Channels Leading to Rapid Error Recovery for Decision Feedback Equalizers: Passivity Analysis

R.A. Kennedy, B.D.O. Anderson, and R.R. Bitmead

Department of Systems Engineering, Research School of Physical Sciences
Australian National University, G.P.O. Box 4, Canberra, A.C.T. 2601, AUSTRALIA.

(* Research supported by the Australian Telecommunications and Electronics Research Board)

Abstract

When a decision feedback equalizer is used on a channel satisfying a simple passivity constraint (equivalently expressible in terms of gain-phase constraints) the error recovery time is bounded, and thus recovery is rapid, regardless of the initial conditions and the particular data sequence. This class of channels includes cases of practical interest and identifies channels for which a decision feedback equalizer is a practical option.

1. Introduction

Decision feedback equalizers (DFEs) are simple hardware devices designed to cancel intersymbol interference (ISI) generated by a distorting channel. However, the major problem with their non-adaptive operation is error propagation [1]. Its presence means that DFE operation in practice may be unsatisfactory, in the sense that the time for the DFE to recover from any error condition may be unacceptably long [2,3]. In fact, for some finite impulse response (FIR) channels of length $N$ the mean error recovery time may be of order $2^N$ data periods (even for some which are minimum phase or near minimum phase), which evidently is totally impractical. It then becomes a problem to identify stronger hypotheses on the channel model for which the error recovery time is sufficiently short, as judged by practical standards. For these channels we can say then that a DFE is a practical option. In our work we find a broad and robust class of channels for which the error recovery time is finite. As such we are defining a class of channels suffering from significant ISI for which a DFE may be effectively used.

In the literature there has been very little written about the error recovery properties of DFEs. In fact only in [2,5,6] has it been indicated theoretically that there are some non-trivial channels for which the DFE operates satisfactorily. In contrast, the two prominent early references analyzing error propagation in DFEs [1,4] both give bounds which are too conservative for practical use. In [1,4] the given bounds on recovery time and error probability actually correspond to the worst realizability channel models as was demonstrated in [2,3]. We note here the work in [7,8] which strives to reduce these bounds given explicit, i.e., specific, knowledge of the channels. In contrast, here we give a broad general condition on the channel parameters—specifically the transfer function is positive real of a certain degree—to ensure good DFE error recovery performance.

The major sections are organized as follows. In §2 we define the finite error recovery time problem. In §3 we give the background passivity theory. In §4 we give our main theorem. We also include four applications, including analysis of a real channel. In §5 we establish convergence rates and explicit bounds given an exponential overlap on the channel impulse response. In §6 we give the result for M-array data and relate the error recovery time bound back to the binary case. Also for high SNR channels satisfying a passivity constraint we give a formula for the error probability. Finally in §7 we give a discussion.

2. Problem Formulation and Definitions

A communication channel and general non-adaptive decision feedback equaliser (DFE) are shown in Fig. 1. The communication channel is modelled as a linear, time-invariant filter with impulse response,

$$ h = \{ h_0, h_1, h_2, \ldots \} \quad (2.1) $$

of possibly infinite dimension. This channel is driven by an input binary sequence $\{ a_n \}$, where $k$ is the discrete time index. No statistical model of $\{ a_n \}$ is assumed nor needed. The M-array $\{ a_n \}$ case will also be treated in a later section. We note that in a more general context $h$ could be thought of as the cascade (convolution) of the linear channel and a linear equalizer preceding the DFE.

The distorted output of the linear channel is $b_k$ and is assumed noiseless. By studying the noiseless case we are creating a pointer to the important practical situation of a high SNR channel. (In a later section we will introduce an additive noise signal into the analysis but only treat the asymptotic case as the noise variance tends to zero.) At the receiving end we have a DFE consisting of a tapped delay line with impulse response

$$ d = \{ 0, d_1, d_2, \ldots \} \quad (2.2) $$

fed by a binary output decision sequence $\{ \hat{a}_n \}$ as described by Fig.1.

The algebraic formulation of the system depicted in Fig. 1 is given by

$$ \begin{align*}
 h_0 + h_1 q^{-1} + h_2 q^{-2} + \cdots & \rightarrow b_k \\
 \Sigma & \rightarrow \text{sgn}(\cdot) \\
 d_1 q^{-1} + d_2 q^{-2} + \cdots & \rightarrow \hat{a}_k
\end{align*} $$
a strictly causal convolutional operator $H$ which maps the error sequence $e(t, k \geq 0)$ to the ISI sequence $r(t, k \geq 0)$ in accordance with (2.4b), i.e., $r(t) = H(s) * h(t) \in h(t)$ where $h(t) \equiv h(0, h_1, h_2, \ldots)$ (which differs from (2.1)). The lower block $L$ in Fig. 2 consists of two parts. The first is a stochastic multiplier, to account for the stochastic input $a_k$ defined by,

$$m_k \equiv \begin{cases} 1 & \text{if } a_k = -\text{sgn}(r_k); \\ 0 & \text{if } a_k = +\text{sgn}(r_k), \end{cases}$$

