right\} = \sigma_a^2 M$ and [2]

$$M = \left(0_{N_d \times \delta} \quad I_{N_d \times N_d} \quad 0_{N_d \times N_h - N_d - \delta}\right).$$

Hence, the MMSE feedback filter can be written as

$$d_\delta^\dagger = M \mathcal{C} f.$$

Similarly, for the feedforward parameters we have

$$\nabla_f \sigma_\varepsilon^2 = 2\sigma_a^2 \mathcal{C}^t \mathcal{C} f - 2\sigma_a^2 \mathcal{C}^t M^t d - 2\sigma_a^2 \mathcal{C}^t e_\delta + 2\sigma_n^2 f$$

where $e_\delta = (0, \ldots, 0, 1, 0, \ldots, 0)^t$ is the standard basis vector, with one at position δ, $0 \leq \delta \leq N_h$. Assuming mutually uncorrelated source and noise, substituting $d_\delta^\dagger = M\mathcal{C}f$ results in

$$\nabla_f \sigma_\varepsilon^2 = 2\sigma_a^2 \mathcal{C}^t (I - M^t M) \mathcal{C} f - 2\sigma_a^2 \mathcal{C}^t e_\delta + 2\sigma_n^2 f$$

yielding the MMSE feedforward equalizer

$$f_\delta^\dagger = \left(\mathcal{C}^t P \mathcal{C} + \lambda I\right)^{-1} \mathcal{C}^t e_\delta$$

[1] Both input and noise sequences are white processes and are uncorrelated with each other.

[2] Note that this formulation of M restricts the feedback filter length to $N_d \leq N_h - \delta$. This simply means that additional feedback taps are not needed and should be set to zero. One could also extend our derivation to longer feedback filters by zero-padding the combined channel-feedforward equalizer impulse response.

where $\lambda = \sigma_n^2/\sigma_a^2$ and $P = (I - M^t M)$.

The problem of minimizing MSE for mutually uncorrelated source and noise and some fixed delay δ translates to

$$f_\delta^\dagger = \arg\min_f \underbrace{\|P(\mathcal{C}f - e_\delta)\|^2}_{\substack{\text{residual}\\ \text{pre- and}\\ \text{postcursor}\\ \text{ISI}}} + \underbrace{\lambda\|f\|^2}_{\substack{\text{amplified}\\ \text{noise}}} \quad , \quad d_\delta^\dagger = \arg\min_d \underbrace{\left\|d - M\mathcal{C}f_\delta^\dagger\right\|}_{\substack{\text{modelled}\\ \text{postcursor}\\ \text{ISI}}}.$$

Note that MSE is also a function of delay δ and that optimization of delay requires an exhaustive search over the range $0 \leq \delta \leq N_h$.

3.1 Example

Using a simple example we will show the main design issues in DFE design. Our task is to design an equalizer with at most $N_f + N_d = 64$ parameters for the ISI channel shown in Figure 3, at 15 dB SNR. The constraint on the total number of parameters mimicks the problem of limited hardware resources which is in practice complicated by other issues such as cost, ease of implementation, and equalization requirements.

As designers, we will have to make two major choices. **First:** how do we partition the total number of taps into the feedforward and feedback filters? **Second:** which overall system delay do we pick? The brute force solution to these questions is to try every possible combination of (N_f, N_d, δ). Some useful guidelines, however, may help to make the search more focused:

- In practice, it is often the case that a channel has a strong cursor with some delay $\nu \geq 0$ (for the channel in Figure 3, $\nu = 7$), and possibly a few secondary rays occuring at later (and less frequently at earlier) time delays. Symbols delayed by the channel cursor delay have the most energy. Consequently, it makes sense to attempt to recover them by picking the overall system delay δ to match the channel cursor delay, *i.e.* $\delta = \nu$. It also makes sense to pick a larger system delay, *i.e.* $\delta > \nu$ since the extra delay may be implemented by the feedforward filter. This extra delay, however is limited by the feedforward equalizer length, *i.e.* $\delta < \nu + N_f$. Hence, the range of *good* delays is $\delta \in [\nu, \nu + N_f]$.

- As we have seen, the role of the feedback taps is to additively cancel ISI without introducing noise gain. Therefore, one can think of placing a DFE *window* to cancel large channel postcursor taps. In our sample channel the cursor is at delay index $\nu = 7$ and the portion of the postcursor with significant ISI lies around delay index 8 to about delay index 25. This gives us a general idea of how to partition the available parameters.

We use these two *rules-of-thumb* in our sample design. To match the channel postcursor, let us use $N_d = 32$ feedback taps, leaving $N_f = 32$ feedforward taps. We choose an overall system cursor $\delta = 25$ approximately in the center of the range of *good* delays $[7, 39]$. For comparison, we will also try a linear design with $N_f = 64$ feedback taps (and $N_d = 0$ feedback taps). MSE versus system delay is shown in Figure 3. Notice the trough of low MSEs as predicted above, *i.e.* $[7, 39]$ for the DFE design and $[7, 71]$ for the linear design. Also notice that the DFE design outperforms the linear design due to a clever tap partitioning.

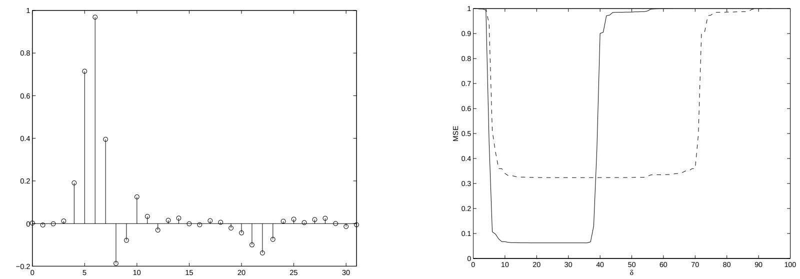

Figure 3: (left) Channel impulse response, (right) MSE $\mathrm{E}\left\{(s_{k-\delta} - z_k)^2\right\}$ versus delay choice δ for (solid) DFE design with $N_f = 32$ feedforward taps and $N_d = 32$ feedback taps and (dashed) linear equalizer design with $N_f = 64$ feedforward taps.

3.2 Remarks

Extended Analysis A more general derivation for the finite length MMSE-DFE that includes fractionally-spaced feedforward filters and colored source and noise can be found in [Al-Dhahir TIT 95a]. By making the assumption of a *sufficiently long feedback filter*, *i.e.* $N_d > N_h$, this paper also gets a handle on delay optimization and statistics of the recovery error ε_k. Under this (strong) assumption a single matrix inversion provides the facility of determining the optimal delay δ. The paper draws comparisons between residual errors for the finite and infinite length DFEs. We briefly note that the residual error will in general be correlated due to the filtering process through the feedforward filter and residual ISI taps, and non-Gaussian due to residual ISI (unless the source is Gaussian).

Unbiased design [Al-Dhahir TIT 95a] also emphasizes the importance of having an *unbiased* design, where MSE is minimized under the equality constraint $h_\delta = 1$. Under this constraint MSE is higher than MSE achieved without the constraint. On the other hand, the probability of error given by the unbiased design is lower the the probability of error from the unconstrained design. Computing the unbiased MMSE-DFE (f_δ^u, d_δ^u) is easy

$$f_\delta^u = \frac{1}{\alpha} f_\delta^\dagger, \quad d_\delta^u = \frac{1}{\alpha} d_\delta^\dagger$$

where $\alpha = [Cf_\delta^\dagger]_{\delta,1}$ is the δ^{th} tap of the combined channel-(unconstrained, MMSE) feedforward filter impulse response. This fact can be proved using *Lagrange multipliers*.

Practical design issues In the design example above we encountered to questions: how to partition parameters between feedorward and feedback filters and how to select system delay. At the current moment, the literature has no practical systematic way of answering these questions, much less for the blind scenario. Schemes which aid the designer to overcome these barriers would certainly be welcomed by the DFE community.

4 MMSE-DFE: infinite length filters

The finite-length MMSE-DFE was expressed as the solution to a finite system of linear equations. The infinite-length solution will be derived using rational transfer functions instead. Let $C(z)$ denote a stable, causal channel transfer function that is related to the impulse response coefficients $\{c_i\}$ through

$$C(z) \quad = \quad \sum_{i=0}^{\infty} c_i z^{-i},$$

our task is to design rational transfer functions representing the forward and feedback filters. Since the feedback filter relies on past decisions, it must be strictly causal.

