New block-based blind equalization algorithms

This item was submitted to Loughborough University's Institutional Repository by the/an author.


Additional Information:

- This is a conference paper [© IEEE]. It is also available at: http://ieeexplore.ieee.org/ Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Metadata Record: https://dspace.lboro.ac.uk/2134/5792

Version: Published

Publisher: © IEEE

Please cite the published version.
This item was submitted to Loughborough’s Institutional Repository (https://dspace.lboro.ac.uk/) by the author and is made available under the following Creative Commons Licence conditions.

For the full text of this licence, please go to:
http://creativecommons.org/licenses/by-nc-nd/2.5/
New Block-Based Blind Equalization Algorithms

John C. Allwright, Antonis Giannopoulos and Jonathon A. Chambers

Control and Communications & Signal Processing Research Groups
Department of Electrical and Electronic Engineering
Imperial College of Science, Technology and Medicine
Exhibition Road, London, SW7 2BT, U.K.
E-mail: {j.allwright,j.chambers}@ic.ac.uk

Abstract - New block-based blind equalization algorithms are introduced based upon the cost function underlying the recently proposed soft constraint satisfaction blind equalization algorithm. The derivation of these algorithms is based on mapping the original constrained optimization problem in $\mathbb{C}^N$ into a much simpler optimization problem in $\mathbb{R}^2$. Versions of the new algorithms are also developed for fractionally-spaced equalizers. Simulations on a baud-spaced and a fractionally-spaced channel support the potential of the resulting block-based techniques.

I. INTRODUCTION

When digital data are transmitted over a finite bandwidth analogue channel by means of discrete pulse-amplitude modulation schemes, the number of detectable levels that can be supported is essentially limited by intersymbol interference. Intersymbol interference arises because of the spreading of the transmitted pulse caused by the dispersive nature of the channel together with filtering in the transmitter and receiver. An effective method to minimize the system degradation is to connect an equalizer in cascade with the channel, so that the latter can model the inverse channel characteristics. The combined action of the channel and the equalizer is then in the zero channel noise situation, essentially an identity operation [1]. This is the objective of linear equalization; where a linear filter, that is the equalizer, is applied to the received signal, in order to improve the probability of correct symbol detection.

Blind equalization is unsupervised adaptation of the tap-coefficients of the equalizer to find the approximate inverse of the channel impulse response. Blind algorithms are conventionally implemented either in a baud-spaced ($T$-spaced) or a fractionally-spaced ($F$-spaced) setting. In a $T$-spaced implementation, the channel output is sampled at the signalling rate, whereas in an $F$-spaced setting, a higher sampling rate is used dependent upon the pulse shaping. Much work has considered the advantages of fractionally-spaced equalization [2,3]. In both $T$- and $F$-spaced cases, the decision-directed [4] (DD), Sato [5], Godard [6] and constant modulus [7] (CMA) adaptive blind equalization algorithms and the decision feedback configuration can be used. In this paper we consider the recently proposed soft constraint satisfaction (SCS) algorithm [1], which attempts to mitigate the problem of undesirable local solutions in the CMA cost function. We address the slow convergence of the sequential SCS algorithm by the provision of a novel block solution based upon mapping the original optimization formulation into a problem in $\mathbb{R}^2$.

The following notation is employed: vector quantities appear in bold, $| \cdot |$ is the absolute value, $\text{sgn} (\cdot)$ is the signum function, $\mathbb{E} \left[ \cdot \right]$ is the statistical expectation operator, $\text{Re} \left[ \cdot \right]$ and $\text{Im} \left[ \cdot \right]$ are respectively the operators which extract the real and imaginary part of a complex quantity. The $L_p$-norm, transpose and the Hermitian transpose of a vector or matrix are respectively denoted by $\| \cdot \|_p$, $(\cdot)^T$ and $(\cdot)^H$. The Moore-Penrose pseudo inverse of a matrix is indicated by $(\cdot)^\#$. The discrete time index is $k$. The equalizer vector $\Theta_k \in \mathbb{C}^{N_e}$ and input vector $X_k \in \mathbb{C}^{N_o}$ are defined respectively as $\Theta_k^T = [\theta_{k,0} \theta_{k,1} \cdots \theta_{k,N_o-1}]$ and $X_k^H = [x_k \ x_{k-1} \cdots x_{k-N_o+1}]$ and the equalizer output is $y_k = X_k^H \Theta_k$.