(3.1)

whose function is clear from Lemma 1(ii), i.e., if $m_k = 0 \Rightarrow a_k = \text{sgn}(r_k) \Rightarrow e_k = 0$. Otherwise $m_k$ does nothing, i.e., $m_k$ takes the value unity (3.1). The second part of the lower block is a time-invariant non-linearity which maps $\{r_k, m_k\}$ into the sequence $e \equiv -r$. Note whenever the input $m_k r_k$ is less in magnitude than $h_k$ the output $e_k = -r_k$ is zero (as in Lemma 1(ii)). Otherwise the output conforms to Lemma 1(ii). Note that in this block the stochastic multiplier and the non-linearity may be commuted.

![Fig.2 Error Propagation Block Diagram.](image)

A significant observation we make concerning the lower block $L$ in Fig. 2 is that it preserves the sign of the input and therefore is a passive operator in the circuit theoretic sense [9]. Our first task will be to transform the system in Fig. 2 such that the upper block $H$ becomes a strictly passive operator whilst the lower block $L$ remains passive. Then we utilize some standard results from input-output stability to show the DFE has a (quantifiable) finite error recovery time.

### 3.2 Definitions and Passivity Theorem

We begin with some definitions which are standard in input-output stability theory [9]. We focus on a Hilbert space structure composed of real valued sequences indexed by $k \in \mathbb{Z}_+^N$ (non-negative integers). Then if we have two sequences $x \equiv \{x_0, x_1, \ldots\}$ and $y \equiv \{y_0, y_1, \ldots\}$ their inner product will be defined as

$$\langle x, y \rangle \equiv \sum_{i=0}^{\infty} x_i y_i.$$  

(3.2)

where it is clear that $\langle x, y \rangle = \langle y, x \rangle$. This inner product (3.2) induces a natural euclidean norm defined by

$$\|x\| \equiv \langle x, x \rangle^{1/2} = \left( \sum_{i=0}^{\infty} x_i^2 \right)^{1/2}.$$  

(3.3)

We define the discrete function space $l^2$ which consists of all sequences satisfying

$$x \in l^2 \iff \|x\| < \infty.$$  

(3.4)

Similarly we have the space $l^1$ which consists of all sequences satisfying $x \in l^1 \iff \|x\|_1 \equiv \sum_{i=0}^{\infty} |x_i| < \infty$. The space $l^1$ is generally too restrictive an arena for deriving results, so we introduce the standard concept of an extended space $l^2_\infty$ [9], defined by

$$x \in l^2_\infty \iff \|x\|_{l^2} < \infty, \forall \in \mathbb{Z}_+^N.$$  

(3.5)

where $\|x\|_{l^2} \equiv \|P_T x\| < \infty$, $\forall \in \mathbb{Z}_+^N$ defined by

$$(P_T x)(k) \equiv \begin{cases} x_k, & \text{if } k \leq T; \\ 0, & \text{if } k > T. \end{cases}$$

Note (3.5) just says that $x \in l^2_\infty$ if $\|x\|_{l^2} < \infty \forall \in$, i.e., $x$ does not have a finite escape time. So for example if $x \equiv \{x_k = k, \forall k \in \mathbb{Z}_+^N\}$ then $x \in l^2_\infty$ but clearly $x \notin l^2$.

From definitions (3.4) and (3.5) it is apparent that $l^2 \subset l^2_\infty$. In our work all signals considered will lie in the extended space $l^2_\infty$ (because we stipulate only that $h \in l^2$). However it is of great interest to show that particular signals also lie in the subset $l^2$. For example with the error signal, it is our aim to show $e \in l^2$. Then because $e_k \in \{-2, 0, +2\}$ we have the following fundamental observation.

$$e \in l^2 \iff e_k = 0, \forall k \geq K \quad K < \infty.$$  

(3.6)

i.e., the DFE has recovered from error at time $K$.

Now define $\|x\|_\infty \equiv \|P_T x\|$, $\|x\|_{l^2} \equiv \|x, y\|_{l^2}$ such that

$$\langle Hz, z \rangle_{l^2} \geq \beta, \forall z \in l^2 \forall T \in \mathbb{Z}_+^N.$$  

(3.7)

If $H$ were linear then $\beta$ could be taken as zero.

### Definition:

An operator $H: l^2 \rightarrow l^2$ is passive if there exists a constant $\beta > 0$ such that

$$\langle Hz, z \rangle_{l^2} \geq \beta, \forall z \in l^2 \forall T \in \mathbb{Z}_+^N.$$  

(3.8)

Again if $H$ were linear then $\beta$ could be taken as zero. We label $\beta$ as the degree of passivity.