The IIR-DFE design can be motivated through consideration of IIR linear equalizer performance. Consider the error system illustrated by Figure 4(a), where $F(z)$ denotes the transfer function of a baud-spaced forward equalizer. (As the overall system delay becomes inconsequential with IIR equalization, we assume zero delay on the reference path.) Assuming mutually uncorrelated source and noise processes, the recovery error power spectrum $S_\varepsilon(e^{j\omega})$ takes the form

$$S_\varepsilon(e^{j\omega}) \quad = \quad \left|C(e^{j\omega})F(e^{j\omega}) - 1\right|^2 \sigma_a^2 + \left|F(e^{j\omega})\right|^2 \sigma_n^2.$$

Completing the square,

$$S_\varepsilon(e^{j\omega}) \quad = \quad \left(|C(e^{j\omega})|^2 \sigma_a^2 + \sigma_n^2\right) \left|F(e^{j\omega}) - \frac{C^*(e^{j\omega})}{|C(e^{j\omega})|^2 + \lambda}\right|^2 + \frac{\sigma_n^2}{|C(e^{j\omega})|^2 + \lambda},$$

where $\lambda = \sigma_n^2/\sigma_a^2$. Recalling that $\sigma_\varepsilon^2 = \int S_\varepsilon(e^{j\omega})d\omega$, the IIR forward equalizer minimizing σ_ε^2 is given by

$$F(e^{j\omega}) \quad = \quad \frac{C^*(e^{j\omega})}{|C(e^{j\omega})|^2 + \lambda}$$

for which the error power spectrum becomes

$$S_\varepsilon(e^{j\omega}) \quad = \quad \frac{\sigma_n^2}{|C(e^{j\omega})|^2 + \lambda}.$$

Noting that the corresponding error sequence $\{\varepsilon_k\}$ is (in general) not white, the theory of linear prediction tells that the error variance may be reduced by proper whitening of the sequence. The filter which accomplishes this is referred to as a prediction error filter [Haykin Book 96] and is illustrated by $D(z)$ in Figure 4(a). Manipulation of the filtered-error-system block diagram reveals its equivalence to the IIR DFE shown in Figure 4(b) [Lee Book 94]. We summarize key properties of properly designed $D(z)$:

- As a prediction error filter, $D(z)$ is guaranteed to be monic[3] and causal, and thus $1 - D(z)$ is an appropriate (*i.e.* strictly causal) DFE feedback filter.
- The IIR DFE error sequence ε'_k is white. This is advantageous for many detection schemes.
- The variance of ε'_k is less than or equal to the variance of ε_k, implying a properly designed IIR DFE is at least as good as an IIR linear equalizer.

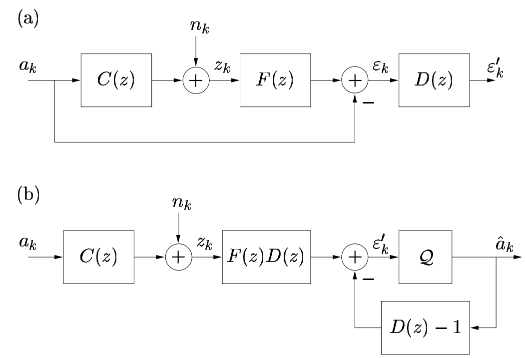

Figure 4: Two equivalent error system models: (a) forward equalizer followed by prediction error filter $D(z)$ and (b) DFE (assuming correct decisions).

Next we derive the MMSE-optimal $F(z)$ and $D(z)$. For any choice of $F(z)$, the prediction error filter $D(z)$ is completely determined by S_ε. Specifically, if a spectral decomposition [4] is used to split S_ε into its minimum-phase and maximum-phase components:

$$S_\varepsilon(z) \quad = \quad \sigma_\varepsilon^2 G_\varepsilon(z) G_\varepsilon^*(1/z^*),$$

where $G_\varepsilon(z)$ is monic, causal, and (loosely) minimum phase, then the optimal prediction error filter is given by

$$D^\dagger(z) \quad = \quad \frac{1}{G_\varepsilon(z)}.$$

It can be shown that MMSE $F(z)$ is designed just as in the linear case [Lee Book 94], so as before

$$F^\dagger(z) \quad = \quad \frac{C^*(1/z^*)}{|C(z)|^2 + \lambda}.$$

For a more concise description of the DFE feedforward filter $F(z)D(z)$, we employ a spectral decomposition of the received signal spectrum:

$$S_z(z) = \sigma_a^2 |C(z)|^2 + \sigma_n^2 = \sigma_z^2 G_z(z) G_z^*(1/z^*)$$

[3] For a polynomial to be monic, its constant term must equal one. Monic rational transfer functions are characterized by monic numerator and denominator polynomials.

[4] A non-negative real transfer function has complex conjugate pole pairs and zero pairs which may be split evenly between (loosely) minimum phase and (loosely) maximum phase transfer functions. Pole/zero pairs on the unit circle are split likewise.

to write

$$
\begin{aligned}
S_\varepsilon(z) &= \frac{\sigma_a^2\sigma_n^2}{\sigma_z^2 G_z(z) G_z^*(1/z^*)}, \\
D^\dagger(z) &= G_z(z), \\
F^\dagger(z) &= \frac{\sigma_a^2 C^*(1/z^*)}{\sigma_z^2 G_z(z) G_z^*(1/z^*)},
\end{aligned}
$$

which implies

$$
F^\dagger(z) D^\dagger(z) = \frac{\sigma_a^2}{\sigma_z^2} \cdot \frac{C^*(1/z^*)}{G_z^*(1/z^*)}
$$

From Figure 4(a), the power spectrum of the whitened error is

$$
S_{\varepsilon'}(e^{j\omega}) = S_\varepsilon(e^{j\omega})|D^\dagger(e^{j\omega})|^2 = \frac{\sigma_a^2\sigma_n^2}{\sigma_z^2},
$$

from which it is apparent that the MMSE of the IIR-DFE

$$
\sigma_{\varepsilon'}^2 = \frac{\sigma_a^2\sigma_n^2}{\sigma_z^2}.
$$

Note that the MMSE DFE feedforward filter is anti-causal stable, while the feedback filter $1 - D^\dagger(z)$ is strictly causal stable. Though the anticausal IIR feedforward filter is not implementable, it can be approximated by a causal FIR filters of sufficient length. In fact, the key properties of the MMSE IIR DFE (such as a white soft-error sequence) are expected to well characterize long FIR approximations.

5 Error Propagation

Up to this point, we have always relied on the *key* assumption of correct past decisions for studying DFE performance. We will now break this assumption and learn about the shortcomings of the DFE due to error propagation. To understand the error propagation mechanism, consider Figure 5 showing error sequences for 4-PAM communication system (*i.e.* $\mathcal{A} = \{\pm 1, \pm 3\}$). This figure compares the DFE under the correct past decision assumption, as shown in Figure 2 with the regular DFE in 1.

The *dashed* line shows soft decision errors, $\varepsilon_k = a_{k-\delta} - z_k$ for the correct past decision DFE. Notice that when residual error is enough to cross the *dotted* decision boundary line, a hard decision error $e_k = a_{k-\delta} - \hat{a}_k$, designated with a $\circ$, occurs. No subsequent errors occur until the residual error is large enough again, about 100 symbols later.

When decision error feedback comes into play, the picture is slightly different. While there are no decision errors both structures share the same error sequence. When a decision error occurs, however, both soft and hard error sequences diverge. The first residual-induced decision error called a *primary* error (shown in Figure 5 with overlapping $\circ$ and $\star$ since this first decision error is shared by both structures in Figures 2 and 1) is fed back by the DFE causing *secondary* errors (denoted with $\star$'s), creating an error burst. Notice the oscillating pattern of the soft residual errors plotted with a *solid* line. The burst terminates when the residual errors add up in the right way to flush out all decision errors in the feedback, so that the error sequences for both structures converge again. The event repeats itself about 100 symbols later.

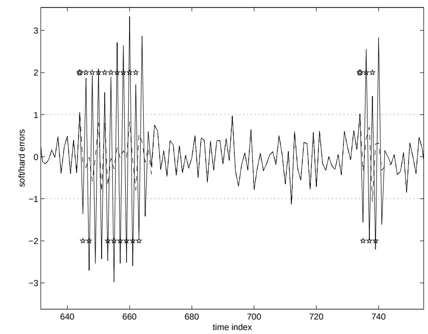

Figure 5: Error burst for regular DFE (Figure 1) and *key assumption* DFE without error feedback (Figure 2).

