SOLVED AND UNSOLVED PROBLEMS IN ADAPTIVE BLIND EQUALIZATION

Brian DO Anderson

Research School of Information Sciences and Engineering
The Australian National University
Canberra ACT 0200

ABSTRACT

This paper presents some results and conjectures dealing with channel equalization. For several possible equalizers, the following question is considered: if the equalizer output sequence has identical statistics to the input signal sequence, has the equalizer correctly converged?

1. THE PROBLEM DOMAIN

Many communication signal processing problems envisage a transmitted sequence $a_0\ a_1, \ldots$ drawn from a finite signal alphabet, with passage of that sequence through a linear channel which produces intersymbol interference in the channel output sequence, call it $\{b_k\}$. There is then possible addition of noise, followed by processing of b_k to try to undo the effects of the channel, through a device termed an equalizer. The equalizer output $\{\hat{a}_k\}$ is supposed to reconstruct the signal alphabet, *eg* $\hat{a}_k = a_{k-\delta}$ for some fixed integer δ and all k.

Many equalizers need to be adaptive, and often also blind, *ie* not even a training sequence of known $\{a_k\}$ is provided to allow equalization. In the blind case, adaptive equalizers rely either on being close to correct to begin with (*eg* open eye diagram) or computing certain statistics which are used to guide changes in equalizer parameters, such as in Godard blind equalizers [1] and decision directed equalizers [2].

When the parameters or taps in a blind adaptive equalizer settle down, *ie* cease making gross movements and remain in the neighbourhood of some point, it is reasonable to ask the question: has convergence occurred to a correct point. Such questions can be subtle. For example, for the Godard equalizers, with an infinite number of taps, the existence of incorrect equilibria was demonstrated in [1], but they were shown to be unstable (and thus irrelevant) in [3]. However, with finite equalizer length, stable undesirable equilibria were shown to exist in [4] (and with an 11 year interval from [1] to [4]).

In this paper we study the issue of establishing whether convergence for certain equalizers has occurred, given that the equalizer output sequence has the same statistics as the signal sequence. The equalizers are decision directed and decision feedback equalizers, embodying linear system processing together with quantization.

2. DECISION DIRECTED EQUALIZERS

The behaviour of decision directed equalizers after adaptation can often be described in the following way. The channel input sequence $\{a_k\}$ is a sequence of independent random variables taking values in an M-ary alphabet, with equal probability. We shall, for convenience, assume M=2, and $a_k \in \{\pm 1\}$. The channel impulse response is $\{d_0, d_1 \ldots\ldots\}$. In the absence of noise, the channel output is

$$b_k = \sum_{l=0}^{\infty} d_l a_{k-l} \tag{2.1}$$

This signal drives the equalizer, which is a finite impulse response filter with impulse response $(f_0, f_1 \ldots, f_L)$, followed by a quantizer. The quantizer input sequence is

$$c_k = \sum_{j=0}^{L} f_j b_{k-j} \tag{2.2}$$

and the quantizer output is

$$\hat{a}_k = Q_M(c_k) \tag{2.3}$$

where Q_M is an M-ary quantizer.

With M=2, correct operation of the equalizer is achieved when

$$\hat{a}_k = a_{k-\delta} \text{ or } \hat{a}_k = -a_{k-\delta} \tag{2.4}$$

for some fixed δ and all k. In an adaptive equalizer, the coefficients f_i are tunable. In a correctly operating equalizer, *ie* after satisfactory tuning, the convolution $h := d * f$ is an impulse response with a dominant term h_δ, *ie*

$$|h\delta| \geq \sum_{i \neq \delta} |h_i| \tag{2.5}$$

In a correctly operating equalizer, it is obvious that the equalizer output must be a sequence of independent random variables, taking values ± 1 with equal probability.

The key converse question is: if the sequence $\hat{a}_k$ is a sequence of independent random variables, taking values ±1 with equal probability, is it necessarily a delayed version of a_k, possibly with inversion, ie. does (2.4) hold?