As an example of passivity (but not strict passivity), which will be important later, let us check the claim at the end of §3.1 concerning the lower block $L$ of Fig. 2. Suppose $x \in l^2$ is the input to an operator $H$ with output $y \equiv Hy$, which satisfies $y \geq 0, \forall k \in \mathbb{Z}_+$ (a sign preserving operator). Then trivially

$$\langle Hx, x \rangle_{l^2} = \sum_{k=0}^{T} y_k x_k \geq 0, \forall x \in l^2 \forall T \in \mathbb{Z}_+.$$  

(3.9)

showing $H$ is passive according to definition (3.7) with $\beta = 0$. That is, if $H$ is a non-linearity confined to the first and third quadrants then it is passive (even if it is time-varying or has memory).

Our second example which we state as a lemma will be important later and relates to the definition of strict passivity (3.8) applied to linear operators.

### Lemma 2:

Suppose $G: l^2 \rightarrow l^2$ is defined by $Gu = g \otimes u$, where $g \equiv \{g_0, g_1, \ldots\} \in l^1$. Let $\delta > 0$. Then $Vu \in l^2$

$$\|Gu, u\|_{l^2} \geq \|u, u\|_{l^2} + \delta, \forall u \in [0, 2\pi]$$

(3.10)

where $\|g\|_{l^2} = \sum_{i=0}^{\infty} g_i r^{-i}$ is the Z-transform of the impulse sequence $g$.

The proof follows from Parseval's Theorem (see [9]). Lemma 2 says that a linear convolutional operator is strictly passive if its Nyquist plot belongs to $\{r \in C: Re(r) \geq \delta\}$.

![Fig.3 Passivity Theorem Block Diagram.](image)

We now come to the main passivity theorem, Figure 3 defines the signals and operators of interest. In it $e$ and $u$ are the input sequences to the operators $H_1$ and $H_2$ and $y = H_2 e$ are the respective output sequences. There is a single external signal $u$. All signals shown are assumed to lie in $l^2$. The following theorem and proof are an adaptation of a more general result in [9, p.182].

### (Passivity) Theorem 3:

Suppose: (i) Operator $H_1$ is linear and strictly passive, i.e.,

$$\langle H_1 e, e \rangle_{l^2} \geq \delta \|e, e\|_{l^2}^2, \forall e \in l^2 \forall T \in \mathbb{Z}_+^N.$$  

(3.11)

where $\delta_1 > 0$, and (ii) operator $H_2$ is a non-linearity confined to the first and third quadrant, implying

$$\langle H_2 v, v \rangle_{l^2} \geq 0, \forall v \in l^2 \forall T \in \mathbb{Z}_+^N.$$  

(3.12)

by (3.7), and is thus passive. Then $\forall u \in l^2 \Rightarrow e \in l^2$.

**Proof:** We show $e \in l^2$ by determining upper and lower bounds on the quantity $\langle H_1 e, e \rangle_{l^2} + \langle H_2 v, v \rangle_{l^2}$. First we determine a lower bound. Using (3.11) and (3.12) we clearly have

$$\langle H_1 e, e \rangle_{l^2} + \langle H_2 v, v \rangle_{l^2} \geq \delta_1 \|e, e\|_{l^2}^2 \forall T \in \mathbb{Z}_+^N.$$  

(3.13)

where, recall, $\delta_1 > 0$ is the constant associated with the degree of passivity of the $H_1$ operator. An upper bound on (3.13) follows from the following simple calculation, see Fig. 3.

$$\langle H_1 e, e \rangle_{l^2} + \langle H_2 v, v \rangle_{l^2} = \langle H_1 e, e \rangle_{l^2} + \langle -e, e \rangle_{l^2} = \langle -e, e - H_1 e \rangle_{l^2} = \langle -e, e \rangle_{l^2} \leq \|e, e\|_{l^2} \forall T \in \mathbb{Z}_+^N.$$  

(3.14)
where the last line is an application of the Cauchy-Schwartz inequality. Then combining (4.13) with (4.14) we obtain \( \|e_T\| \leq \delta_1 \|u_T\| \), \( \forall T \in \mathbb{Z}_+ \) whenever \( \|e_T\| > 0 \). Letting \( T \to \infty \) we find

\[
\|e\| \leq \delta_1 \|u\|
\]  
(3.15)
i.e., \( u \in l_2 \Rightarrow e \in l_2 \) as desired.

4. Sufficient Conditions for a Finite Recovery Time

In this section we transform the system in Fig.2 so that we may apply the general passivity theorem of the last section. This involves two steps. The first step is to apply a loop transformation because \( \mathcal{H} \) (Fig.2) is non-passive. The second step is to model the effect of initial conditions at time \( k = 0 \), i.e., an initial (arbitrary) error state, by an external signal \( u \) as in the passivity theorem.

