Performance Improvement of Blind Equalizers Using the Constant Modulus Algorithm

Andrew W. Hull and W. Kenneth Jenkins
Department of Electrical and Computer Engineering and
The Coordinated Science Laboratory
1308 West Main St.
University of Illinois
Urbana, Illinois 61801

Abstract - A novel technique for orthogonalizing the update relation for steepest descent algorithms using pseudo-random (PR) modulation has been developed. In this paper it is shown that the application of PR modulation to the Constant Modulus Algorithm (CMA) improves the rate of convergence and reduces the steady state misadjustment. Computer simulation results are presented to illustrate the performance improvement of blind equalization using the CMA with PR modulation.

1. Introduction

An effective tool for investigating the performance of adaptive filters is the mean square error (MSE) surface. The MSE surface is an N-dimensional function of the filter coefficient vector \( \mathbf{w} \), as generated by

\[
E[(e^2(n))] = E[d(n)^2] - 2 \mathbf{w}^T \mathbf{p} + \mathbf{w}^T \mathbf{R}_{xx} \mathbf{w},
\]

where \( \mathbf{R}_{xx} \) is the input autocorrelation matrix, \( \mathbf{p} \) is the cross correlation matrix between the input and the training signal, and \( E[ \cdot ] \) denotes the statistical expected value. The function is constructed by fixing the adjustable coefficients, exercising the adaptive filter with a noise input, and measuring the error. This process is repeated until the entire error surface is created. A typical error surface is shown below:

The object of an adaptive filter is to search over this error surface and locate the values of \( \mathbf{w} \) which yield the minimal MSE. The properties of the surface affect the performance of the optimization procedure implemented by the adaptive filter. For finite impulse response (FIR) adaptive filters the error surface is quadratic and convex with a unique minimizing solution. For the case of an IIR adaptive filter the error surface may have multiple minimizing solutions. About those points the surface is well approximated by a quadratic surface. In either case the shape of the error surface in the neighborhood of the minimum is determined by the autocorrelation matrix, \( \mathbf{R}_{xx} \). The eigenvalues of \( \mathbf{R}_{xx} \) determine the axes of the ellipses of the error surface, and the eigenvectors of \( \mathbf{R}_{xx} \) determine its orientation in N-space.

The condition number, \( r \), is defined as the ratio of largest and smallest eigenvalues of \( \mathbf{R}_{xx} \),

\[
r = \lambda_{\text{max}}/\lambda_{\text{min}},
\]

and is a measure of the "skewness" of the error surface associated with \( \mathbf{R}_{xx} \). For small condition numbers the error surface has circular cross-sections. This corresponds to small variation among the eigenvalues of \( \mathbf{R}_{xx} \). Large condition numbers indicate very narrow elliptical cross-sections for the error surface, and greatly disparate eigenvalues of \( \mathbf{R}_{xx} \).

The optimization algorithms examined in this paper follow the familiar form of

\[
w(n+1) = w(n) + 2\mu \mathbf{f}[e(n)]x(n)
\]

where \( \mu \) is a small gain term, \( \mathbf{f}[ \cdot ] \) is the chosen norm of the error, \( e(n) \), and \( x \) is the input to the adaptive filter. Steepest descent is the most popular numerical optimization technique used for adaptive filtering. In the adaptive filtering literature it is known as the LMS algorithm and takes the form of

\[
w(n+1) = w(n) + 2\mu e(n)x(n)
\]

The performance of the LMS algorithm may be determined by examining the mean behavior of the error. By expressing the error in terms of a set of axes centered about the optimal solution, \( \mathbf{w}^* \), the decay of \( E[e(n)] \) follows a geometric series generated by
approximately orthogonal and it is possible to whiten the
approach an orthogonal transform is applied to the input
approaches are also useful, the most common being the
Transform Domain Adaptive filter (TDAF). In this
Newton-type algorithms, or self-orthogonalizing
filter algorithms discussed in this paper is to achieve rapid
approximation improves the matrix product in Eq. (7)
relation of Eq.

$$\text{(I - } \mu \text{R}_{xx} \text{)}^n.$$  \hspace{1cm} (5)

Thus "modes" of convergence exist which are determined
by the statistics of the input signal. Aligning these axes
parallel to those defined by the major and minor axes of
the quadratic form by rotation results in

$$\text{(I - } \mu \Lambda \text{)}^n.$$  \hspace{1cm} (6)

The entries of the diagonal matrix, \( \Lambda \), are the eigenvalues
of the autocorrelation matrix \( \text{R}_{xx} \). Each mode is
determined then by one of the \( N \) eigenvalues of \( \text{R}_{xx} \).