One way of studying error propagation is to analyze the finite state machine formed by the overall communication system, with states formed by source symbols in the channel-equalizer memory and symbol estimates in the feedback filter tap-delay line. This is apparent if we write

$$
\hat{a}_k = \mathcal{Q}\left(a_{k-\delta} + \sum_{i=1}^{N_d} h_{i+\delta}(a_{k-\delta-i} - \hat{a}_{k-i}) + \varepsilon_k\right)
$$

where we have used $d_i = h_{i+\delta}, 1 \leq i \leq N_d$, and where ε_k is the symbol recovery error, clumping precursor ISI, filtered noise and unmodeled postcursor ISI. The finite state machine structure is described as follows: the state of the system

$$
X_k = (a_{k-\delta-1}, \ldots, a_{k-\delta-N_d} | \hat{a}_{k-1}, \ldots, \hat{a}_{k-N_d}) \in \mathcal{A}^{2N_d}
$$

depends on the source symbols and their estimates, the transitions are brought about by the joint input $a_{k-\delta}$ and ε_k, and the system output consists of the hard decisions $\hat{a}_k$. To reduce the complexity of the model, we will assume that the sequence ε_k is i.i.d and has zero-mean Gaussian distribution with variance σ_ε^2. The system can now be described as *Markov chain.*

Once all the possible states (there are a total of $|\mathcal{A}|^{2N_d}$ states, wher $|\mathcal{A}|$ is the cardinality of the alphabet) and transition probabilities between states are determined, a transition matrix for the Markov chain provides the steady state distribution of the system. In other words, the transition matrix says how often a certain X_k will occur, so that one can simply pick out error states (where $\hat{a}_{k-1} \neq a_{k-\delta-1}$) to determine the error probability of the DFE *accounting for error propagation.*

5.1 Example

We illustrate the concepts above with the following case:

Source BPSK, *i.e.* $a_k \in \{\pm 1\}$

Channel $c = (1, 0.8)^t$, 5 dB SNR

DFE $f = 1$, $d = 0.8$.

Channel-equalizer response $h = (1, 0.8)^t$

Residual error $\varepsilon_k = n_k$

States $X_k = (a_{k-1}|\hat{a}_{k-1})$. We conveniently label the states as binary digits, *i.e.*

$$(-1,-1) \equiv <\mathbf{0}>,\ \ (-1,1) \equiv <\mathbf{1}>,\ \ (1,-1) \equiv <\mathbf{2}>,\ \ (1,1) \equiv <\mathbf{3}>$$

Input a_k

Output $\hat{a}_k = \mathrm{sgn}\,(a_k + 0.8(a_{k-1} - \hat{a}_{k-1}))$

Transition probabilities

Current State X_k	Input a_k	Soft Estimate z_k	Hard Estimate $\hat{a}_k$	Next State X_{k+1}	Probability $\Pr(X_{k+1}\|X_k, a_k)\Pr(a_k)$
$(-1,-1)$	1	$1.0 + n_k$	-1	$(1,-1)$	$p_0/2$
			1	$(1,1)$	$(1-p_0)/2$
$<\mathbf{0}>$	-1	$-1.0 + n_k$	-1	$(-1,-1)$	$(1-p_0)/2$
			1	$(-1,1)$	$p_0/2$
$(-1,1)$	1	$-0.6 + n_k$	-1	$(1,-1)$	$(1-p_1)/2$
			1	$(1,1)$	$p_1/2$
$<\mathbf{1}>$	-1	$-2.6 + n_k$	-1	$(-1,-1)$	$\approx 1/2$
			1	$(-1,1)$	≈ 0
$(1,-1)$	1	$2.6 + n_k$	-1	$(1,-1)$	≈ 0
			1	$(1,1)$	$\approx 1/2$
$<\mathbf{2}>$	-1	$0.6 + n_k$	-1	$(-1,-1)$	$p_1/2$
			1	$(-1,1)$	$(1-p_1)/2$
$(1,1)$	1	$1.0 + n_k$	-1	$(1,-1)$	$p_0/2$
			1	$(1,1)$	$(1-p_0)/2$
$<\mathbf{3}>$	-1	$-1.0 + n_k$	-1	$(-1,-1)$	$(1-p_0)/2$
			1	$(-1,1)$	$p_0/2$

where $p_0 = \Pr(1 + n_k < 0) = 0.0375$, $p_1 = \Pr(0.6 + n_k < 0) = 0.1436$.

Transition matrix $\mathcal{M}$ has entries $[\mathcal{M}]_{i,j}$ representing the *probability that a state* $<\mathbf{i}>$ *goes to a state* $<\mathbf{j}>$. Here,

$$\mathcal{M} = \begin{pmatrix} (1-p_0)/2 & p_0/2 & p_0/2 & (1-p_0)/2 \\ 1/2 & 0 & (1-p_1)/2 & p_1/2 \\ p_1/2 & (1-p_1)/2 & 0 & 1/2 \\ (1-p_0)/2 & p_0/2 & p_0/2 & (1-p_0)/2 \end{pmatrix}$$

Steady-state To find steady state probabilities of a Markov chain one solves for the left eigenvector of $\mathcal{M}$ with unity eigenvalue, under the constrained that steady-state probabilities for each state sum to unity, *i.e.*

$$\pi = \pi\mathcal{M},\ \sum_{i=0}^{3}\pi_i = 1.$$

For this problem, we find

$$\pi = (0.46923, 0.03077, 0.03077, 0.46923)^t$$

Error probability $P_e = \pi_1 + \pi_2 = 0.0615$. Through simulation with 300,000 symbols we estimate an error probability of 0.0617. Compare this error probability with the lower error probability assuming correct past decisions, $P_e = \Pr(1 + n_k < 0) = 0.0375$.

5.2 Remarks

Bounds As the constellation size and the feedback filter length increases, the total number of Markov states $|\mathcal{A}|^{2N_d}$ becomes very large and application of the ideas described above is impractical. Some research has been done to simplify the analysis by clumping error states and providing upper and lower bounds on the error probabilities given some or no knowledge of the channel dynamics, as found in [Duttweiler TIT 74, Cantoni TCOM 76, OReilly IEE 85, Oliveira IEE 85] for example. Other similar work based on Markov chain techniques classifies channel types and provides other error measures, such as time to recover from a primary error [Kennedy ISSPA 87, Kennedy TCOM 87b, Kennedy TCOM 87c].

Stability arguments It is also possible to study error propagation from a dynamical systems perspective. For instance, [Kennedy TCOM 87a] shows that a DFE equalizing a noiseless exponentially decaying impulse response channel can recover in a finite time from an error burst. The idea is extended for channels satisfying a passivity condition in [Kennedy TCOM 89a].

Countermeasures Little work has been done on detecting and reacting to error propagation. [Dogancay TIT 97] proposes techniques for detecting decision errors in equalization schemes, including DFEs, while [Fertner TSP 98] introduces a scheme which attempts to prevent primary errors and error propagation.

6 Joint DFE and Coding

Consider the problem of adding a DFE to a coded communication system, as represented in Figure 6. Ideally, one would like to ignore the behavior of the *effective channel* which is a combination of the communications channel and the equalization (DFE) scheme. In this case it is possible to predict the performance of the coded scheme since the effective channel consists solely of additive noise, n. When the DFE can perfectly cancel ISI and whitens the symbol estimatation errors, as is the case of the IIR-DFE [Cioffi TCOM 95] (assuming that fedback decisions are correct) the effective channel continues to consist solely of additive white noise, and therefore performance is predictable. When the assumption of correct past decision is violated (while keeping the assumption of perfect ISI cancellation), the effective channel consists of the sum of channel noise n and *unstructured* noise from DFE error bursts. Little is known of the effects of error propagation on error correcting decoders. One would like to know for what types of channels result in error bursts which may significantly undermine the performance of the decoder. We would also like to understand the structure of these error bursts, and to have schemes which improve SER/BER performance.

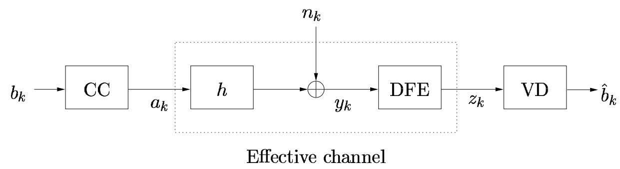

Figure 6: Joint DFE and coding communication system. A convolutional code (CC) codes bits b_k to symbols a_k and transmitts these over the effective DFE channel. Soft estimates z_k (or hard symbol estimates $\hat{a}_k$) are fed to a Viterbi Decoder for decoding into bit estimates $\hat{b}_k$.

6.1 Example

A simulation of a joint DFE and coding system will illustrate these concerns. The simulation setup consists of a source, encoder, channel and equalizer followed by a viterbi decoder as shown in Figure 6. The binary sequence was encoded using a [57] convolutional code and mapped to a 4-PAM constellation. This sequence is then transmitted through the channel and equalized by the DFE. The probability of bit errors versus SNR is studied for low SNR values in the waterfall curves in Figure 7 for

- the ideal feedback case, where we artifically feed back correct symbols through the feedback filter,
- the non-ideal case, where symbol estimates are fed back and may cause error propagation.

The error bursts are coloured and hence damage viterbi decoder performance. This can be improved to some extent by whitening the soft inputs to the viterbi decoder. One approach is to interleave the encoded symbols, and to introduce a deinterleaver prior to the viterbi decoder. The performance curves corresponding to various interleaver lengths of 48, 96 and 192 are shown in the plot. The whitening of the soft decisions improves the performance of the coded system to some extent, but there seems to be a limit to amount of reduction in error probability since the variance of soft estimation errors due to error propagation is not affected by interleaving/deinterleaving.