We apply a loop transformation \( [9] \) to the system in Fig.2 to obtain the new system shown in Fig.4. Note that the effect of the newly introduced feedforward and feedback paths with gains \( \mathcal{H}_f \) is to cancel exactly. The upper block labelled \( \mathcal{H}_f \) has impulse response given by

\[
(\mathcal{H}_f, h_1, h_2, \ldots)
\]  
(4.1)
where \( \mathcal{H}_f \) is a finite gain associated with the feedforward path. For the passivity theorem to apply we need (4.1) strictly passive, i.e., \( \mathcal{H}_f \) sufficiently passive and we have available Lemma 2 as a test.

In the lower block labelled \( \mathcal{H}_s \), which includes the positive feedback of gain \( \mathcal{H}_f \), we need to be concerned that we have not destroyed the passivity of the original lower block (Fig.3). The following lemma with proof now applies. The symbol definitions are given in Fig.4.

**Lemma 4:** If \( 0 \leq h \leq h_f \) then \( \mathcal{H}_s \) (Fig.4) is passive.

**Proof:** The \( \mathcal{H}_s \) block has input \( v \) and output \( z \in \{-2,0,2\} \). We attempt to show \( v_t z_t \geq 0, \forall \mathcal{V} \) which ensures passivity. For \( v_t \) the input to the sector non-linearity within the \( \mathcal{H}_s \) block is given by \( v_t = m_2(z_t + h z_2) \) from which we have after multiplying through by \( z_t \),

\[
m_2 z_t = (v_t - m_2 h z_2) z_t, \quad \forall \mathcal{V} \in \mathbb{Z}_+
\]  
(4.2)
We have three cases (see Fig.4): (i) \( z_t = 2 \Rightarrow m_2 = 1 \) and \( w_t \leq h_0 \), which implies from (4.2) that \( v_t z_t = (v_t - 2 h_0 z_t) z_t \geq 0 \), given \( 0 \leq h \leq h_f \), i.e., \( v_t z_t \geq 0 \); (ii) \( z_t = 0 \Rightarrow m_2 = 1 \) and \( w_t \leq h_0 \) leading to \( z_t \geq 0 \) by symmetry; and (iii) \( z_t = 0 \) which gives \( v_t z_t = 0 \) because \( v_t \in \{0,1,2\} \), i.e., \( v_t z_t \leq 0 \). Thus \( u_t z_t \geq 0 \), \( \forall \mathcal{V} \) in every case.

Another condition which needs to be fulfilled in Theorem 3 is \( u \in l_2 \). This condition will necessitate some hypothesis on the channel \( h \) to be fulfilled. The signal \( u \) for our application will model the effects of initial conditions in the \( \mathcal{H}_s \) block since all our sequences are defined only for \( k = 0 \), whereas the real system may have been operating from the distant past, i.e., \( k = -\infty \).

Note that this signal \( u \), as shown in Fig.4, is unaffected by the introduction of \( \mathcal{H}_f \). From Fig.2 we superpose on the upper \( \mathcal{H}_f \) linear operator of impulse response \( (h_0, h_1, h_2, \ldots) \) to represent the effects of arbitrary initial conditions, i.e., an arbitrary initial internal state via the signal

\[
u_0 = \sum_{k=-\infty}^{\infty} h_k e_{k-t}, \quad k \in \mathbb{Z}_+
\]  
(4.3)
where values \( e_{-2}, e_{-1}, e_0, e_1, \ldots \), taking values in \( \{-2,0,2\} \) define the initial state at time \( k = 0 \). To ensure \( u \) to be in \( l_2 \) we impose some sufficient conditions on the channel \( h \).

**Lemma 5:** Suppose \( \mathcal{H}_s \in l_2 \) satisfies \( |h_m| = O(m^{-\alpha}) \) as \( m \to -\infty \) where \( \alpha \) is constant. Then:

(i) \( \eta > 1 \Rightarrow h \in l_1 \),

(ii) \( \eta > 0.5 \Rightarrow u \in l_2 \).

**Proof:** (i) is elementary. (ii) It is easy to show from (4.3) that \( |u_0| \leq 2 \sum_{\mathcal{K} \in \mathbb{K}} |h_k| = O((k+1)^{\alpha-1}) = \infty \) as \( k \to -\infty \), using continuous approximations to the summations. Then \( p_2 \leq \alpha \leq O(k^{-\alpha+1}) \) as \( k \to -\infty \). However, \( u \in l_2 \) iff \( p \leq l_2 \). Using (4.4a) on this implies \( 2 \eta - 2 > 1 \), i.e., \( \eta > 0.5 \).
Theorem 6 gives no indication of the maximum time slgoal the 8gn(zt) We take the multiplier
The signals
transfor&. of
maps
strictly passive
Lemma 5, the only sensible condition for stability takes the following form:

This type
DFE
Consider Fig.8 which shows the
channel
Fig.7
is shown in
ia
Let
H;:et
> 0

The meaning attached to the quantity
is provided
provided \(|\gamma| < 1\), i.e., \(||u\| < \infty\). Thus with an exponential overbound of the channel and \(|\gamma| < 1\), Theorem 3 applies to the starred system in Fig.8 and we conclude from (5.15) that

\[
\|u^*\| \leq \sum_{k=0}^{\infty} |e^k| \leq \gamma^k = \frac{1}{1-\gamma^2} \leq 1
\]

i.e., \(e^k \in l_2\) (provided \(|\gamma| < 1\)). This provides an exponential rate of decay on \(e_k = \rho^k e_1 \leq \rho^k \|u\|\). However, \(e_k\) is restricted to the set \([-2, 0, +2]\) and therefore must be zero after some time \(K(p)\) in \(l_2\), which is the least integer satisfying

\[
K(p) \geq \frac{\log((1-\gamma)/(1-\gamma^2))}{\log(\rho)}
\]

i.e., the least integer \(K(p)\) in \(l_2\) such that,

\[
K(p) \geq \frac{\log((1-\gamma)/(1-\gamma^2))}{\log(\rho)}
\]

Table 1: Error Recovery Time Bounds

<table>
<thead>
<tr>
<th>Analysis Technique</th>
<th>(\gamma = 0.50)</th>
<th>(\gamma = 0.81)</th>
<th>(\gamma = 0.95)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passivity Theory</td>
<td>8</td>
<td>43</td>
<td>258</td>
</tr>
<tr>
<td>Exponential Results</td>
<td>6</td>
<td>11</td>
<td>71</td>
</tr>
<tr>
<td>Markov Processes</td>
<td>4</td>
<td>4094</td>
<td>5 \times 10^{21}</td>
</tr>
</tbody>
</table>

These bounds are conservative by the nature of the analysis. In [6] but only for exponential channels it is shown that the tight bounds on the maximum error recovery time are 2, 11 and 71, respectively. It is interesting to compare both sets of bounds (see the first two rows of Table 1) with mean error recovery time bounds which can be deduced from the DFE literature based on Markov Processes ([4, 6, 8]). Of these being statistical bounds we need a statistical model of the input sequence \(a_n\)—an independent, equiprobable binary distribution being standard. This does not invalidate the comparison because the error recovery time bound \(K(p)\) in (5.7) always overbounds the true mean error recovery time.

To compute the mean error recovery time bounds based on the work in [1] we define an effective channel length \(n\) for the exponential channel. This is given by the minimum \(n\) such that

\[
2 \sum_{i=0}^{n} \gamma_i = 2^{n+1} \frac{1}{1-\gamma} < 1.
\]

The meaning attached to the quantity \(n\) is simply that the DFE needs to make \(n\) consecutive correct decisions to recover from any error state with \(a_{n-1} \neq a_{n} \) (being the present instant of time). Now for the worst case channels implicitly considered in [1-4, subject to (5.9), the probability of making an error is precisely \(1/2\) for any decision before recovery (i.e., before \(n\) consecutive correct decisions have been made). By the theory of success runs [4] the mean recovery time is given by 2(2^n-1). Looking at our three examples we have: (i) \(\gamma = 0.50\) implying \(n = 2\) and thus a mean recovery time of 6, (ii) \(\gamma = 0.81\) implying \(n = 11\) and thus a mean recovery time of 4094, and (iii) \(\gamma = 0.95\) implying \(n = 71\) and thus a mean recovery time of 5 \times 10^{21}. These three bounds are displayed in the third row of Table 1.
6. Some Generalizations

6.1 Error Recovery Under Imperfect Equalization

This subsection represents a threefold generalization of the previous results. These modifications involve, in part, relaxation of some of the previous assumptions regarding the model of the system under study. The generalizations are as follows: (i) the DFE tapped delay line is assumed to be FIR of length $N$ rather than IIR, whilst the channel may be IIR; (ii) the assumptions that $d_i = h_i, \forall i > 1$ is relaxed to a condition which stipulates the $d_i$ are sufficiently close but not necessarily equal to some ideal values, and (iii) the results are generalized to the situation where error-free behaviour is characterized by $A_k = \text{sgn}(h_i)\alpha_{k-1}, \forall k \geq K$ for some fixed delay $i \in \{0, 1, \ldots, N\}$ rather than $A_k = \alpha_k, \forall k \geq K$.

All these generalizations will be treated in parallel. A key feature of the analysis performed in this subsection is showing explicitly the close relationship between eye diagrams and rates of error recovery.