II. FUNDAMENTALS

A. Problem Statement

Consider the baseband representation for a digital transmission structure shown in Figure 1,
where $C(z)$ and $\Theta(z)$ represent respectively the $z$-transforms of the impulse response of the channel and equalizer, $a_k$ is the channel input with $a_k \in \{\pm 1 \pm j\}$, namely a complex Bernoulli sequence, $v_k$ is statistically independent additive channel noise, and $x_k$ and $y_k$ are respectively the equalizer input and output. The objective for the equalizer, in the noise free case, is to achieve a pure delay operation, i.e. where $C(z)\Theta(z) = e^{\jmath \omega \delta}$. We desire the recovered symbol $\hat{a}_k$ to be identical to $a_k$ up to some constant delay, $\delta$, and phase shift, $\theta$.

![Fig. 1. Baseband model of a digital communications system](image)

A linear, causal and finite impulse response (FIR) model of length $N_C$ is assumed for the channel:

$$C(z) = \sum_{m=0}^{N_C-1} c_m z^{-m}$$  \hspace{1cm} (1)

The $T$-spaced equalizer of length $N_\Theta$, $\Theta(z)$ and the combined space $S(z)$ are also linear, causal FIR and defined respectively as

$$\Theta(z) = \sum_{m=0}^{N_\Theta-1} \theta_m z^{-m}$$  \hspace{1cm} (2)

$$S(z) = C(z)\Theta(z) = \sum_{m=0}^{N_C-1} S_m z^{-m}$$  \hspace{1cm} (3)

Note that $c_m$, $\theta_m$ and $s_m$ are the complex impulse response coefficients and that $N_S = N_C + N_\Theta - 1$.

### B. Existing Approaches to the Blind Equalization Problem

The generic update equation of the existing Bussgang-type blind equalization algorithms is [8]

$$\Theta_{k+1} = \Theta_k + \mu(p(Y_k) - y_k)X_k \text{ for } k = 0, 1, \ldots$$  \hspace{1cm} (4)

where $p(Y_k)$ is a zero-memory non-linearity (ZNL) and $p(Y_k) - y_k$ is the error function.

**Godard's algorithms**[6]: Godard proposed to minimize the family of cost functions defined by

$$J_G(\Theta) = \frac{1}{2\mu} E\left[ \left| \frac{p(y_k)^p - R_p}{\|a_k\|^p} \right|^2 \right]$$

where $R_p = E\left[ \frac{a_k^{2p}}{\|a_k\|^p} \right]$.

The most frequently utilized member of this family is the Constant Modulus (CM) Algorithm and corresponds to the case where $p = 2$ [7] The corresponding update equation is the following

$$\Theta_{k+1} = \Theta_k + \mu(R_2 - \|y_k\|^2)X_k$$

$$\text{where } R_2 = E\left[ \frac{|a_k|^4}{\|a_k\|^2} \right]$$  \hspace{1cm} (6)

**Normalized CM (NCM) algorithms**[9,10]: Hilal and Duhamel proposed a normalized algorithm for constant modulus signal constellations. For the constant modulus algorithm defined above, they showed that the fastest convergence rate occurs if the update equation is modified to include a normalization term

$$\Theta_{k+1} = \Theta_k + \frac{1}{\|X_k\|^2} (\text{sgn}(y_k) - |y_k|^2)X_k$$

$$\text{The same algorithm has been proposed in } [11] \text{ as a link between blind equalization and classical adaptive equalization where a desired (or training) signal } d_k \text{ is available. The NCM algorithm has been formally derived from the exact solution of}$$

$$\begin{align*}
\min_{\Theta_{k+1}} & \left\{ \frac{1}{\mu} \left| X_k \right|^2 \left\| \Theta_{k+1} - \Theta_k \right\|^2 \right. \\
\text{subject to } & X_k^H \Theta_{k+1} X_k = R^2
\end{align*}$$

**Sequential Soft Constraint Satisfaction Algorithm**[11]: The SCS algorithm follows the formulation of the blind equalization problem for a complex input alphabet as

$$\begin{align*}
\min_{\Theta_{k+1}} & \left\{ \left\| \Theta_{k+1} - \Theta_k \right\|^2 \right. \\
\text{subject to } & X_k^H \Theta_{k+1} X_k = R^2
\end{align*}$$

where the optimization problem in (9) is motivated by the principle of minimum disturbance together with the enforcement of a constraint upon the updated equalizer weight vector. The SCS algorithm is based upon
softening the enforcement of the constraint during adaptation as the constraint may be inappropriate when the equalizer weight vector is far from a desired solution.