7 Adaptive DFEs

This section presents adaptive algorithms for implementing DFEs and discusses some of the the difficulties encountered by these algorithms due to the nonlinear feedback nature of the DFE. The algorithms are based on stochastic gradient descent schemes (SGA) where a cost function $J(\theta)$ is minimized over the parameters $\theta = (\ldots, \theta_i, \ldots)$ by taking small steps in the direction of the negative gradient, *i.e.*

$$\theta_i(k+1) = \theta_i(k) - \mu \frac{\partial}{\partial \theta_i} J(\theta).$$

In terms of DFEs, parameters θ_i consist of feedforward and feedback filter taps. For a more detailed discussion of adaptive equalization see [Haykin Book 96, Johnson PROC 98].

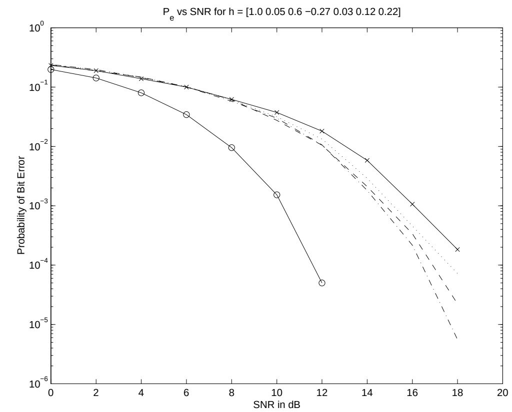

Figure 7: Channel $h = (1.0, 0.05, 0.6, -0.27, 0.03, 0.12, 0.22)^t$, ($\times$) BER of regular DFE, ($\circ$) BER of ideal DFE, and successive improvement in BER for interleavers of length ($\ldots$) 48, ($--$)96 and ($-.$) 192.

LMS with training When it is possible to transmit a training sequence, the receiver has knowledge of a sequence of source symbols a_k. In that case, it is possible to directly minimize MSE for some given delay δ by setting up the LMS-DFE cost function

$$\mathcal{J}_{lms} = \mathrm{E}\left\{(a_{k-\delta} - z_k)^2\right\}$$

(which is in fact equivalent to σ_ε^2). The adaptive algorithm ensues

$$\begin{aligned} f_i(k+1) &= f_i(k) + \mu r_{k-i}(a_{k-\delta} - z_k), \quad 0 \le i \le N_f \\ d_j(k+1) &= d_j(k) - \mu a_{k-\delta-j}(a_{k-\delta} - z_k) \quad 1 \le j \le N_d. \end{aligned}$$

It should be clear that the training symbols are not only used to update the algorithm, but also to generate soft estimates z_k (*i.e.* no feedback is actually used, *i.e.* $\hat{a}_k = a_{k-\delta}$).

Since nonlinear feedback is unnecessary under this adaptive implementation with training the LMS cost function is quadratic, unimodal and the adaptive algorithm is globally convergent. All the standard techniques for analyzing transient and steady-state response are valid for this algorithm. See [Haykin Book 96].

Decision-Direction When a training signal is not available, one can substitute delayed source symbols $a_{k-\delta}$ by their estimates $\hat{a}_k$ so as to minimize the cluster variance cost $\mathcal{J}_{dd} = \mathrm{E}\left\{(\hat{a}_k - z_k)^2\right\}$, a proxy to the true MSE. The DD-DFE algorithm

$$\begin{aligned} f_i(k+1) &= f_i(k) + \mu r_{k-i}(\hat{a}_k - z_k), \quad 0 \le i \le N_f \\ d_j(k+1) &= d_j(k) - \mu \hat{a}_{k-j}(\hat{a}_k - z_k) \quad 1 \le j \le N_d \end{aligned}$$

is blind in the sense that it does not use a training signal for parameter updates.

The weakness of DD-DFE is that it relies on symbol estimates for its update. If DD-DFE is initialized with a closed-eye it is very likely that it will get trapped in a bad minimum, possibly closed-eye as well. Usually DD-DFE is used after some other blind algorithm has opened the channel eye. One common scheme is to open the eye by adapting the feedforward equalizer with the Constant Modulus Algorithm (CMA) and then switching to DD-DFE, where the DFE is initialized at the origin (*i.e.* $d_i = 0, \forall i$. One can show that in the absence of noise DD-DFE is globally convergent to the MMSE-DFE (for a fixed delay) from an open-eye initialization. See [Kennedy IJACSP 93, Casas IJACSP 98, Casas SPAWC 97]).

CMA In the same way that CM criterion is used for linear equalization, in CMA-DFE, the cost function, $\mathcal{J}_{cma} = \mathrm{E}\left\{(z_k^2 - \gamma)^2\right\}$, $\gamma = \mathrm{E}\left\{a_k^4\right\}/\mathrm{E}\left\{a_k^2\right\}$ is used as a proxy to MSE (refer to [Johnson PROC 98]. The updates are given by

$$\begin{aligned} f_i(k+1) &= f_i(k) + \mu r_{k-i}(z_k^2 - \gamma) z_k, \quad 0 \le i \le N_f \\ d_j(k+1) &= d_j(k) - \mu \hat{a}_{k-j}(z_k^2 - \gamma) z_k \quad 1 \le j \le N_d \end{aligned}$$

Little is known about the properties of CMA-DFE. The algorithm is not globally convergent [Casas ASIL 95] under the CMA-FSE perfect equalization conditions [Johnson PROC 98], but can be shown to converge under these conditions to a perfect solution when initialized sufficiently close to it. Again, the main problem is the use of decision feedback right from a (possibly) closed-eye initial condition, where estimates are not reliable [Papadias ASIL 95, Casas ASIL 95, Tong ASIL 96].

IIR-DFE Hybridization One way to avoid the problem of feedback of incorrect decisions at startup is to adapt a linear IIR filter until it sufficiently opens the eye and then to switch to decision-direction, using the IIR filter parameters to initialize the DFE.

Several such methods have been proposed [Labat ICC 96, Mottier ICC 97] based on IIR equalization [da Rocha ICASSP 94]–[Cavalcanti SPAWC 97] where a linear prediction filter is used to whiten the received signal thereby correcting amplitude distortions from the channel, so that the combined channel-predictor response is all-pass; this filter is then followed by an IIR all-pass filter for correcting phase distortions. The idea uses a factorization of a FIR channel as the product of a minimum phase filter C_m, an all-pass filter C_a and a gain α

$$C(z) = \alpha C_m(z) C_a(z).$$

C_m contains the minimum phase roots of C (*i.e.* roots inside the unit circle) and the minimum phase relflection over the unit circle of the maximum phase roots of C (*i.e.* roots outside the unit circle). The roots of the numerator of C_a are the maximum phase roots of C, which are the inverse of the roots of its denominator (thus cancelling the artificially introduced inverse roots in C_m). Note that we have assumed the channel has no roots on the unit circle. Naturally, one way of equalizing C is to have three filters inverting C_m, C_a and α separately as shown in Figure 8. [Labat ICC 96, Mottier ICC 97] proposes blind adaptation of these elements in three separate stages

1. Take care of C_m with a blind adaptive Auto Regressive (AR) linear prediction filter

$$\begin{aligned} x_k &= r_k - g_1(k) x_{k-1} - \ldots - g_{N_g}(k) x_{k-N_g} \\ g_i(k+1) &= g_i(k) - \mu x_{k-i} x_k, \quad 1 \le i \le N_g. \end{aligned}$$

In the absence of noise and when the order N_g of the filter matches the number of roots of C_m one expects the parameters to converge to an equilibrium yielding the (stable) inverse of C_m

$$\frac{1}{1 + \sum_{i=1}^{N_g} g_i z_{-i}} = C_m^{-1}.$$

This has the effect of whitening the received signal and giving a channel-predictor combination equal to an all-pass filter. Note that the adaptive algorithm is minimizing the output power $\mathcal{J}_p = \mathrm{E}\{x_k^2\}$ of the predictor. Important questions to be resolved are what happens when the prediction filter order is smaller that the order of C_m, and can the prediction filter become unstable? Partial answers to these questions are probably found in the adaptive IIR literature.

2. Correct the amplitude with an Automatic Gain Control (AGC) to give a unity power output. This can be done by minimizing the mean square difference between the power at the output of the adaptive gain element and the desired (unity) power level

$$\begin{aligned} x'_k &= \sqrt{|\beta(k)|} x_k \\ \beta(k+1) &= \beta(k) - \mu x_k (x_k^2 - 1). \end{aligned}$$

3. Adaptation of an FIR approximation of the inverse C_a via CMA

$$\begin{aligned} y_k &= q_0 x'_k + q_1 x'_{k-1} + \ldots + q_{N_q} x'_{k-N_q} \\ q_i(k+1) &= q_i(k) + \mu x'_{k-i}(y_k^2 - \gamma) y_k, \quad 0 \le i \le N_q. \end{aligned}$$

This filter should converge to an approximate inverse of C_a.

Once these stages are completed, the receiver reverts to DFE structure: the DFE feedback coefficients are replaced by the predictor coefficients, *i.e.* $d_i = g_i$, $1 \le i \le N_d = N_g$, the feedforward coefficients of the DFE are replaced the all-pass approximation coefficients corrected by the gain, *i.e.* $f_i = \beta q_i$, $1 \le i \le N_f = N_q$. Then, given this DFE initialization, both feedforward and feedback coefficients are adapted through decision-direction, as described above (see Figure 8).