2.1 FIR Channels

The converse question has been answered in the affirmative for the case of FIR channels, thus $d_l = 0$ for $l > K$, K being some finite positive integer. In this case, treated in [5], the combined channel-equalizer impulse response is necessarily also FIR, and, with $N = K + L$,

$$\hat{a}_k = sgn(h_0 a_k + h_1 a_{k-1} + \ldots\ldots + h_N a_{k-N}) \tag{2.6}$$

Without reproducing the proof in [5], we record here some salient aspects. First, define as inessential any impulse expanse coefficient h_i with the property

$$\text{sgn}\Big(\sum_{j=0}^{N} h_j a_{k-j}\Big) = \text{sgn}\Big(\sum_{\substack{j=0 \\ j \neq i}}^{N} h_j a_{k-j}\Big) \quad \forall \{a_k\} \tag{2.7}$$

Delete any inessential h_i starting with h_0 and working on, and starting with h_N and working back. We conclude that

$$\hat{a}_k = sgn(h_\alpha a_{k-\alpha} + \ldots\ldots + h_\beta a_{k-\beta}) \tag{2.8}$$

for $0 \leq \alpha \leq \beta \leq N$. In effect, $h_0, \ldots, h_{\alpha-1}$ and $h_{\beta+1}, \ldots, h_N$ can be thought of as negligible precursors and postcursors of the impulse response. Suppose now that $\hat{a}_k$ is white, but that $\alpha \neq \beta$, so that $\hat{a}_k$ is not $\pm a_{k-\delta}$ for some δ and all k. Next, study $\hat{a}_k$ and $\hat{a}_{k-(\beta-\alpha)}$, which depend respectively on $a_{k-\alpha}, a_{k-\alpha-1}, \ldots\ldots, a_{k-\beta}$ and on $a_{k-\beta}, a_{k-\beta-1}, \ldots\ldots, a_{k-2\beta+\alpha}$. Since the a_k sequence is independent and identically distributed (iid), $\hat{a}_k$ and $\hat{a}_{k-(\beta-\alpha)}$ are conditionally independent given $a_{k-\beta}$. It is not then hard to show that

$$p(\hat{a}_k > 0 / \hat{a}_{k-(\beta-\alpha)} > 0 = p(\hat{a}_k > 0 / a_{k-\beta} > 0) p(\hat{a}_{k-(\beta-\alpha)} > 0 / a_{k-\beta} > 0) + p(\hat{a}_k > 0 / a_{k-\beta} < 0) \; p(\hat{a}_{k-(\beta-\alpha)} > 0 / a_{k-\beta} < 0) \tag{2.9}$$

Given that h_α and h_β both fail to be inessential, $\hat{a}_k$ can be influenced by $a_{k-\beta}$, and $\hat{a}_{k-(\beta-\alpha)}$ can also be influenced by $a_{k-\beta}$. Thus $p(\hat{a}_k > 0 / h_\beta a_{k-\beta} > 0) = 1/2 + \delta_1$ for some $\delta_1 > 0$ and $p(\hat{a}_{k-\beta} > 0 / h_\beta a_{k-\beta} > 0) = 1/2 + sgn(h_\alpha h_\beta)\delta_2$ for some $\delta_2 > 0$. We have then

$$\begin{aligned} p(\hat{a}_h > 0 / \hat{a}_{k-(\beta-\alpha)} > 0) &= (1/2 + \delta_1)(1/2 + \delta_2 sgn(h_\alpha h_\beta)) \\ &+ (1/2 - \delta_1)(1/2 - \delta_2 sgn(h_\alpha h_\beta)) \\ &= 1/2 + 2\delta_1 \delta_2 sgn(h_\alpha h_\beta) \\ &\neq 1/2 \end{aligned} \tag{2.10}$$

and lack of independence is demonstrated. Hence, if the $\hat{a}_k$ sequence is iid, there must hold $\alpha = \beta$,
ie.

$$\operatorname{sgn}(h_\alpha a_{k-\alpha}) = \operatorname{sgn}(\sum_{j=0}^{N} h_j a_{k-j}) \quad \forall\{a_k\} \qquad (2.11)$$

ie

$$|h_\alpha| > \sum_{\substack{j=0 \\ j\neq\alpha}}^{N} |h_j|$$

In fact, it is enough to know that the $\hat{a}_k$ are uncorrelated to be able to conclude correct equalizer operation.