As a motivation to studying delay-type behaviour, we mention that a DFE has its taps adapted blindly, i.e., without a training sequence. In this case, it was shown in [13] that the DFE taps may adapt not only to an (ideal) equilibrium where $d_i = h_i, i \in \{1, 2, \ldots, N\}$ but also to a delay equilibrium where $d_i = \text{sgn}(h_i)\alpha_{i-1}, i \in \{1, 2, \ldots, N\}$ provided certain conditions are met. We will show that when in the vicinity of a delay equilibrium, after some finite time $K$, all decisions will be of the form $A_k = \text{sgn}(h_i)\alpha_{k-1}, \forall k \geq K$, hence the terminology.

To analyze non-ideal behaviour we take (2.2) and set $d_i = 0$ for $i > N$, i.e., the tapped delay line is FIR of length $N$ rather than IIR. Define $\sigma_k = \text{sgn}(h_i)$. We can decompose (2.2) as follows:

$$A_k \equiv \text{sgn}(\sum_{i=1}^{\infty} h_i a_{i-1} - \sum_{i=1}^{\infty} d_i a_{i-1})$$

(6.1a)

$$= \text{sgn}(h_k a_{k-1} + r_k(\xi) + s_k(\xi) + t_k(\xi))$$

(6.1b)

where

$$r_k(\xi) \equiv \sum_{i=1}^{\infty} h_i a_{i-1} - \sum_{i=1}^{\infty} h_i a_{i-1}$$

(6.1c)

and

$$s_k(\xi) \equiv \sum_{i=1}^{\infty} h_i a_{i-1} + \sum_{i=1}^{\infty} h_i a_{i-1}$$

(6.1d)

In (6.1): (i) $r_k(\xi)$ acts as the basic residual ISI term (note if we let $\delta = 0$ and $N \to \infty$ then (6.1c) becomes (2.4b)); (ii) $s_k(\xi)$ is a term which generally gets smaller as the taps $d_1, d_2, \ldots, d_N$ approach the $\delta$-delay equilibrium at $a_1(h_k, a_2, \ldots, a_N)$, and includes any precursors; and (iii) $t_k(\xi)$ is that part of the tail of the channel which cannot be modelled by the DFE after the tapped delay line is FIR.

Beginning with $t_k(\xi)$ in (6.1c), it is clear that we need

$$t_k(\xi) \leq \sum_{i=1}^{\infty} |h_i| \leq \Phi, \forall k \in Z_2, \xi$$

(6.2)

with $\Phi$ sufficiently small else the DFE problem is not well-posed, i.e., $N$, the number of DFE taps, needs to be chosen large enough in the first place so that the DFE can effectively cancel the ISI.

Now when in the vicinity of a delay equilibrium we claim $A_k = \sigma_k a_{k-1}, \forall k \geq K$ provided certain conditions are met, which we now determine. Define new (delay) errors

$$e_k(\xi) \equiv a_{k-1} - \sigma_k A_k$$

(6.3)

then the basic residual ISI term $r_k(\xi)$ (6.1c) may be written

$$r_k(\xi) \equiv \sum_{i=1}^{\infty} h_i a_{i-1}$$

(6.4)

and will be zero whenever we make $N$ consecutive correct $\delta$-delay decisions.

Now suppose

$$A_{\text{min}}(\delta) \equiv |h_k - \sum_{i=0}^{\infty} |h_i| - \sum_{i=0}^{\infty} |h_i - \sigma_k a_{i-1}| - \sum_{i=0}^{\infty} |h_i| > 0$$

(6.5)

Then some perusal will show that whenever $N$ consecutive correct decisions are made, all future decisions will be correct (in the delay sense) because $r_k(\xi) = 0$ and $h_k$ is larger in magnitude than $s_k(\xi) + t_k(\xi)$ can ever be seen to (6.1b). This defines a new form of error recovery, i.e., (6.5) is a sufficient condition for all decisions to be (delay) correct whenever $N$ consecutive $\delta$-delay decisions have been made. Note that if all decisions are to be of the form $A_k = \sigma_k a_{k-1}$ for all input sequences, given $N$ consecutive correct decisions have been made, then condition (6.5) is also a necessary condition (see [13] which treats a similar problem).

Define $A_k(\delta) \equiv (h_k a_{k-1} + s_k(\delta) + t_k(\delta))\sigma_k a_{k-1},$ noting that (6.5) we have $A_k(\delta) > A_{\text{min}}(\delta) > 0.$ From (6.1b), $A_k = \text{sgn}(a_k a_{k-1} + A_k(\delta) + t_k(\delta));$ then clearly the analogue of Lemma 1 is:

**Lemma 8:** Suppose condition (6.5) holds. Then

(i) $|r_k| < A_k(\delta)$ or $a_k = a_{k-1}$ $\Rightarrow A_k = a_k a_{k-1}.$

(ii) $|r_k| > A_k(\delta)$ and $a_k = a_{k-1} =-\sigma_k \text{sgn}(r_k) = A_k = -\sigma_k a_{k-1}.$

