Adaptive Decision Feedback Equalization Under Parallel Adaptation

Wee Gin Lim
Faculty of Engineering and Computer Science
University of Nottingham Malaysia Campus
Jalan Broga, 43500 Semenyih, Selangor
Email: lim.wee-gin@nottingham.edu.my

R. A. Kennedy
RSISE
Australian National University
0200 ACT Australia
Rodney.Kennedy@anu.edu.au

T. D. Abhayapala
RSISE
Australian National University
0200 ACT Australia
Thushara.Abbayapala@anu.edu.au

Abstract—In this paper, we first propose an alternative decision feedback equalizer (DFE) scheme to the DFE scheme of Labat, Macchi and Laot. Both schemes are broken into two distinct operation modes which feature a linear equalizer during acquisition and a DFE after the initial convergence has been achieved. While avoiding some potential problems with the latter DFE, our scheme achieves similar overall performance in terms of convergence speed and steady state error. A second contribution is the development of a strategy that simultaneously adapts the linear equalizer and the DFE in a parallel fashion such that the switch-over between modes is no longer necessary. Consequently, smoother and usually faster convergence is achieved as switching disruptions are reduced. This scheme is especially advantageous under noisy and severe conditions as supported by our simulations.

Index Terms—Adaptive equalization, blind decision feedback equalization, smooth switching, dual-mode algorithm.

I. INTRODUCTION

Blind equalization compensates for channels distortions without relying on a training sequence. Relative to adaptation using training sequences, blind equalizers tend to exhibit slower convergence, higher steady state errors and possibly ill-convergence. In the domain of linear equalization the blind algorithm design and analysis problem has been exhaustively studied for several decades [1]. In the domain of decision feedback equalization, it was only in 1998 that Labat, Macchi and Laot found a compelling realization of an effective blind decision feedback equalizer (DFE) [2]. In this paper we build on their work to develop an alternative design which exhibits significant performance advantages.

In the innovative blind scheme presented in [2], the adaptation is broken up into two modes, an acquisition mode and a tracking mode. Not only is the adaptation algorithm switched between modes but also the actual equalization filtering structure is switched. In the acquisition mode a recursive linear structure is employed and later, once the eye diagram has opened sufficiently, the DFE is employed in the tracking mode. The algorithm in the acquisition mode uses a combination of a constrained convex minimum energy cost [3] and the constant modulus algorithm (CMA) [4], [5], whilst in the tracking mode a decision directed (DD) algorithm is used. These choices are not arbitrary but are guided by the design objective to have the blind DFE converge as close as possible to the optimum MMSE solution [2].

In our design we seek to improve on the design in [2], while retaining the simplicity of the equalization scheme, with the following goals in mind:

A) Non-Recursive Linear Adaptation – In [2], the recursive form of the linear filter in the acquisition structure is used to permit direct transfer of the filter tap values to the DFE feedback filter in the tracking structure. Our design goal is to retain the ability of direct transfer of filter tap values but avoid some of the well-known problems with the adaptation of the recursive form (where the gradient can only be approximated, the dynamic range at the output may be large and the stability needs to be monitored). We show that this is indeed possible using a non-recursive linear filter provided we modify the DFE feedback filter appropriately.

B) Reduced Switching Transients – Switching structures and switching algorithms lead to transients in signals which disrupts convergence. The design goal here is to develop a strategy which provides for “smooth switching” and thereby significantly improved convergence. We achieve this on the structural side by employing a parallel adaptation strategy where the acquisition structure and the tracking structure are jointly adapted. Further on the algorithm side, rather than abruptly switch algorithms, we employ a technique to smoothly combine the acquisition and tracking algorithms into a single algorithm. In this way it is not necessary to make a distinction between acquisition and tracking and, therefore, the new scheme needs not be considered a dual-mode equalizer. This eliminates the requirement for the user to pre-specify certain parameter values, that strictly depend on the channel, to determine the exact and appropriate sampling instant to perform the switch-over.