2.2 Non FIR Channels

It is a long-established custom in considering communications speech processing problems over linear channels to model all such channels on FIR channels where this is convenient, on the grounds that there is no essential loss of generality, in view of the presumed decay of the impulse response coefficients in an IIR channel to zero. Indeed, [5] simply says "We shall assume that $h_j = 0$ for $j > N_1$. (Extensions to the infinite impulse response case could also be conducted)". For the problem at hand of inferring correctness of equalization from matching of statistics, this extension is indeed far from trivial, and in fact remains (to the best of the author's knowledge) unestablished.

We can however offer some insights. Let $\{h_0, h_1, \ldots\ldots.\}$ denote the combined impulse response of the IIR channel and FIR linear equalizer. In case for some N, there holds

$$\inf\|[\pm h_0, \pm h_1, \ldots\ldots. \pm h_N]\| > \sum_{i>N} |h_i|$$

the post cursors beyond h_N are negligible (since $\operatorname{sgn}\sum_{i=0}^{N} h_i a_{k-i} = \operatorname{sgn}\sum^{\infty} h_i a_{k-i}$ for all sequences a_k). In this case, we are effectively back with the FIR problem. In general, one cannot expect such behaviour of the impulse response. A channel with $h_i = \alpha\beta^i$, β=0.6, for example does not offer such behaviour. It is then fairly straightforward to check that there is no simple variant on the technique of Section 2.1.

It may be possible to work with a sequence of FIR problems, thus for the J-th sequence member one would assume

$\hat{a}^J_k = sgn(h_0 a_k + h_1 a_{k-1} + h_J a_{k-J})$

and study $p(\hat{a}^J_k > 0 \mid \hat{a}^J_{k-\beta} > 0)$ for arbitrary but fixed β. Then one lets $J \to \infty$. Certainly, the distribution function of $\hat{a}^J_k$ approaches that of $\hat{a}_k = \operatorname{sgn}\sum_{i=0}^{\infty} h_i a_{k-i}$ under reasonable conditions on the impulse response *eg.* $\sum |h_i| < \infty$, [6]. It is highly likely that

$$p(\hat{a}^J_k > 0 | \hat{a}^J_{k-\beta} > 0) \to p(\hat{a}_k > 0 | \hat{a}_{k-\beta} > 0) \text{ as J} \to \infty$$

However, understanding how the probabilities $p(\hat{a}^J_k > 0 | \hat{a}^J_{k-\beta} > 0)$ vary with J is difficult, except in restricted situations such as when all h_i have the same sign.

3. DECISION FEEDBACK EQUALIZER (DFE)

For the channel of (2.1), an alternative form of equalizer is provided by the feedback equation:

$$\hat{a}_k = \operatorname{sgn}(b_k - \sum_{i=1}^{N} f_i \hat{a}_{k-i}) \qquad (3.1)$$

This equation and (2.1) imply that

$$\hat{a}_k = \operatorname{sgn}(d_o a_k + \sum_{i=1}^{N} d_i a_{k-l} - \sum_{i=1}^{N} f_i \hat{a}_{k-i} + \sum_{N+1}^{\infty} d_i a_{k-i}) \qquad (3.1a)$$

If $f_i = d_i$ for $l = 1,\ldots.,N$, if $\hat{a}_i = a_i$ for $l = 0,1,\ldots,N-1$, and if $\sum_{N+1}^{\infty} |d_i|$ is small in relation to $|d_o|$, then $\hat{a}_k = (\text{sgn } d_0) a_k$ for all $k \geq N$, indicating successful equalizer operation. The feedback cancels out much of the intersymbol interference, provided that initial conditions are satisfactory.