Convince yourself that the switch initializes the DD algorithm near MMSE DFE solution by first noticing that in the IIR design the predictor and phase filters can trade positions due to linearity and yield the same output. If each of the three stages of blind IIR adaptation are successful, the output of of the linear IIR filter should be close to a delayed version of the source sequence. Thus, the prediction filter will be operating on estimates of the source, just like the feedback DFE filter. Then, the phase correcting filter acts as the feedforward DFE filter.

7.1 Example

We illustrate some of the properties of the adaptive algorithms discussed above in the following example.

Source BPSK, *i.e.* $a_k \in \{\pm 1\}$

Channel $c = (1, 0.9, -0.8)^t$, no noise

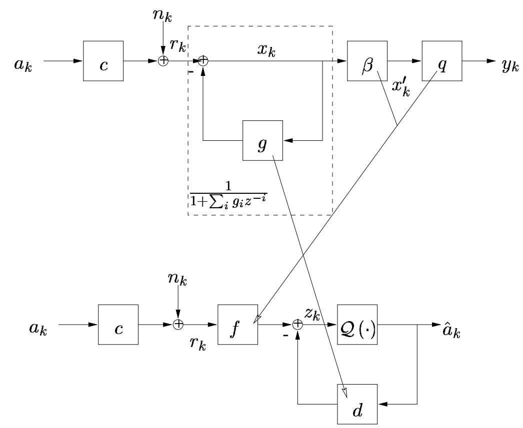

Figure 8: (top) IIR equalization converts to (bottom) DFE.

DFE $f = 1$ (fixed, non-adaptive), $d = (d_1(k), d_2(k))$

Channel-equalizer response $h = (1, 0.9, -0.8)^t$

Figure 9 shows the three cost functions associated with the first three adaptive DFE algorithms presented. As explained, the LMS-DFE cost function is unimodal while the DD-DFE and CMA-DFE cost functions have multiple local minima. All three cost functions have a global minimum at $(d_1, d_2) = (0.9, -0.8)$ where channel ISI is perfectly cancelled. Because no knowledge of the channel dynamics is available at start-up, a typical initialization of the adaptive DFE is at the origin. The algorithm will then try to *identify* the channel postcursor as best as it can to cancel ISI. If this is done, both DD-DFE and CMA-DFE get trapped at an undesired local minimum near the origin, while LMS-DFE converges to the optimal solution. We make a note that for DD-DFE and CMA-DFE, which rely on decision feedback (unlike LMS-DFE which does

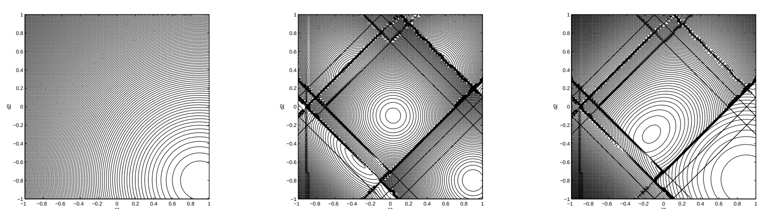

Figure 9: (left) $\mathcal{J}_{lms}$, (middle) $\mathcal{J}_{dd}$, (right) $\mathcal{J}_{cma}$.

not actually feed back decisions, but training symbols instead) the DFE quantizer, in this case a *signum* function, partitions parameter space into various regions called *polytopes* where the parameter adaptation behaves differently (refer to [Kennedy IJACSP 93, Casas IJACSP 98] for a more detailed account of the convergence properties of these algorithms).

Next, we apply the hybrid IIR-DFE blind adaptive algorithm to this test channel. At the first stage, a two-tap recursive predictor travels over the cost function shown in Figure 10, and converges to what appears to be a unique minimum at $(g_1, g_2) = (0.1378, -0.3798)$. Recall that the linear predictor should converge to the inverse of the minimum phase roots of the channel (factored into a minimum phase polynomial and an all-pass filter). This is exactly what it does since the polynomial $1 + g_1 z^{-1} + g_2 z^{-2}$ has roots $-0.6891 = 1/(-1.4512)$ and 0.5512 and the channel has roots -1.4512 and 0.5512. The triangular region drawn in dashed lines represents the region over which the polynomial $1 + g_1 z^{-1} + g_2 z^{-2}$ has both roots inside the unit circle. Note that outside the region the cost function $\mathcal{J}_p$ grows rapidly. When steady-state is reached, adaptation of the linear predictor is turned off and the phase equalizer is adapted. Figure 10 shows the magnitude of the steady-state frequency response of this element, which seems to approximate an all-pass filter. Finally, a switch to decision-direction tightens the eye, as can be seen in the same figure.

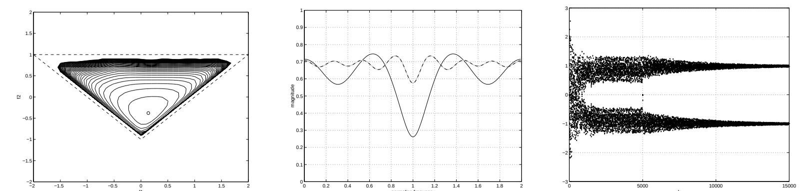

Figure 10: (left) $\mathcal{J}_p$. (middle) Frequency response magnitude of steady-state phase equalizer updated via CMA with 8 taps (solid) and 16 taps (dashed). (right) Eye diagram for adaptation of 8 tap phase equalizer from $k = 0$ to $k = 5000$ and switch to decision-direction after $k = 5000$.

7.2 Remarks

Initialization It should be clear by now that the main problem in blindly adapting a DFE is initialization, since feedback decisions at a cold start, where the channel eye may be closed, are unreliable. Several schemes have been proposed [Papadias ASIL 95, Casas SPAWC 97, Tong ASIL 96, Labat ICC 96] but they invariably fail in either speed, computational burden, or global convergence. A robust blind adaptive DFE algorithm is still to be born.

Tap partitioning Another practical problem related to adaptive DFE design is tap partitioning. Again, reliable decisions are needed to update the feedback filter, thus, one could use the feedforward filter to sufficiently open the channel eye. This filter, however, could end up requiring more taps the feedforward filter in the desired MMSE-DFE. Thus, in order to achieve a design with low complexity, one must resort to an intermediate design with high complexity. Could this be avoided? Notice that the Hybrid IIR-DFE scheme has a similar problem, where the predictor filter order must match the order of the channel. Is this filter robust to undermodeling?

8 Simulation tools: The DFEcator

The DFEcator (©1998 by R.A. Casas) is a simulation tool for designing and investigating DFEs. This Matlab based software package (work in progress) is publicly available at

`http://www.backhoe.ee.cornell.edu/ raulc/research/DFEcator`

Currently the software only deals with filter design. Future versions will include facilities for studying adaptive DFE algorithms.

A brief example elucidates the various abilities of The DFEcator.

8.1 Example: design issues revisited...

Consider equalization of the fractionally-spaced channel shown in Figure 11 with 20 dB SNR and a 64-QAM constellation.

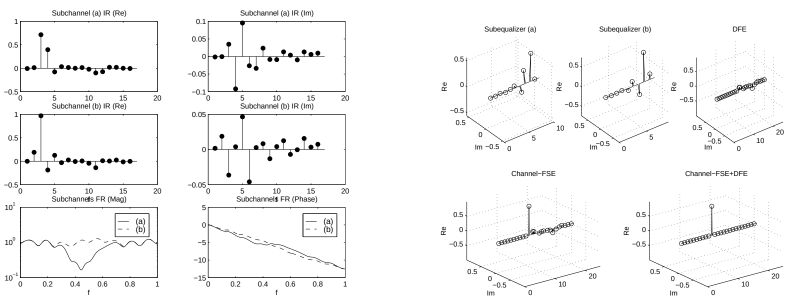

Figure 11: (left) Fractionally-spaced channel impulse and frequency response, (middle) resulting impulse response of fractionally-spaced feedforward filter, baud-spaced feedback filter, combination of channel-feedforward response and effective system response after feedback tap cancellation.

As mentioned above, the first step in the design process is to select feedforward and feedback filter lengths as well as an overall system delay. There is no straightforward method for this selection process so we will constrain the total number of equalizer taps to $N_f + N_d = 32$ and look at all resulting MMSEs as a function of N_f and δ ($N_d = 32 - N_f$), as shown in the MSE surface in Figure 12. Observe that there is a region of delay and (fractionally-spaced) feedforward filter lengths which provide low and similar performance. The best solution uses $N_f = 19, N_d = 13, \delta = 11$, and is displayed in Figure 11, with eye-diagram in Figure 12.