To accomplish our first goal, we develop new equalizer structures in the acquisition and tracking modes as described in Section III. As the underlying concepts of our alternative scheme are identical to that of [2], simulation results would subsequently show almost identical performance between both schemes. Our second goal features a novel strategy that enhances these fast-convergence DFE schemes (that include
the scheme in [2] and our alternative scheme in Section III) where they must undergo a switch in both their adaptation algorithms and filtering structures. This new switching strategy that results in a single-mode DFE by combining the respective equalizers in both modes in a parallel fashion is described in Section IV.

II. SYSTEM MODEL

Consider a typical baseband symbol-rate blind equalizer whose main objective is to recover a corrupted version of the transmitted signal, \(\{a(k)\}\), based only on the observable received signals, \(\{r(k)\}\). Let the linear distortive channel coefficients and additive white Gaussian noise be denoted as \(\{h_i\}\) and \(n(r)\), respectively. Then the objective of equalization is to yield an output, \(z(k)\), that well estimates the transmitted signal such that \(z(k) \approx a(k-\Delta)\), where \(\Delta\) is some time delay. We will treat the channel \(\{h_i\}\), via the channel decomposition property [6], as a cascade of an equivalent minimum phase channel and an all-pass channel for the design of our linear equalizer in the starting mode in an identical fashion to [2], [7]. Thus, we separate the principle tasks of our equalizer into compensation for amplitude distortions only (due to the equivalent minimum phase channel), and compensation for the remaining phase distortions [2], [8].

The equalization scheme involves two distinct modes – the linear acquisition mode and the tracking DFE mode. In order to distinguish the filter parameters and signals of these two modes, we append the superscripts “(1)” and “(2)” to all filter parameters and signals in the “first” acquisition mode and “second” tracking mode, respectively.

III. DEVELOPMENT OF ALTERNATIVE FAST-CONVERGENCE DFE SCHEME

Following [2], a possible equalizer setup in the acquisition mode is the cascade of four filters, namely, a real gain control filter (\(\mathcal{GC}\)), a whitening filter (\(\mathcal{W}\)), a phase equalizer (\(\mathcal{T}\)), and a complex phase rotator (\(\mathcal{PR}\)). Thus in acquisition mode the equalizer is linear and decision feedback is not employed. This overall linear equalizer consisting of filters \(\mathcal{GC}, \mathcal{W}, \mathcal{T}, \text{and } \mathcal{PR}\) is shown in Fig. 1.

In contrast to [2] for the reasons listed in the introduction, we propose to use a non-recursive whitening filter in the place of the originally proposed recursive filter. This requires a modification to the whitening algorithm. The whitening algorithm, also known as the minimum output energy (MOE) algorithm, minimizes the cost function \(E\{|e^{(1)}(k)|^2\}\), where \(e^{(1)}(k)\) is the output of \(\mathcal{W}\). In order to avoid the convergence of the equalizer to the trivial solution, it is necessary to impose a linear constraint upon the equalizer taps so that convexity of the cost function is not lost. Following [3], the simplest and recommended linear constraint fixes the leading tap, \(a_0(k)\), to be unity for all \(k\) so that the transfer function of \(\mathcal{W}\) becomes \(1 + A(z)\) as shown in Fig. 1. Define \(a^{(1)}(k) = [a_1(k), a_2(k), \cdots, a_N(k)]^T\) as the weight vector of \(\mathcal{W}\) of length \(N\) that excludes the leading tap, \(a_0(k)\).

Then under the usual stochastic gradient adaptation, \(a^{(1)}(k)\) is updated according to

\[
a^{(1)}(k + 1) = a^{(1)}(k) - \mu_{a^{(1)}} E_{MOE}(k)s^{(1)}(k - 1)
\]

where the MOE error takes the form

\[
E_{MOE}(k) = t^{(1)}(k)
\]

and \(a^{(1)}(0) = [0, \cdots, 0]^T\). \(\mu_{a^{(1)}}\) is the adaptation step size, \(s^{(1)}(k - 1) = [s^{(1)}(k-1), s^{(1)}(k-2), \cdots, s^{(1)}(k-N)]^T\) and \(\ast\) is the complex conjugate operator. This algorithm attempts to minimize the cost \(J_{MOE}(a^{(1)}) = E\{|t^{(1)}(k)|^2\}\) subject to the first tap being fixed at unity.