Variations on (2.1) to accommodate precursors are also possible. Thus if d_0 is very small and d_1 is large, with $|d_1| > |d_o| + \sum_{N+1}^{\infty} |d_i|$ one could look to

$$\hat{a}_k = \operatorname{sgn}\left(d_1 a_{k-1} + \sum_{i=2}^{N} d_i a_{k-i} - \sum_{i=2}^{N} f_i \hat{a}_{k-i} + d_0 a_k + \sum_{N+1}^{\infty} d_i a_{k-i}\right)$$

and with $d_i = f_i$ for $i=2,\ldots\ldots,N$, $\hat{a}_i = a_i$ for $i=0,1,\ldots.N-2$ there would hold $\hat{a}_k = (sgn\ d_1)\ a_{k-1}$ for all k. (One performance issue, not considered here, involves characterising performance when initial conditions are not matched, *ie* understanding transient performance).

The issue we want to consider in this paper is: if $a_k\{\pm 1\}$ and is iid, and if $\hat{a}_k$ is observed to obey $p(\hat{a}_k = 1) = {}^1/_2$, and $\hat{a}_k$ is iid or at least uncorrelated, is it true that $\hat{a}_k = \pm a_{k-\delta}$ for some δ and all k?

There is no complete answer to this question, but partial results can be given.

One can think of a DFE as a finite state Markov process with state X_k determined by $(a_{k-1}, \ldots, a_{k-N}, \hat{a}_{k-1}, \ldots, \hat{a}_{k-N})$ and each component can take on two values. There is of course a substantial theory, depending significantly on the theory of stochastic matrices, to allow conclusions to be drawn about such systems. However, it is important to understand that, almost certainly, this theory will not be enough. To see why, consider the (non DFE) system

$\hat{a}_k = a_k a_{k-1}$

where $a_k \in \{\pm 1\}$ and is iid. Obviously, $\hat{a}_k$ will be iid and one can model the set-up with a finite state Markov process. Nevertheless, manifestly one will not secure $\hat{a}_k = \pm a_{k-\delta}$ for some δ and all k. The linear blocks within the finite state Markov structure evidently play a significant role.

For the case

$$\hat{a}_k = sgn(h_0 a_k + h_1 a_{k-1} - d_1 \hat{a}_{k-1}) \quad (3.2)$$

it certainly is true that uncorrelatedness of $\hat{a}_k$ implies $\hat{a}_k = sgn\ h_\delta\ a_{k-\delta}$ for δ=0 or 1. How may this be seen? Since we are working with a finite state Markov process, with state vector $(a_{k-1}, \hat{a}_{k-1})$, it turns out there is only a finite number of patterns of state transition that can ever by accommodated with the structure of (3.2). More precisely, the behaviour of (3.2) is necessarily identical with its behaviour for one of the sets of values in the following table

Table 1: Classes of behaviour of finite state Markov process of (3.2)

h_0	h_1	d_1	Steady state behaviour
1	0	0	$\hat{a}_k = a_k$
1	2	0	$\hat{a}_k = a_{k-1}$
1	1	1	$\hat{a}_k = a_k$
1	0	2	$\hat{a}_k = -\hat{a}_{k-1}$
1	-1	1	$Pr(\hat{a}_k = 1 \mid \hat{a}_{k-1} = 1) = 1/4$
1	-2	0	$\hat{a}_k = -a_{k-1}$
1	-1	-1	$\hat{a}_k = a_k$
1	0	-2	$\hat{a}_k = \hat{a}_{k-1}$
1	1	-1	$Pr(\hat{a}_k = -1 \mid \hat{a}_{k-1} = 1) = 1/4$

The behaviour for each set of values is quickly established. Either $\hat{a}_k = sgn\ h_\delta\ a_{k-\delta}$, or $\hat{a}_k$ and $\hat{a}_{k-1}$ are correlated. The above argument is due to Kennedy [7].

In principle, this methodology can be used to resolve the question of whether uncorrelatedness or whiteness implies correct equalization for an equalizer of any specified complexity. It is possible also that an induction argument may be available to establish this conclusion for all DFEs of arbitrary complexity, but no such argument is known to have been constructed.