The final step in the design procedure is to study robustness of the design. We compute the watervall (SER vs. SNR) curve for this particular channel for a range of operating SNRs as shown in Figure 13. Note that this calculation is done under the assumption of correct past decisions. To have some idea of the severity of error propagation, The DFEcator also simulates the achieved system and plots decision errors and soft estimation errors as shown in Figure 13. In this case, it appears that error propagation effects are not crippling, which can be account for by small taps in the feedfack DFE filter (which cancel small coefficients of the combined channel-feedforward equalizer impulse response).

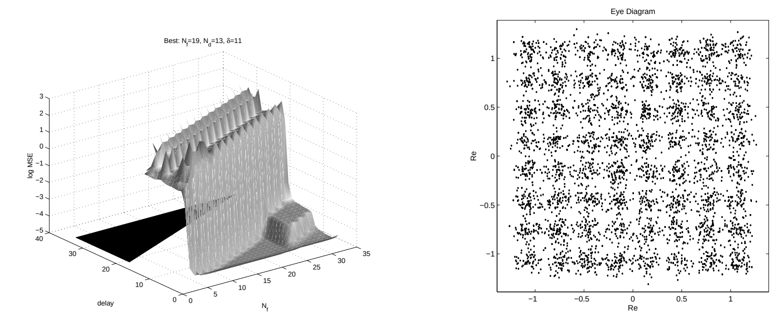

Figure 12: (left) MSE versus feedforward equalizer length and choice of system delay, (right) eye diagram for best solution.

References

General DFE

[Austin MIT 67] M. Austin, "Decision feedback equalization for digital communication over dispersive channels," *MIT Research Laboratory of Electronics Technical Report 461*, August 1967

One of the first papers defining a theory of decision feedback equalization, this paper derives an optimal equalizer (maximum likelihood estimator), a conventional equalizer (baud spaced feed forward equalizer) and a DFE. The paper then discusses issues for the conventional and decision feedback equalizers, such as the cancellation of intersymbol interference, noise gain and error propagation, and then compares these two forms of equalization through waterfall plots (SER vs. SNR). For these comparisons, it uses (and derives) maximum distortion *channels for the forward equalizer case. The paper assumes* no past decision errors *for derivations of the decision feedback equalizer.*

[Monsen TIT 71] P. Monsen, "Feedback equalization for fading dispersive channels," *IEEE Transactions on Information Theory*, IT-17,pp.56-64, January 1971

This paper attacks two problems. It first derives a MMSE joint feedforward and feedback equalizer solution assuming no past decision errors *and then it presents a blind adaptive algorithm (equivalent to DD-LMS) for adapting both forward and backward equalizers using past decisions. Monsen uses throughout all his derivations the assumption of no past decision errors, so that he can both derive the MMSE solution and analyze the convergence of his adaptive algorithm. There is no mention of algorithm initialization.*

[Price ICC 72] R. Price, "Nonlinearly feedback-equalized PAM vs. capacity from noisy filter channels," *Proceedings of IEEE International Conference on Communications*, Philadelphia, USA, 1972, pp.22/12-17

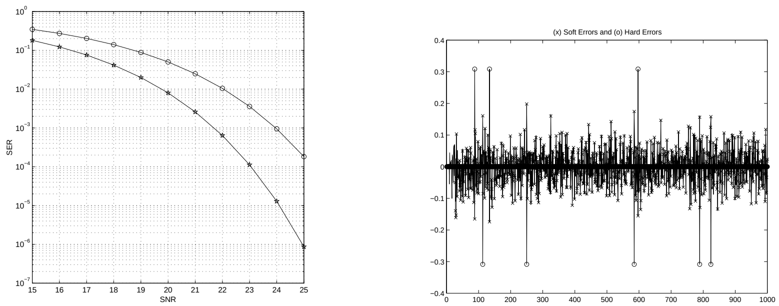

Figure 13: (left) SER versus range of SNRs near operation 20 dB point, the top curve corresponds to equalization of the channel and the bottom corresponds to a pure delay channel with equal power, (right) occurrence of decision errors (∘) and soft estimation errors (×) due to error propagation (real part).

This conference paper concentrates on analyzing the performance of a PAM communication system using feedforward and feedback equalization as a function of the channel capacity (information rate). It is a bit difficult to read, and requires some knowledge of rate distortion and Shannon channel capacity theory.

[Messerschmitt Bell 73] D. Messerschmitt, "A geometric theory of intrasymbol interference. Part I: Zero-forcing and decision-feedback equalization," *Bell System Technical Journal*, Vol.52, No.10, pp.1483-1519, November 1973

[Messerschmitt TCOM 78] D. Messerschmitt, "Minimum MSE equalization of digital fiber optic systems," *IEEE Transactions on Communications*, COM-26, pp.1110-1118, 1978

[Belfiore PROC 79] C.A. Belfiore and J.H. Park, Jr., "Decision feedback equalization," *Proceedings of the IEEE*, Vol. 67, pp.1143-1156, August 1979

Major survey on DFEs, looking at zero-forcing and MMSE for finite and infinite length realizations. Some discussion on error propagation.

[Leclert TCOM 85] A. Leclert and P. VanDamme, "Decision feedback equalization of dispersive radio channels," *IEEE Transactions on Communications*, Vol.33, pp.676-684, July 1985

Adaptive DFEs

[George TCOM 71] D. George, R. Bowen, and J. Storey, "An adaptive decision feedback equalizer," *IEEE Transactions on Communications*, Vol.COM-19, pp.281-292, June 1971

This paper develops an algorithm for the blind adaptation of cascaded feedforward and feedback equalizers using the assumption of no past decision errors. *The algorithm uses a zero forcing approach by cross-correlating the error with decisions. The paper compares the use of an adaptive forward equalizer versus the adaptive joint forward/backward equalizer for a typical telephone channel by plotting MSE during adaptation. There is no mention of how the equalizer taps are initialized.*

[Ling TCOM 85] F. Ling and J.G. Proakis, "Adaptive lattice decision-feedback equalizers – their performance and application to time-variant multipath channels," *IEEE Transactions on Communications*, Vol.33,pp.348-356, April 1985

[Kennedy IJACSP 93] R.A. Kennedy, B.D.O. Anderson, and R.R. Bitmead, "Blind adaptation of decision feedback equalizers: gross convergence properties," *International Journal of Adaptive Control and Signal Processing*, Vol.7, pp.497-523, 1993

This paper may serve as a fairly good introduction to adaptive DFEs. It analyzes the convergence behavior of adaptive DFEs, in specific DD-DFE, The paper shows the occurrence of bad local minima in DD-DFE and constructs a theory on delay-type minima. The discussion of the results is relatively easy to follow and simple examples illustrate these concepts.

[Castelein ICASSP 94] T.S. Castelein, Y. Bar-Ness, R. Prasad, "Blind linear recursive equalizer with decorrelation algorithm," *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, pp.1069-1072, Detroit, MI, May 1994

[Kamel MIL 94] R.E. Kamel, Y. Bar-Ness, "Error performance of the blind decision feedback equalizer using decorrelation," *IEEE MILCOM*, pp.618-622, Long Branch, NJ, October 1994

[Marcos ICASSP 95] S.Marcos, S.Cherif, M.Jaidane, "Blind cancellation of intersymbol interference in decision feedback equalizers," *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*,Detroit, MI, May 1995

The paper discusses the use of soft decisions for updating a blind decision-directed-type DFE algorithm. Examples use two tap channels.

[Papadias ASIL 95] C.B.Papadias and A.Paulraj, "Decision-feedback equalization and identification of linear channels using blind algorithms of the bussgang type," *Proceedings of the 29th Annual Asilomar Conference on Signals, Systems, and Computers*, November 1995

The paper suggests the use of CMA for the update of adaptive cascaded feedforward (baud) and feedback equalizers. Under the assumption that there are no past decision errors, *it proceeds to analyzing the structure as a multichannel problem, thus enabling the use of results guaranteeing the convergence of CMA to a perfect solution under certain length and zero conditions. This perfect solution, however, consists of the inverse of the first channel tap with the DFE matching the remaining taps, which may be extremely ill-behave when the first channel tap is small. An interesting suggestion on the trapping of the algorithm at bad minima is that a* normalized *algorithm with a larger step-size can be used to avoid ill-convergence.*

[Casas ASIL 95] R. Casas, Z. Ding, R.A. Kennedy, C.R.Johnson, Jr., and R. Malamut, "Blind adaptation of decision feedback equalizers based on the constant modulus algorithm," *Proceedings of the 29th Annual Asilomar Conference on Signals, Systems, and Computers*, November 1995

This paper finds a class of channels for which DD-DFE does not converge to the desired solution when initialized at the origin. The class is used to determine if CMA-DFE may have better convergence behavior than DD-DFE.