Thus we have the picture in Fig.9 which differs marginally from Fig.2. Note the lower block is sector bounded within the 1st and 3rd quadrants whilst $A_{\text{min}}(\delta) > 0.$

![Fig.9 Imperfect Equalization Error Propagation](image_url)

**Imperfect Equalization** Theorem 9: Suppose the parameters of a binary linear channel $\mathbf{h} = (h_0, h_1, \ldots)$ and the DFE tapped delay line parameters $d = (d_1, d_2, \ldots)$ satisfy (6.5) for some (at most one) delay $\delta \in \{0, 1, \ldots, N\}$ and $A_k \equiv \text{sgn}(h_k).$ Further, suppose $\Xi \geq 0$ such that

$$\frac{\Delta_{\text{min}}(\delta)}{2} + \sigma_k \sum_{i=1}^{N} h_i a_{i-1} \cos(\theta) > \Xi, \forall \theta \in [0, 2\pi].$$

Given a non-ideal DFE output sequence $\{a_k\}$ generated through (6.1a), then for some $\Xi < \infty,$ we have $A_k = \sigma_k a_{k-1}, \forall k \geq K.$

**Remarks:**

(i) Note condition (6.5) stipulates that the $d_i$ need to be sufficiently close to the $\sigma_k a_{k-1}$ (in an $L_1$-norm sense) if a certain operator is to be strictly passive. Note that the worse the mismatch, the less $A_{\text{min}}(\delta)$ will be. This forms a convenient geometrical picture to replace the messy algebra.

(ii) Note $A_{\text{min}}(\delta)$ may be interpreted precisely as the amount that a certain eye diagram is open after recovery. Thus the wider the post-recovery eye can be, the more rapid one can expect recovery to be. Theorem 9 is saying that given the eye is initially closed (an arbitrary error state) it will always open after at most $K.$

6.2 M-ary Results

The theory developed for binary systems can be extended to larger alphabets where $M$ symbols are used. We outline some of the important differences for brevity we restrict attention to zero delay systems. Let $\{a_k\} \in \{1, M, 3, M, 2M, \ldots, M-1\}$ where $M$ is positive and even. The standard decision function $Q_M(z)$ which replaces $\text{sgn}(z)$ in the binary analysis is defined by

$$Q_M(z) \equiv \sum_{k=1-M/2}^{M/2-1} \text{sgn}(z + 2k).$$

The M-ary version of (2.4a), where we have ideal equalization, becomes

$$A_k = Q_M(h_k a_{k-1} + r_k), \quad h_0 > 0$$

(6.7)

where $r_k$ is as in (2.4b) with the exception that $\sigma_k = \sigma_k a_{k-1}$ takes values in $\{0, 2, \ldots, 2M-1\}$. Now suppose we had no residual ISI, i.e., $r_k = 0$, then (6.7) reduces to

$$A_k = Q_M(h_k a_{k-1})$$

(6.8)

from which it is clear that (with $M > 2$) we need $h_0 \approx 1$ for error-free behaviour. (This differs from the binary case, $M = 2$, where it was only necessary that $h_0 > 0).$ Elaborating, we have (without proof)

**Lemma 10:** Given $e_k = 0, \forall k \in Z_2,$ and $M \geq 4$ even, then

$$A_k = a_k, \quad A_k \equiv M - 2 < h_0 < M - 2 \quad \text{M-3}$$

(6.9)

So, in summary, we require the right-hand condition in (6.9) to be in force if the M-ary error recovery process is to be well-posed. (Note, with an obvious redefinition of $r_k$ an analogous result for $h_0 < 0$ can be generated.)

Consider the error propagation mechanism for the well-posed M-ary problem. We now verify that an operator $L$ which maps $z \mapsto -z$ (the residual ISI to the negative of the errors) is passive, indeed sector bounded. The proof is omitted but may be found in [10].
Lemma 11: The operator \( L : r \mapsto e \) is sector bounded according to

\[
0 \leq \frac{r_1}{r_2 \Delta'} \leq \frac{2}{\Delta'(h_0)} \tag{6.10}
\]

where

\[
\Delta'(h_0) \triangleq (M - 1)h_0 - (M - 2) \quad \text{if } h_0 \leq 1;
\]

\[
- (M - 3)h_0 + (M - 2) \quad \text{if } h_0 \geq 1
\]

provided \( h_0 \) lies in the interval \((0, 1)\).

With Lemma 12 we may again invoke the passivity theorem to get the \( M \)-ary analogue of Theorem 6.

(M-ary) Theorem 12: Suppose \( a_o \in \{1 - M, 3 - M, \ldots, M - 1\} \) is the input to a linear channel \( h = (h_0, h_1, \ldots) \) which satisfies the right hand side of (6.9) to be well-posed and \( |h_m| = O(m^{1-1}) \) as \( m \to \infty \) where \( c > 0 \).