A further unpublished result due to the author is that for

$$\hat{a}_k = \mathrm{sgn}(h_0 a_k + h_j a_{k-j} - \sum_{l=1}^{m} d_i \hat{a}_{k-i}) \quad (3.3)$$

with $a_k \in \pm 1$, a_k iid, there holds $\hat{a}_k = sgn\ h_0\ a_k$ or $\hat{a}_k = sgn\ hj\ ak_{-j}$ if and only if $\hat{a}_k$ is white. These two cases correspond respectively to

$$|h_0| > \sum_{\substack{i=1 \\ i \neq j}}^{m} |d_i| + |h_j - d_j| \quad (3.4)$$

and

$$|h_j| > |h_0| + \sum_{i=1}^{m} |d_i| \quad (3.5)$$

What about the general case where (3.2) holds? The answer is unknown. No counter example has been found.

4. CONNECTION BETWEEN DECISION DIRECTED AND DECISION FEEDBACK EQUALIZATION

Reference was made in the previous section to the issue of (transient) convergence of DFEs, see *eg* [8, 9]. With a view to obtaining improved transient convergent properties, some variants on the DFE structure have been suggested, see [10]. One of these variants is the following:

$$\tilde{a}_k = b_k - \sum_{i=1}^{N} f_i a_{k-i} \tag{4.1}$$

$$\hat{a}_k = \operatorname{sgn} \tilde{a}_k \tag{4.2}$$

Compare this with (3.1). In (3.1), the feedback comes after the quantization and in (4.1) and (4.2), the quantization comes after the feedback. One would think (3.1) would be better, since the range of $\hat{a}_k$ is identical with that of a_k and this is not so for $\tilde{a}_k$. Nevertheless, for convergence from arbitrary initial conditions (and in the presence of dislodgement by a noise spike from correct equalization), (4.1) and (4.2) are more attractive. Now (4.1) together with (2.1) indicates that

$$\tilde{a}_k = [1 + f_1 z^{-1} + \ldots\ldots + f_N z^{-N}]^{-1}[d_0 + d_1 z^{-1} + \ldots\ldots]a_k \tag{4.3}$$

$$\hat{a}_k = sgn\ \tilde{a}_k \tag{4.4}$$

and we see that this set up is precisely the same as the IIR decision directed problem, where now

$$\sum h_i z^{-i} = [1 + \sum_{i=1}^{N} f_i z^{-i}]^{-1} \sum_{j=0}^{\infty} d_j z^{-j} \tag{4.5}$$

Even if the channel is genuinely FIR, we have an IIR problem.

5. REFERENCES

[1] D N Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems", <u>IEEE Transactions on Communications,</u> vol COM-28, (1980), pp 1867 – 1875

[2] J E Mazo, "Analysis of decision –directed equalizer convergence", <u>Bell Systems Technical Journal</u>, vol 59, 1980, pp 1857 – 1876

[3] G J Foschini, "Equalization without altering or detecting data", <u>AT & T Technical Journal</u>, vol 64, 1985, pp 1885 – 1911

[4] Z Ding, R A Kennedy, BDO Anderson and C R Johnson, "Ill Convergence of Godard Blind Equalizers in Data Communication Systems", <u>IEEE Transactions on Communications</u>, vol 39, 1991, pp 1313 – 1327

[5] R A Kennedy, G Pulford, BDO Anderson and R R Bitmead, "When has a Decision –Directed Equalizer Converged?". <u>IEEE Transactions on Communications</u>, vol 37, 1989, pp 879 – 884

[6] A M Garsia. "Arithmetic Properties of Benoulli Convolutions", <u>Transactions of the American Mathematical Society</u>, vol 91, 1962, pp 409 – 432

[7] R A Kennedy, "Operational Aspects of Decision Feedback Equalizers", PhD Thesis, Australian National University, 1988

[8] R A Kennedy, BDO Anderson and R R Bitmead, "Channels leading to rapid error recovery for decision feedback equalizers", <u>IEEE Transactions on Communications</u>, vol 37, 1989, pp 1146 – 1155

[9] J J O'Reilly and A M de Oliveria Duarte, "Error Propagation in Decision Directed Feedback Equalizers", <u>IEE Proceedings F. Communications, Radar and Signal Processing</u>, vol 132, 1985, pp 567 – 575

[10] R A Kennedy, Z Ding and S Hasnie, "Quantizer Optimization: Application of Passivity in Telecommunications", in <u>Model Identification and Adaptive Control</u>, G.C. Goodwin ed, Springer, London, 2001