[Jinbiao ICSP 96] X. Jinbiao, M. Fulong, W. Yumin, "Soft decision-directed blind equalization," *Proceedings of the Internation Conference on Signal Processing*, pp.32-34, Beijing, China, October 1996

[Zhuang CISS 96] W. Zhuang, "Adaptive H_∞ filtering algorithm for decision feedback equalizer over dispersive fading channels," *Conference on Information Sciences and Systems*, pp.107-112 , Princeton, NJ, March 1996

[Kim ICCE 96] D.W. Kim, S.H. Han, M.S. Eun, J.S. Choi, Y.S. Cho, "An adaptive decision feedback equalizer using error feedback," *IEEE Transactions on Consumer Electronics*, Vol.42, pp.468-477, August 1996

[Tong ASIL 96] L. Tong, D. Liu, "On blind decision feedback equalization," *Proceedings of the 29th Annual Asilomar Conference on Signals, Systems, and Computers*, pp.xxx, November 1996

[Labat ICC 96] J. Labat, O. Macchi, C. Laot, N. Le Squin, "Is training of adaptive equalizers still useful?" *Proc. IEEE International Conference on Communications*, Dallas, TX, pp.968-972, June 1996

[Mottier ICC 97] D. Mottier, M. Hélard, J. Labat, "Self-adaptive decision feedback equalization: application to high-order QAM signal," *Proc. IEEE International Conference on Communications*, Montréal, Canada, pp.1100-1104, June 1997

[Casas SPAWC 97] R.A. Casas, C.R. Johnson, Jr., "On the blind adaptation of an FSE+DFE combination," *IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications*, pp.113-116, Paris, France, April 1997

This paper shows that an decision-directed adaptive FSE+DFE initialized from an open eye condition converges on average to the MMSE solution for noiseless channels. Simulations with noise are conducted that show that with a 3% decision error probability the decision-directed FSE+DFE will converge close to the MMSE solution.

[Kamel TCS 97] R.E. Kamel, Y. Bar-Ness, "Anchored blind equalization using the constant modulus algorithm," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, Vol.44, pp.397-402, May 1997

[Casas IJACSP 98] R.A. Casas, C.R. Johnson, Jr., R.A. Kenney, Z. Ding, R. Malamut, "Blind adaptive decision feedback equalization: a class of channels resulting in ill-convergence from a zero initialization," *International Journal of Adaptive Control and Signal Processing*, Vol. 12, pp.173-193, 1998

Adaptive IIR Equalization

[da Rocha ICASSP 94] C.A. Faria da Rocha, O. Macchi, "A novel self-learning adaptive recursive equalizer with unique optimum for QAM," *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, Adelaide, Australia, pp.481-484, April 1994

[da Rocha ITS 94] C.A. Faria da Rocha, O. Macchi, J.M.T. Romano, "An adaptive nonlinear IIR filter for self-learning equalization," *Proc. IEEE International Telecommunications Symposium*, Rio de Janeiro, Brazil, pp.6-10, August 1994

[da Rocha GLOBE 96] C.A. Faria da Rocha, J.M.T. Romano, "A unimodal adaptation criterion for a non-linear IIR self-learning equalizer," *Proc. IEEE Global Telecommunications Conference-Communications Theory Mini-Conference*, London, UK, pp.41-45, November 1996

[Cavalcanti SPAWC 97] F.R.P. Cavalcanti, J.C.M. Mota, "Predictive blind equalization with the constant modulus criterion," *Proc. IEEE Signal Processing Workshop on Signal Processing Advance in Wireless Communications*, Paris, France, pp.49-52, April 1997

[Tong ICASSP 97] L. Tong, D. Liu, "Blind predictive decision-feedback equalization via the constant modulus algorithm," *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, Munich, Germany, pp.3901-3904, May 1994

Error Propagation

[Duttweiler TIT 74] D.L.Duttweiler, J.E. Mazo, and D.G.Messerschmitt, "An upper bound to the error probability in decision-feedback equalization," *IEEE Transactions on Information Theory*, IT-20, pp.490-497, July 1974

[Cantoni TCOM 76] A.Cantoni and P.Butler, "Stability of decision feedback inverses," *IEEE Transactions on Communications*, Vol COM-24, pp.970-977, September 1976

[OReilly IEE 85] J.J. O'Reilly and A.M. de Oliveira Duarte, "Error propagation in decision feedback receivers," *Proceedings of the IEE*, pt. F, pp.561-566, December 1985

[Oliveira IEE 85] A.M. de Oliveira Duarte and J.J. O'Reilly, "Simplified technique for bounding error statistics for DFB receivers," *Proceedings of the IEE*, pt. F, No.7, pp.567-575, December 1985

The two papers above use the Markov chain model of the DFE to define error probabilities and time for recovery from a channel model. Application of the results is very limited due to complexity at higher dimensions.

[Kennedy ISSPA 87] R.A. Kennedy, B.D.O. Anderson and R.R. Bitmead, "Stochastic analysis of non-adaptive decision feedback equalizers," *Proceedings of ISSPA87*, pp.842-847, Brisbane, Australia, August 1987

[Kennedy TCOM 87a] R.A. Kennedy and B.D.O. Anderson "Error recovery of decision feedback equalizers on exponential impulse response channels," *IEEE Transactions on Communications*, vol. COM-35, pp.846-848, August 1987

[Kennedy TCOM 87b] R.A. Kennedy, and B.D.O. Anderson, "Recovery times of decision feedback equalizers on noiseless channels," *IEEE Transactions on Communications*, Vol.COM-35, No.10, October 1987, pp.1012-1021

[Kennedy TCOM 87c] 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, pp1022-1029, October 1987

[Kennedy TCOM 89a] R.A. Kennedy, B.D.O. Anderson, and R.R. Bitmead, "Channels leading to rapid error recovery for decision feedback equalizers," *IEEE Transactions on Communications*, pp. 1867-1875, November 1989

The paper uses passivity analysis to find a class of channels for which a DFE having past decision errors will flush out these errors in a finite time. The main result is that SPR channels do have this characteristic, but this is not a necessary condition for a finite recovery time.

[Altekar TIT 93] S.A.Altekar and N.C.Beaulieu, "Upper bounds to the error probability of decision-feedback equalization," *IEEE Transactions on Information Theory*, IT-39, pp.145-156, January 1993

[Beaulieu TCOM 94] N.C. Beaulieu, "Bounds on recovery times of decision feedback equalizers," *IEEE Transactions on Communications*, Vol.42, pp.2786-2794, October 1994

[Krueger ICC 94] D. Krueger, J.R. Cruz, "Upper bounds for binary and ternary decision feedback equalization of the magnetic recording channel," *Proceedings of the 1994 IEEE International Conference on Communications*, pp.1206-1210, New Orleans, LA, May 1994

[Beaulieu TIT 95] N.C. Beaulieu, W.W. Choy, "The autocorrelation of error propagation in a one-tap decision-feedback equalizer," *IEEE Transactions on Information Theory*, Vol. 41, pp.847-848, May 1995

[Smee VTC 95] J.E. Smee, N.C. Beaulieu, "New methods for evaluating equalizer error rate performance," *IEEE Vehicular Technology Conference*, pp.87-91, Chicago, IL, July 1995

[Choy GTC 96] W.W. Choy, N.C. Beaulieu, "New recursive method for estimating the error probability of decision feedback equalization," *IEEE Global Telecommunications Conference*, pp.51-56, London, UK, November 18-22, 1996

[Dogancay TIT 97] K. Dogancay, R.A. Kennedy, "Blind detection of equalization errors in communication systems," *IEEE Transactions on Information Theory*, Vol.43, pp.469-482, March 1997

[Fertner TSP 98] A. Fertner, "Improvement of bit-error-rate in decision feedback equalizer by preventing decision-error propagation," *IEEE Transactions on Signal Processing*, July 1998, pp.1872-1878

[Proakis SP 98] J.G. Proakis, "Equalization techniques for high-density magnetic recording," *IEEE Signal Processing Magazine*, July 1998, pp.73-82

DFE Design

[Falconer Bell 73] D. Falconer and G. Foschini, "Theory of minimum mean-square-error QAM systems employing decision feedback equalization," *Bell System Technical Journal*, Vol.52, No.10, December 1973

[Salz Bell 73] J. Salz, "Optimum mean-square decision feedback equalization," *Bell System Technical Journal*, Vol.52, no.8, October 1973

[Salazar Bell 74] A. Salazar, "Design of transmitter and receiver filters for decision feedback equalization," *Bell System Technical Journal*, Vol.53, No.3, March 1974

[Anderson ISSPA 87] B.D.O. Anderson, R.A. Kennedy, and R.R. Bitmead, "Decision feedback equalizers: concepts towards design guidelines," *Proceedings of ISSPA87*, pp.1-7, Brisbane, Australia, August 1987

[Bergmans Philips 87] J.W.M. Bergmans, A.J.E.M. Janssen, "Robust data equalization, fractional tap spacing and the Zak transform," *Philips Journal of Research*, Vol. 42, no.4, pp.351-398, 1987

[Abdulrahman JCOM 92] M. Adbulrahman and D.D.Falconer, "Cyclostationary crosstalk suppression by decision feedback equalization on digital subscriber loops," *IEEE Journal on Selected Areas in Communications*, Vol 10, No.3, April 1992, pp.640-649