Each mode must converge for the adaptive filter to
converge. The step size parameter \( \mu \) must, therefore, lie
between 0 and \( 1/\lambda_{\text{max}} \). Since the rate of convergence is
controlled by the slowest mode, convergence is most swift
when all modes are equal. This rate is thus given by
white input signals, where \( \lambda_{\text{min}} = \lambda_{\text{max}} \), corresponding to
an error surface with circular contours. Slow rates of
convergence are observed for highly correlated input data.
The eigenvalues of \( \text{R}_{xx} \) for these signals are highly
disparate and the associated error surface is elliptical.

Many improvements upon Eq. (4) have been posed in
the literature to reduce or eliminate the dependence of the
convergence rate of the LMS algorithm on the input signal
statistics. The goal is to cancel the matrix \( \text{R}_{xx} \) from Eq
(5), thus de-coupling and normalizing the modes of
convergence. This is accomplished by incorporating the
inverse of an approximation of \( \text{R}_{xx} \) into the update
relation of Eq. (4). The modal relations then become

$$\text{(I - } \mu \text{R}^\prime(n)_{xx}^{-1}\text{R}_{xx})^n.$$  \hspace{1cm} (7)

where \( \text{R}^\prime(n)_{xx} \) is an approximation of \( \text{R}_{xx} \). As the
approximation improves the matrix product in Eq. (7)
becomes diagonal, and, in the case of exact representation,
identity. The update relation necessary to achieve is

$$w(n+1) = w(n) + 2\mu \text{R}(n)_{xx}^{-1}e(n)x(n).$$  \hspace{1cm} (8)

Algorithms sharing this form are known generically as
Newton-type algorithms, or self-orthogonalizing
algorithms. This form is also shared with Recursive Least
Squares (RLS) algorithms. The use of lattice filters to
generate the innovations sequence is another approach to
effect orthogonalization. Many other suboptimal
approaches are also useful, the most common being the
Transform Domain Adaptive filter (TDAF). In this
approach an orthogonal transform is applied to the input
signal before the adaptive filter. The transformed input is
approximately orthogonal and it is possible to whiten
the signal by power normalization. The goal of the adaptive
filter algorithms discussed in this paper is to achieve rapid
correlation by orthogonal update relations without greatly
increasing the computational burden beyond that of the
LMS algorithm.

2. Nonlinear Orthogonalization

Applying pseudo-random (PR) modulation to the input
of an adaptive LMS filter was first proposed in [1]. PR
modulation is applied in Direct Sequence Spread Spectrum
(DS-SS) communications systems to increase the
bandwidth of transmitted signals and reduce interference
after appropriate processing, as measured by the signal-to-
noise ratio. It is this bandwidth-increasing effect of PR
modulation which is exploited to great effect in adaptive
filtering.

Recall that the simplicity of the LMS is mitigated by
its possibly slow rate of convergence [2]. The PR
modulation is used to de-correlate the input signal input
of the adaptive filter, ideally so that \( \lambda_{\text{min}} = \lambda_{\text{max}} \). The
correlation matrix \( \text{R}_{xx} \) of the spread-spectrum (SS) input
now has an improved condition number, the associated
quadratic form (the MSE error surface) has more circular
contours, and the modes of the error system are
approximately de-coupled. The computation burden
required to achieve orthogonalization is therefore removed
from Eq. (8), i.e., it is removed from "inside" the adaptive
algorithm, placed "outside" of the adaptive filter, and
performed by the PR modulation. Uniform application of
this idea is constrained by the nonlinear nature of the PR
modulation. Access to the requisite signals may not be
possible for all configurations of adaptive filters in
various applications.