As for \(\mathcal{GC}, \mathcal{T}\) and \(\mathcal{PR}\), our adaptation algorithms are no different from that of [2]. Briefly, \(\mathcal{GC}\) is a real filter that fixes the average power level of the samples at \(t^{(1)}(k)\) at a particular value, while \(\mathcal{PR}\) is the complex phase rotator that compensates for any demodulation phase errors. As for the phase equalizer, \(\mathcal{T}\), the CMA is the preferred algorithm to achieve the desired removal of the residual ISI. The update equations of \(\mathcal{GC}, \mathcal{T}\) and \(\mathcal{PR}\) are as follows:

- For \(\mathcal{GC}\):
  \[
  G^{(1)}(k + 1) = G^{(1)}(k) + \mu_{G^{(1)}} (1 - |u^{(1)}(k)|^2) \quad (3a)
  \]
  \[
  g^{(1)}(k + 1) = \sqrt{|G^{(1)}(k + 1)|} \quad (3b)
  \]
  where \(G^{(1)}(0) = 1\), and \(\mu_{G^{(1)}}\) is the positive step size that governs the rate of adaptation of \(\mathcal{GC}\). The output of \(\mathcal{GC}\) is \(s^{(1)}(k) = g^{(1)}(k)t^{(1)}(k)\).

- For \(\mathcal{T}\):
  \[
  b^{(1)}(k + 1) = b^{(1)}(k) - \mu_{b^{(1)}} E_{CMA}(k)t^{(1)}(k) \ast \quad (4)
  \]
  where the CMA error takes the form
  \[
  E_{CMA}(k) = u^{(1)}(k)[|u^{(1)}(k)|^2 - R] \quad (5)
  \]
  and \(b^{(1)}(0) = [0, 0, \cdots, 0, 1, 0, \cdots, 0]^T\) and \(\mu_{b^{(1)}}\) is a small positive step size, \(t^{(1)}(k) = [t^{(1)}(k), t^{(1)}(k - 1), \cdots, t^{(1)}(k - M + 1)]^T\), where \(M\) is the equalizer length and \(R \triangleq E\{|u^{(1)}(k)|^4\}/E\{|u^{(1)}(k)|^2\}\) is the so-called dispersion constant of the CMA, where \(a(k)\) is the transmitted data symbol [4].
Unfortunately, unlike the recursive structure in [2], our 
non-recursive transfer function of $W$ is transformed to a non-linear 
filter and placed downstream. In the tracking mode, the position of the 
whitening filter, $W$, is interchanged with the positions of $T$ and $\mathcal{PR}$, 
so that $W$ is transformed into a non-linear filter and placed downstream. 

In the tracking mode, the position of the whitening filter, $W$, is interchanged with the positions of $T$ and $\mathcal{PR}$, so that $W$ is transformed to a non-linear filter and placed downstream. Unfortunately, unlike the recursive structure in [2], our non-recursive $W$ is not immediately ready to facilitate decision feedback. Our subsequent objectives therefore are to firstly transform the non-recursive filter block, $W$, into a new but equivalent filter block that is equipped with a feedback path; and secondly, to retain the ability for the direct transfer of the parameters of $W$ between the acquisition and the tracking modes. The objectives can be jointly achieved by performing the following block transformation:

$$1 + A(z) = \frac{1 + A(z)}{1 + A(z) - A(z)} = \frac{1 + A(z)}{1 - \frac{A(z)}{1 + A(z)}}. \quad (7)$$

The term on the right hand side, which is equivalent to the non-recursive transfer function of $W$ of the acquisition mode, is now equipped with a feedback path whose transfer function is $-\frac{A(z)}{1 + A(z)}$. The transformed block $W$ in terms of $A(z)$ only of the DFE is illustrated in Fig. 2.