The derivation of the SCS algorithm has been based upon a gradient descent interpretation; the update equation of which is given by

$$\Theta_{k+1} = \Theta_k + \frac{\mu}{\|X_k\|^2} (1 - \frac{\|y_k\|}{R}) y_k X_k$$ (10)

In this work, however, we investigate the direct solution of the constrained optimization problem represented by (9), motivated by the potential improved convergence properties of a block estimator.

III. NEW BLOCK APPROACH

An exact block solution to the constrained optimization problem posed by (9) is presented.

A. The Minimization Solution

The following vector definitions are adopted

$$z = \begin{bmatrix} x^R \\ z^I \end{bmatrix} \in \mathbb{R}^{2N_e}$$
$$z^* = \begin{bmatrix} -z^R \\ z^I \end{bmatrix} \in \mathbb{R}^{2N_e}$$ (11)

where $x^R = \text{Re}[z] \in \mathbb{R}^{N_e}$ and $z^I = \text{Im}[z] \in \mathbb{R}^{N_e}$. Some useful identities that will be employed are

$$z^H z = z^T z$$ (12),

$$X^H \Theta = X^T \Theta + jX^T \Theta$$ (13),

$$\Theta^H X = X^T \Theta - jX^T \Theta$$ (14),

$$X^T X = 0$$ (15),

and

$$\|z\|^2 = \|z^R\|^2 + \|z^I\|^2 = \|z_R\|^2 + \|z_I\|^2$$ (16).

Using (16), (13) and (14), the optimization problem (9) can be rewritten as

$$\min_{\Theta_{k+1}} \left\{ \Theta_{k+1} - \Theta_k \right\}$$
subject to

$$\left( X_k^T \Theta_{k+1} \right)^2 + \left( X_k^T \Theta_{k+1} \right)^2 = R^2$$ (17)

Let

$$Y = [X_k, X_k] \in \mathbb{R}^{2N_e \times 2}$$ (18)

so that we can write

$$\Theta_{k+1} = Y \eta + \gamma$$ (19)

for some $\eta \in \mathbb{R}^2$ and some $\gamma$ belonging to the null space of $Y$, i.e. $\gamma \in \mathbb{R}^\perp \mathcal{Y}$.

From (15) and (16) we get

$$X_k^T \Theta_{k+1} = X_k^T [X_k \eta_1 + X_k \eta_2 + \gamma] = X_k^T X_k \eta_1 = \|X_k \|^2 \eta_1$$ (20)

$$X_k^T \Theta_{k+1} = X_k^T [X_k \eta_1 + X_k \eta_2 + \gamma] = X_k^T X_k \eta_2 = \|X_k \|^2 \eta_2$$

and consequently

$$(X_k^T \Theta_{k+1})^2 + (X_k^T \Theta_{k+1})^2 = \|X_k \|^2 \|\eta\|^2$$ (22)

Hence a solution $\Theta_{k+1}$ of the initial minimization problem is given by $\Theta_{k+1} = Y \eta + \gamma$ where $\eta$ and $\gamma$ solve

$$\min_{\eta \in \mathbb{R}^2} \left\{ \|Y \eta - \gamma\|^2 : \|\eta\| = R \|X_k \|^2 \right\}$$ (23)

The given vector $\Theta_k$ can be written as $\Theta_k = \alpha_y + \beta_y$ for unique $\alpha_y \in \mathbb{R}(Y)$ and $\beta_y \in \mathbb{R}(Y)$. Since these are orthogonal projections we have $\alpha_y = Y(Y^* \Theta_k)$ and $\beta_y = (I - Y^* \Theta_k) \Theta_k$ (24). Therefore, the minimization problem can be further expressed as follows

$$\min_{\eta \in \mathbb{R}^2} \left\{ \|Y \eta - \gamma\|^2 : \|\eta\| = R \|X_k \|^2 \right\}$$

$$= \min_{\eta \in \mathbb{R}^2} \left\{ \min_{\eta \in \mathbb{R}(Y)} \left( \|\alpha_y + \beta_y - Y \eta - \gamma\|^2 \right) : \|\eta\| = R \|X_k \|^2 \right\}$$

$$= \min_{\eta \in \mathbb{R}(Y)} \left\{ \min_{\eta \in \mathbb{R}(Y)} \left( \|\alpha_y - Y \eta\|^2 + \|\beta_y - \gamma\|^2 \right) : \|\eta\| = R \|X_k \|^2 \right\}$$

since $\alpha_y - Y \eta \in \mathbb{R}(Y)$ and $\beta_y - \gamma \in \mathbb{R}(Y)$.