[Lee Book 94] E.A. Lee, D.G. Messerschmitt, *Digital Communication*, Kluwer Academic Publishers, Boston, 1994

[Gozzo TCOM 95] F. Gozzo, J.B. Anderson, "The impact of white gaussian noise mismatch on linear and decision feedback equalizers," *IEEE Transactions on Communications*, Vol.43, pp.173-177, February-April 1995

[Rohrs GTC 95] C.E. Rohrs, "Decision feedback equalization and capacity on the linear stationary Gaussian channel," *IEEE Global Telecommunications Conference*, pp.871-873, Singapore, Singapore, November 1995

[Al-Dhahir TIT 95a] N. Al-Dhahir and J.M. Cioffi, "MMSE decision-feedback equalizers: finite-length results," *IEEE Transactions on Information Theory*, pp.961-975, July 1995

[Al-Dhahir TIT 95b] N. Al-Dhahir and J.M. Cioffi, "Fast computation of channel-estimate based equalizers in packet data transmission," *IEEE Transactions on Information Theory*, November 1995, pp.2462-2473

[Johnson ASIL 95] C.R.Johnson,Jr., H.J.Lee, J.P.LeBlanc, T.J.Endres,R.A.Casas,E.Tai,Z.Reznic, W.E.Meyer,F.López de Victoria,J.R.Treichler,I.Fijalkow, and Z.Ding "On fractionally-spaced equalizer design for microwave radio channels," *Proceedings of the 29th Annual Asilomar Conference on Signals, Systems, and Computers*, November 1995

The paper presents some empirical microwave channel data and derives a fractionally-spaced MMSE equalizer design equation. It discusses some issues as length and delay selection for design based on SER performance.

[Stojanovic TCOM 95] M. Stojanovic, J.G. Proakis, J.A. Catipovic, "Analysis of the impact of channel estimation errors on the performance of a decision-feedback equalizer in fading multipath channels," *IEEE Transactions on Communications*, Vol.43, pp.877-886, February-April 1995

[Berberidis GTC 95] K. Berberidis, J. Palicot, "Frequency-domain decision feedback equalizer for multipath echo cancellation," *IEEE Global Telecommunications Conference*, pp.98-102, Singapore, Singapore, November 14-16, 1995

[Cioffi TCOM 95] J.Cioffi, G. Dudevoir, M. Eyuboglu, and G.D. Forney, Jr., "Minimum mean-square-error decision feedback equalization and coding – Parts I and II," *IEEE Transactions on Communications*, Vol. 43, pp.2582-2604, October 1995

[Al-Dhahir TSP 96] N. Al-Dhahir, J.M. Cioffi, "Efficient computation of the delay-optimized finite-length MMSE-DFE," *IEEE Transactions on Signal Processing*, Vol. 44, pp.1288-1292, May 1996

[Al-Dhahir TIT 96] N. Al-Dhahir, J.M. Cioffi, "Block transmission over dispersive channels: transmit filter optimization and realization, and MMSE-DFE receiver performance," *IEEE Transactions on Information Theory*, Vol. 42, pp.137-160, January 1996

[Barry TIT 96] J.R. Barry, E.A. Lee, D.G. Messerschmitt, "Capacity penalty due to ideal zero-forcing decision-feedback equalization," *IEEE Transactions on Information Theory*, Vol. 42, pp.1062-1071, July 1996

[Voois TIT 96] P.A. Voois, I. Lee, J.M. Cioffi, "Effect of decision delay in finite-length decision feedback equalization," *IEEE Transactions on Information Theory*, Vol.42, pp.618-621, March 1996

[Hasnie ISSPA 96] S. Hasnie, R.A. Kennedy, R.A. Casas, C.R. Johnson, Jr., "Constrained MMSE approach to equalizer design," *Proceedings of ISSPA96*, pp.???, Brisbane, Australia, August 1996

This paper proposes a DFE design strategy which mixes MMSE linear equalizer parametrizations with MMSE-DFE parametrizations to yield a DFE design with low error propagation and better SER that MMSE-DFE.

[Shamai TIT 96] S. Shamai, R. Laroia, "The intersymbol interference channel: lower bounds on capacity and channel precoding loss," *IEEE Transactions on Information Theory*, Vol.42, pp. 1388-1404, September 1996

[Smee TCOM 97] J.E. Smee, N.C. Beaulieu, "On the equivalence of the simultaneous and separate MMSE optimizations of a DFE FFF and FBF," *IEEE Transactions on Communications*, Vol. 45, pp.156-158, February 1997

[McGinty tbd tbd] N. McGinty, S. Hasnie, R. Kennedy, "Quantizer design for fractionally spaced decision feedback equalization," tbd, tbd

[Al-Dhahir TSP 97] N. Al-Dhahir, "Mismatched finite-complexity MMSE decision feedback equalizers," *IEEE Transactions on Signal Processing*, Vol.45, pp.935-944, April 1997

[Al-Dhahir TCOM 97] N. Al-Dhahir, A.H. Sayed, J.M. Cioffi, "Stable pole-zero modeling of long FIR filters with application to the MMSE-DFE," *IEEE Transactions on Communications*, Vol. 45, pp.508-513, May 1997

DFE/Coding/RSSE

[Eyuboglu TCOM 88] M.V. Eyuboglu and S.U.H. Quereshi, "Reduced-state sequence estimation with the set partition and decision feedback," *IEEE Transactions on Communications*, Vol.36, pp.13-20, January 1988

[Duel-Hallen TCOM 89] A. Duel-Hallen and C. Heegard, "Delayed decision-feedback sequence estimation," *IEEE Transactions on Communications*, Vol.37, pp.428-436, May 1989

[Cioffi TCOM 90] J.M. Cioffi, P.H.Algoet and P.S. Chow, "Combined equalization and coding with finite-length decision feedback equalization," *IEEE Transactions on Communications*, Vol.37, pp.807.2.1-807.2.4, December 1990

[Williamson TCOM 92] D. Williamson, R.A. Kennedy, and G.W. Pulford, "Block decision feedback equalization," *IEEE Transactions on Communications*, Vol. 40, pp.255-264, February 1992

[Yellin TIT 97] D. Yellin, A. Vardy, O. Amrani, "Joint equalization and coding for intersymbol interference channels," *IEEE Transactions on Information Theory*, Vol. 43, pp.409-425, March 1997

[Bergmans TCOM 97] J.W.M. Bergmans, J.O. Voorman, H.W. Wong Lam, "Dual decision feedback equalizer," *IEEE Transactions on Communications*, Vol. 45, pp.514-518, May 1995

[Ariyavisitakul COM 98] S.L. Ariyavisitakul, G.M. Durant, "A broadband wireless packet technique based on coding, diversity, and equalization," *IEEE Communications Magazine*, July 1998, pp.110-115

Decision Directed Algorithm

[Mazo Bell 80] J.E. Mazo, "Analysis of decision-directed equalizer convergence," *Bell System Technical Journal*, 1980, pp. 1857-1876

Mazo calls to attention the fact that an adaptive linear forward equalizer using the decision directed algorithm may converge to a bad minimum. He explains how overall channel-equalizer parameter space is partitioned into polytopes or cones each containing a different cost function with a possible unique minimum. He shows these minimum for two tap and three tap parameter space using BPSK and 4-PAM. He also mentions the fact that for sufficiently large step-size the algorithm will leave bad minima in a finite amount of time.

[Kumar CDC 83] R. Kumar, "Convergence of a decision-directed adaptive equalizer," *Proceedings of the Conference on Decision and Control*, Long Beach, CA, 1983, IEEE New York, 1983, pp.1319-1324

A difficult paper to read. It discusses the idea of analyzing the decision directed algorithm (as done by Mazo) by adding noise and looking at what happens as the number of equalizer taps is increased progressively. Kumar argues that the algorithm is more stable this way (the problem of being traped at bad minima is not too dauntling.

[Macchi TIT 84] O.Macchi and E.Eweda, "Convergence analysis of self-adaptive equalizers," *IEEE Transactions on Information Theory*, Vol IT-30, No.2, March 1984, pp.161-176

[Kennedy TCOM 89b] R.A. Kennedy, G. W. Pulford, B.D.O. Anderson, and R.R. Bitmead, "When has a decision directed equalizer converged?," *IEEE Transactions on Communications*, 1989, pp.879-884

Miscellaneous

[Qureshi PROC 85] S.U.H. Qureshi, "Adaptive equalization," *Proceedings of the IEEE*, Vol.73, No.9, September 1985

Classic survey on adaptive equalization. A must for anyone doing research in the area.

[Haykin Book 96] S. Haykin, *Adaptive Filter Theory*, 3rd ed., Englewood Cliffs, NJ, Prentice-Hall, 1996

[Johnson PROC 98] C.R. Johnson, Jr., P. Schniter, T.J. Endres, J.D. Behm, D.R. Brown, R.A. Casas, "Blind equalization using the constant modulus criterion: a review,"