The adaptation of the real gain control $\mathcal{GC}$ is no longer required in the tracking mode. As for $T$, $W$, and $\mathcal{PR}$, they are jointly adapted by minimizing the decision directed MSE criterion

$$J_{DD}(a^{(2)}, b^{(2)}, \theta) = E \left\{ \left| z^{(2)}(k) - Q(z^{(2)}(k)) \right|^2 \right\} \quad (8)$$

where $a^{(2)}$ and $b^{(2)}$ are the weight vectors of $A(z)$ and $B(z)$, respectively, as in Fig. 2. They are updated as follows:

$$a^{(2)}(k + 1) = a^{(2)}(k) - \mu_m a^{(2)} e_{DD}(k) s^{(1)}(k)$$

$$b^{(2)}(k + 1) = b^{(2)}(k) - \mu_b a^{(2)} e_{DD}(k) t^{(2)}(k) \quad (9)$$

where $e_{DD}(k) = z^{(2)}(k) - Q(z^{(2)}(k))$ and

$$x^{(2)}(k) = [z^{(2)}(k-1), z^{(2)}(k-2), \ldots, z^{(2)}(k-N)]^T$$

$$y^{(2)}(k) = [y^{(2)}(k-1), y^{(2)}(k-2), \ldots, y^{(2)}(k-N)]^T$$

$$t^{(2)}(k) = [t^{(2)}(k-1), \ldots, t^{(2)}(k-M+1)]^T \quad (11a)$$

$$t^{(2)}(k) = [t^{(2)}(k-1), \ldots, t^{(2)}(k-M+1)]^T \quad (11b)$$

The phase rotator is adapted as it would in the acquisition mode according to (6). Suppose switching occurred at $k = k_0$, then their initial values are $[2]: a^{(2)}(k_0 - 1) = a^{(1)}(k_0 - 1)$ with $x^{(2)}(k_0 - 1) = 0$, $b^{(2)}(k_0 - 1) = b^{(1)}(k_0 - 1)$ and $t^{(2)}(k_0 - 1) = 0$, and lastly $\theta(k) = \theta(k_0 - 1), \forall k > k_0 - 1$.

IV. PARALLEL ADAPTATION STRATEGY FOR DUAL-MODE EQUALIZATION SCHEMES

Switching between the starting mode and the tracking mode involves both a rearrangement of the equalizer structure as well as a change between the acquisition and tracking algorithms. The switching often results in a slower rate of convergence because of the disruptions in the filtering structure and the algorithms employed. In addition, there is also a disruption in the states of the filters which adversely affect the algorithms and the output signals over several sample periods until they are flushed from the regressor vectors. Consequently, the output signals are more error-prone and the DFE, which is sensitive to incorrect decisions, may therefore exhibit pathological behavior [9], [10].

In the light of these problems, we propose a parallel adaptation strategy to ameliorate the transients by employing parallel adaptation of the linear equalizer of the acquisition mode and the DFE of the tracking mode such that only one set of filter parameters $\{A(z), B(z)\}$ is adapted and shared by the linear and decision feedback equalizers, as shown in Fig. 3. This means it is possible to always obtain the equalizer output from the output of the DFE, $z^{(2)}(k)$, as depicted in Fig. 3. Thus, this strategy will transform the original dual-mode equalization scheme in [2] to a single-mode DFE scheme whose initial acquisition is assisted by a linear acquisition equalizer. The new update equations of the DFE filter weights, $a(k)$ and $b(k)$, under parallel adaptation are

$$a(k + 1) = a(k) - \mu_a \{a_1(k) e_{MSE}(k) s^{(1)}(k)$$

$$+ a_2(k) \gamma_\alpha e_{DD}(k) x^{(2)}(k) \} \quad (12)$$

$$b(k + 1) = b(k) - \mu_b \{b_1(k) e_{MCA}(k) t^{(1)}(k)$$

$$+ b_2(k) \gamma_\beta e_{DD}(k) t^{(2)}(k) \} \quad (13)$$

where $a(0) = [0, \ldots, 0]^T$, $b(0) = [0, 0, \ldots, 0, 1, 0, \ldots, 0]^T$, $e_{DD}(k) = z^{(2)}(k) - Q(z^{(2)}(k))$, $a_1(k)$, $a_2(k)$, $b_1(k)$, $b_2(k)$ are data dependent parameters in a manner that is described below, $\gamma_\alpha$ and $\gamma_\beta$ are parameters assigned to compensate for the difference in the expected values of the respective error functions. $s^{(1)}(k)$ and $t^{(1)}(k)$ are regressor vectors of $1 + A(z)$ and $B(z)$ of length $N$ and $M$, respectively, of the linear equalizer along the top path of the new DFE. A phase rotator in the linear equalizer is not required for this setup. The signals
from the top path are obtained by employing the shared DFE
taps such that