$$= \min_{\eta \in \mathbb{R}} \left\{ \|\alpha_y - Y \eta\|^2 + \|\beta_y - \gamma\|^2 : \|\eta\| = R \|X_k \|^2 \right\}$$

where the minimising $\gamma$ is $\gamma = \beta_y$.

$$= \min_{\eta \in \mathbb{R}} \left\{ \|\alpha_y - Y \eta\|^2 : \|\eta\| = R \|X_k \|^2 \right\}$$ (25)

The above analysis has reduced the original minimization on $\mathbb{R}^{N_e}$ to a minimization problem on $\mathbb{R}^2$, which is easier to solve. In fact, this problem can be solved by geometry. From the above derivation it follows that $Y^* Y = \|X_k \|^2 I_2$, hence from (24)

$$\|\alpha_y - Y \eta\|^2 = \|Y^* \Theta_k - \eta\|^2 =$$

$$\|X_k \|^2 \|\eta - Y^* \Theta_k\|^2$$ (26)
Hence, the unconstrained global minimizer of $\|x - y\|_2^2$ with respect to $y \in \mathbb{R}^2$ is $y = y^* \Theta x$.

The contours and the feasible set for the constraint $\|y\|_2 = R\|x\|_2$ are as shown on Figure 2. It can be deduced that the unique constrained globally optimal $y$ for (9) is the multiple $R\|x\|_2^2$ of the normalised direction $y^* \Theta x \|x\|_2^{-1}$, which is given by

$$
\eta = y^* \Theta x \|x\|_2^{-1} R \|x\|_2^2
$$

Fig. 2. Contours used in the minimization problem

Hence the unique $\Theta_{k+1}$ that solves the initial minimization problem (9) is

$$
\Theta_{k+1} = Y \eta + \gamma = Y (Y^* \Theta_x \|x\|_2^{-1}) R \|x\|_2^2 + (I - Y Y^*) \Theta_x
$$

Equation (28) can be simplified by evaluating the pseudo-inverse $Y^* = (Y^T Y)^{-1} Y^T = Y^T \|x\|_2^{-2}$ and $Y^* \Theta_x = (Y^T Y)^{-1} Y^T \Theta_x = (Y^T \Theta_x \|x\|_2^{-2}$. This provides the basis for the new proposed block algorithm

$$
\begin{bmatrix}
\Theta_{k+1} \\
\Theta_{k+1}^{odd}
\end{bmatrix} = \Theta_x + Y (Y^T \Theta_x) \left( (Y^T \Theta_x \|x\|_2^{-2}) R - 1 \right) \|x\|_2^2
$$

B. Developed Block Algorithms

The solutions, which are obtained by the SCS algorithm, correspond to the points of zero gradient of the following cost function

$$
J_{\text{SCS-1}} = E \left[ \frac{1}{\|x\|_2^3} \left( \frac{y_{2k}^2}{2} + \frac{y_{2k+1}^3}{3} \right) \right]
$$

It has been shown in [1] that this cost function has identical global minima to those of the CMA cost function. In [12] the global solutions of CMA are demonstrated to be close to those predicted by the Wiener filter. Therefore, we propose to introduce averaging to the block-based algorithm (29), which gives an exact solution for a block of data, in order to obtain solutions that will also approximate those of the Wiener filter. To this end the following simple schemes are proposed:

The first implementation recalculates the equalizer weight vector for each new block of data, which is offset by one sample from the previous block, with length equal to that of the equalizer. The most recent weight vector is averaged with the previously obtained one. This approach is termed Allw3. The second implementation is similar to Allw3 but the averaging is performed over the three most recent weight vectors. This is referred to as Allw4. Centre tap initialisation is used to initialise $\Theta_0$. These two averaging schemes have been chosen because they provide fastest convergence, other schemes can be found in [13]. The averaging in these schemes is in a sense counteracting the solution given by (29); but to approach a solution defined by the Wiener filter, more richness in the data is necessary than that used to formulate (29). This must be offset, however, by allowing the new information contained within the most recent solution of (29) to dominate the dynamics of the algorithm, hence such short averaging as posed by the Allw3 and allw4 algorithms is postulated as the best approach for fast initial convergence.