With \( \Delta'(h_0) \) as in (6.11), suppose \( \Delta > 0 \) such that

\[
\Delta'(h_0) \geq \frac{\Delta}{2} \sum_{m=1}^\infty h_m \cos(md) \geq \delta, \quad \forall h \in [0, 2x].
\]

Given an ideal DFE output sequence \( a_k \) generated through (6.7), then for some \( K < \infty \), we have \( a_k = a_k, \forall k \geq K \).

Remarks:

(i) Theorems relating the rates of convergence and robustness for the \( M \)-ary case can be generated by analogy with the binary case.

(ii) The error recovery rate is most rapid with \( h_0 = 1 \) which implies \( \Delta'(h_0) = 1 \) because this makes (6.12) the most strictly passive, which is in accord with intuition. If \( h_0 \) differs from 1 there will be a diminishing of passivity and hence a drop in the rate of error recovery. This highlights the role that gain compensation plays in the \( M \)-ary case.

(iii) Normalized channels, where \( h_0 = 1 \), which result in a finite recovery time for binary symbols will also have a finite recovery time for the \( M \)-ary case because then conditions (6.12) and (4.5) are identical. Letting \( K_M(p) \) denote the error recovery time bound for the \( M \)-ary case, analogously to (5.7), then this is related to the binary \( K(p) \) via \( K_M(p) = K(p) + \log_2(M - 1) \).

To prove this, note that in a calculation which mimics (5.6) the factor of 4 is replaced by \( 4(M - 1)^2 \).

6.3 Noise and Asymptotic Error Probability Bounds

In [3] it is shown how the mean error recovery time is related to the error probability in the most important case of a high SNR channel. To calculate an error probability bound we include additive channel noise with variance \( \sigma^2 \) into the analysis, and following [1] we define the fully open eye error probability as \( \epsilon \triangleq P(X \neq a_k | r_k = 0) \) where \( r_k \) is the residual ISI (2.4b), and \( \epsilon \triangleq O(\sigma^2) \) (Chebyshev's Inequality, [3]). We can then use the techniques in [3] to bound the stationary error probability \( P_E \triangleq P(r_k \neq a_k) \) for channels satisfying the conditions (5.2) and (5.3), via

\[
P_E < \frac{\epsilon}{2}(K(p) + 2) \quad \text{as } \sigma^2 \to 0,
\]

where \( K(p) \) is the passive analysis error recovery time bound which appears in (5.7). From Table 1, bound (6.16) may be anything up to a factor of 10^8 tighter than the oft-cited result in [1] which says \( P_E \leq \epsilon^2 \) (effectively derived by replacing \( K(p) \) by \( 2\epsilon^2 - 1 \) the worst case implicit in [4]).

7. Discussion

Up until now there has been scant theoretical justification that non-trivial, non-adaptive DFEs behave satisfactorily because of passivity propagation, perhaps only [8] and [6] being relevant. This is in stark contrast with the purported popularity of DFEs in practice. Previous theoretical work [1,4,7,8] concentrated on bounds which turn out to be hopelessly conservative in the majority of cases. These latter bounds will not be improved without relying heavily on explicit knowledge of the channel to be equalized—this was emphasized in [2,3] and [6]. In this paper we have determined some non-trivial broad classes of channels for which a DFE can be effectively used (the results in [5] and [6] are very narrow and are subsumed by our present analysis). This class includes channels which have near exponential impulse responses, thus are capable of modelling twisted pair cable [12]. This provides some theoretical justification to the (controlled) use of DFEs in practice.

As well as defining a non-trivial class of channels for which the DFE behaves satisfactorily, the passivity analysis appears to provide an opportunity to clarify the role and function of a DFE. Recently the intuition that sensibly the DFE can only be used on minimum phase channels was shown to be misguided [2]. In [2] it is highlighted that minimum phase or near minimum phase nature of \( h = (h_0, h_1, h_2, \ldots) \) is not enough to imply satisfactory DFE error recovery. In comparison we have shown that the stronger notion of strict passivity of the object (or its generalizations) \( \{ h_0, h_1, h_2, \ldots \} \) is a concept which leads to a sensible equalization problem (for both binary and \( M \)-ary alphabets). Naturally strict passivity implies minimum phase-ness. If the channel fails the passivity condition (for some delay \( \delta \)) it is our contention that a linear equalizer preceding the DFE must be used. Note that our analysis covers the case of a cascade of a linear equalizer with a DFE because we can interpret \( h \) as being not just the channel impulse response but alternatively as the convolution of the channel impulse response with the linear equalizer. We interpret the function of the linear equalizer as being to transform the channel into a passive object which aligns well with the intuition that the linear equalizer is needed to remove precursor ISI.

Acknowledgements

The assistance of Monsieur Graham Pulford in the derivation leading to the \( M \)-ary extension was greatly appreciated.

References