\[
\begin{align*}
t^{(1)}(k) &= s(k) + \sum_{j=1}^{N} a_j(k)s(k-j) \quad (14a) \\
u^{(1)}(k) &= \sum_{j=0}^{M} b_j(k)t^{(1)}(k-j). \quad (14b)
\end{align*}
\]

where \(a_j(k), b_j(k)\) are the \(j^{th}\) taps of \(A(z)\) and \(B(z)\) of the new DFE, respectively. The gain control and the phase
rotator are adapted as in (3) and (6). Therefore, due to the
parallel adaptation strategy, the outputs of the “top” linear
equalizer and the “bottom” DFE will be approximately equal, i.e., \(u^{(1)}(k)\exp(-j\theta(k)) \approx z^{(2)}(k)\).

The choice of \(\alpha_1(k), \alpha_2(k), \beta_1(k)\) and \(\beta_2(k)\) is paramount
for the success of our parallel adaptation strategy in terms of
convergence speed and steady state errors, in addition to
guaranteeing a smooth transition between “modes” which
ultimately affects the convergence speed. In fact, certain soft
switching techniques that were previously proposed for dual-
mode algorithms are suitable for the acquisition and tracking
\textit{increment} vectors\(^1\) in (12) and (13) where intermediate
values between 0 and 1 can be assigned to \(\alpha_1(k), \alpha_2(k), \beta_1(k)\) and \(\beta_2(k)\) according to a reliability measure
of the equalizer output [?]. Alternative simpler techniques
(albeit with poorer performance) that include the ‘Stop-And-
Go’ algorithm [11], dual-mode type algorithm [12], as well as
the simple Benveniste-Goursat algorithm [13] can also be
employed. As a simple illustration, we will describe the
Benveniste-Goursat soft-switching technique which assigns

\[
\begin{align*}
\alpha_1(k) &= \beta_1(k) = c_1|e_{DD}(k)| \quad (15a) \\
\text{and} \quad \alpha_2(k) &= \beta_2(k) = c_2 \quad (15b)
\end{align*}
\]

where \(c_1\) and \(c_2\) are user defined positive constants \([13]\). The
new update equations of (12) and (13) under parallel adap-
tation strategy using the Benveniste-Goursat type parameters

\[
\begin{align*}
a(k+1) &= a(k) - \mu_a \left\{ c_1|e_{DD}(k)|s^{(1)}(k) + c_2|e_{DD}(k)|x^{(2)}(k) \right\} \quad (16) \\
b(k+1) &= b(k) - \mu_b \left\{ c_1|e_{DD}(k)|c_{MCA}(k)t^{(1)}(k) + c_2|e_{DD}(k)|t^{(2)}(k) \right\}. \quad (17)
\end{align*}
\]

V. SIMULATION RESULTS

The following simulations are generated to yield performance
results regarding the new scheme in Section III with and
without the enhancements using the parallel adaptation strategy
as described in Section IV. Two stationary non minimum
phase channels are selected for our simulations. The first is
an auto-regressive moving average (ARMA) channel, \(h'\), which
is more heavily colored, and the second is a moving average
(transversal) channel, \(h''\), which is lightly colored \([14]\). The
transfer functions (TF) of both channels are given by

\[
\text{TF} (h') = \frac{0.4 + z^{-1}}{1 - 0.3z^{-1} + 0.5z^{-2} - 0.2z^{-3}} \\
\text{TF} (h'') = 0.04 - 0.05z^{-1} + 0.07z^{-2} - 0.21z^{-3} - 0.5z^{-4} + 0.72z^{-5} + 0.36z^{-6} + 0.21z^{-8} + 0.03z^{-9} + 0.07z^{-10}.
\]

The transmit data format used is 16-QAM and assumed inde-
pendent and identically distributed. The real and imaginary
components of the transmit data are drawn from the set
\([-3, -1, 1, 3]\). To characterize the equalizer performance in
terms of convergence speed and steady state errors, we employ
the decision directed MSE which can be estimated using the
following recursion

\[
MSE_{DD}(k+1) = 0.99 MSE_{DD}(k) + 0.01 |e_{DD}(k)|^2. \quad (18)
\]