IV. FRACTIONALLY-SPACED EQUALISATION

An equivalent multichannel model for a fractionally-spaced equalizer is described in [3], for the case of two channels and a $T$-spaced source. The source sequence is subdivided into even and odd $T$-spaced counterparts of relative delay $T/2$, so that $c_{2n}^{\text{even}} = c_{2n}$ and $c_{2n}^{\text{odd}} = c_{2n+1}$ for $n = 0, 1, \ldots$ and in a similar manner we have $\Theta_{2n} = \Theta_{2n}$ and $\Theta_{2n+1} = \Theta_{2n+1}$. Therefore, we have two branches each having the equivalent of a channel and an equalizer. To derive a block-based solution for the equalizers in the $T/2$ multichannel model the minimization problem can be formulated as follows.
\[
\min \left\| \begin{bmatrix} \Theta_{k+1}^{\text{even}} - \Theta_{k}^{\text{even}} \\ \Theta_{k+1}^{\text{odd}} - \Theta_{k}^{\text{odd}} \end{bmatrix} \right\|_2^2 \quad \text{subject to} \\
X_k = \begin{bmatrix} X_k^{\text{even}} \\ X_k^{\text{odd}} \end{bmatrix} \quad \text{and} \quad \Theta_{k+1} = \begin{bmatrix} \Theta_{k+1}^{\text{even}} \\ \Theta_{k+1}^{\text{odd}} \end{bmatrix}
\]

and the block solution given by (29) can be easily adapted to solve the above by concatenating the equalizer weight and channel outputs of the two sub-branches.

V. EXPERIMENTAL RESULTS

The complex Bernoulli sequence \( \{ \alpha_k, j\beta_k \} \) was applied to the channel, with the real random variables \( \alpha_k = \pm 1 \) and \( \beta_k = \pm 1 \) with zero mean and unit variance. The impulse response of the 5-tap channel is described by a raised cosine \( c = [0.3149 \ 0.2194 \ 1.0 \ 0.2194 \ 0.3149] \). In order to facilitate the comparison, constellation plots that present the output of the Allw3 and SCS algorithms, when the signal to noise ratio (SNR) is 30 dB can be found in Figure 3.

![Fig. 3. Comparison of the output constellations from an SCS and an Allw3 7-tap equalizer](image)

A careful observation of these constellation plots reveals that the Allw3 algorithm has slightly denser constellation groupings than the SCS algorithm. We can calculate the average deviation of the obtained constellation points from the ideal constellation points; this is done by calculating the Euclidean distance between each obtained point and its closest transmitted constellation point and then averaging over all such points. In the 7-tap case, SCS gives a deviation of 0.1032 while Allw3 yields 0.0916. In our experience the block approaches are consistently found to have improved initial convergence, as seen on Figure 4 for the Allw3 and Allw4 algorithms, obtained from 200 Monte Carlo trials, and the more tightly packed constellation suggests reduced steady-state error. For fractionally-spaced equalization, the two F-spaced channels which are used are \( c^a = [0.24 -0.20 -0.16 \ 0.14 \ 0.15] \) and \( c^b = [0.21 \ 0.51 \ 0.85 \ 0.53 \ 0.31] \). These channels have a more degrading effect but still no common zeros, [3], the locations of the zeros of these channels being 1.0178 \( \pm \) j0.5135 and -0.6011 \( \pm \) j0.3459 for \( c^a \), and -0.9243 \( \pm \) j1.2215 and -0.29 \( \pm \) j0.7383 for \( c^b \).

In Figure 5, the corresponding output constellations for the SCS and Allw4 algorithms are shown when the SNR is 30 dB. Figures 5(a) and (c) show the output constellations for the entire simulation of length 10000 whilst Figures 5(b) and (d) display only points after 3000 iterations when the SCS algorithm is in steady state. It is clear that the constellation densities are considerably denser in the latter case.

![Fig. 4. Squared error performance for the SCS, Allw3 and Allw4 equalizers for the first 30 iterations](image)

VI. CONCLUSION

New block-based blind equalization algorithms have been shown to exhibit much improved convergence characteristics than the SCS algorithm. This improvement in performance can be seen for both baud-spaced and fractionally-spaced equalizers. An exact solution to the optimization problem is used in order to derive the block-based algorithms, which with the introduction of averaging of the given solutions is also made sequential. The experimental results support the theoretical findings. Tuning the algorithm by careful choice of the averaging methods it uses, further improves the performance. The increased computational complexity posed by equation (29) could be avoided in a real-time application.
if the block-based scheme is used for the first iterations before the pure sequential SCS algorithm takes over. A real-time implementation is worthy of investigation as a continuation and a possible practical extension to this research.

Fig. 5. Comparison of the output constellations (a) and (b) from an SCS 5-tap equalizer and those of an Allw4 5-tap equalizer (c) (d); (a) and (c) correspond to the entire simulation of length 10000 whilst (b) and (d) are only for points after discrete time 3000

REFERENCES