We tested with three types of adaptive equalizers. The first is
the DFE in [2]. The second is the DFE of Section III without
the use of parallel adaptation strategy. The third is similar to
the second equalizer except that the parallel adaptation strategy
with the Benveniste-Goursat type soft switching technique is
used. The system parameters of the three equalizers are as
follows:

The length of \(A(z)\) is 20, for the recursive whitening
filter of [2] as well as our non-recursive one (see Fig.
1). The length of \(B(z)\) is 21 for all three equalizers. For
the first two DFE’s (without parallel adaptation strategy), it
switches from the linear acquisition mode to the tracking
mode when \(MSE_{DD}(k) < 0.5657\), i.e., \(-4.9\text{dB}\). Once it is
in the tracking mode, it may be switched back to the starting
mode if \(MSE_{DD}(k) > 0.7778\), i.e., \(-2.2\text{dB}\). For the third
equalizer there is no distinct switching point. Their adaptation
equations are governed by the Benveniste-Goursat parameters
\(c_1 = c_2 = 1\) and \(\gamma_{a} = 1, \gamma_{b} = 10\) for both channels, with the
exception \(c_1\) in (16) for \(h''\) where we assigned \(c_1 = 0.5\). The
step sizes of the first two equalizers for \(h'\) in the acquisition
mode are \(\mu_{a}(1) = 5 \times 10^{-5}\) and \(\mu_{b}(1) = 5 \times 10^{-6}\), while in
the tracking mode they become \(\mu_{a}(2) = \mu_{b}(2) = 10^{-4}\), for
\(h''\), we assigned \(\mu_{a}(1) = \mu_{b}(2) = 2.5 \times 10^{-4}\) and \(\mu_{b}(1) = 2.5 \times 10^{-5}\). As for the third equalizer, \(\mu_{a} = 5 \times 10^{-5}\)

\(^1\)The increment vector of the stochastic gradient update equation is the
regressor vector multiplied by the error function.
and $\mu_b = 5 \times 10^{-6}$ for $h'$; and $\mu_a = 2.5 \times 10^{-4}$ and $\mu_b = 2.5 \times 10^{-5}$ for $h''$.

Fig. 4 shows the MSE plots of the three above equalizers averaged over 100 independent trials using 15,000 and 10,000 symbols for channels $h'$ and $h''$, respectively, at SNR levels of 15 dB and 25 dB. Two main conclusions are drawn. Firstly, the first and second equalizers exhibit almost identical performances in terms of convergence speed and steady state errors for both channels under both SNR levels. This shows that our simplified DFE as described in Section III has maintained the standard set by the original DFE in [2]. Secondly, the use of the parallel adaptation strategy using the Benveniste-Goursat technique improves performance of the third equalizer over the first two equalizers. The success of the smooth switching strategy is implicit in the faster rate of convergence achieved by the third equalizer since the first two equalizers would usually switch between their acquisition and tracking modes several times before convergence is finally achieved. For channel $h''$, particularly at high SNR levels, the advantages of parallel adaptation is less because error propagation in the DFE is less severe. On a different note, the third equalizer that employs the Benveniste-Goursat technique yields higher steady state MSE because of the non-zero contribution from the non-DD portions of the update equations (16) and (17) even after the eye is open. The problem with the high steady state MSE can be readily solved by employing the reliability measure technique [15] which is more computationally intensive.

VI. CONCLUSION

In this paper, we have firstly proposed an alternative scheme to the original DFE in [2] that exhibits similar overall convergence and steady state performances. The subsequent development of the parallel adaptation strategy ensures the smooth switching between the acquisition and tracking modes of the fast-convergence DFE scheme that involves not only a switch in algorithms but also a switch in the actual filtering structure. The smooth switching is achieved through a parallel adaptation strategy that can employ any suitable cost function combining technique previously developed for dual-mode type algorithms. This was illustrated using the Benveniste-Goursat technique which resulted in a parallel adaptation strategy that exhibited superior convergence speed to classical dual mode equalization schemes particularly for more severe channels.

REFERENCES