For the case of system identification both the system to
be identified, \( h \), and the adaptive filter, \( w \), are excited by
identical input noise signals. When access to those
signals is available the spectrally enriched input provides
greatly improved performance of the adaptive filter. In this
case the signals of interest are

$$d(n) = h' (c \circ x) \quad \text{and} \quad y = w' (c \circ x),$$  \hspace{1cm} (9)

where \( \circ \) denotes the "component-wise vector product," and
\( c \) is a vector of samples constructed from the modulating
PR sequence. Each signal is modulated equivalently, and
the solution is unchanged and rapidly found. For the case
when access to the probing signal is not possible then

$$d(n) = c \circ (h' x) \quad \text{and} \quad y = w' (c \circ x).$$  \hspace{1cm} (10)

In general \( \circ \) does not commute with the normal vector
product, and convergence will not result in the Wiener
solution. Parameter drift is typically observed. In these
instances application of DS/SS is unsound. The other
configurations of adaptive filters, prediction equalization and
interference cancellation, degrade similarly.

The recovery of the plant signals is another difficulty.
When access to the probing signal is possible, the plant
output may be recovered. Since each system input is
identical the system converges to the Wiener solution. To recover the plant output another adaptive filter may be run in parallel to the first adaptive process. This filter operates on unscrambled probing data, and its parameters are copied from the first adaptive system. After the rapid convergence of the first adaptive filter the output of the second represents the output of the unknown system, as depicted in Figure 1.

An important aspect of this research then is to discover adaptive filter applications that are well suited for DS-SS signal conditioning. The Constant Modulus Algorithm (CMA), as reported and analyzed in [3] and [4] for blind equalization, appears to be one such application that is ideally suited for the DS-SS signal conditioning approach. The sequel is a brief examination of the potential improvements in performance of the CMA for application with spread spectrum techniques.

Of first interest is the fact that PR modulation does not affect the modulus of the input signal. Scrambling is but a pseudo-random toggling of the sign of the data sample. The original corruption of the input data by channel distortion remains unaffected. Furthermore, the CMA is analogous in many ways to the LMS algorithm. It is reasonable to expect performance improvements in CMA comparable to those of LMS when DS-SS concepts are applied.

The general form of the error function for the CMA is given by

$$E(1 | y(n))^p - 1 | q),$$

where $p$ and $q$ are typically small integer values. Note that the error function of Eq. (11) measures how far away from unity the modulus of the output signal is so that the coefficients can be updated in a way to reduce this error. In the experiments presented in the next section, examples of filter adaptation are presented for cases with $p, q = 1, 2$. In general these cases are denoted as CMA p-q.

The "modal" analysis of [4] derives several approximate expressions for the time constants of the CMA, 1-1, 2-1, 1-2, and 2-2 algorithms for various technical conditions. They are repeated below:

$$\tau_{i, 11} = \mu \sqrt{\lambda_i}$$
$$\tau_{i, 12}, \tau_{i, 21} = \frac{2}{\mu \lambda_i}$$
$$\tau_{i, 22} = \frac{1}{12 \mu \lambda_i^2 |w|^2}$$

Of particular interest is the presence once again of $\lambda_i$, the i-th eigenvalue of $R_{xx}$. Convergence is fastest for equal modes, and it is expected that the orthogonalization of DS-SS will accelerate the convergence of CMA.

**Results for the misadjustment, or "excess mean square error", of CMA are developed in [5]. There it was found that the misadjustment may be approximated by**

$$M = \frac{c (\epsilon_{\min} / \epsilon_{\min}) \mu}{1 - 8 \lambda \mu}.$$  

The presence and location of an eigenvalue of the input autocorrelation matrix is reminiscent of the LMS algorithm. This eigenvalue may again be controlled by PR modulation.

3. Computer Simulations

The theory of DS-SS signal conditioning for adaptive filters was tested by performing several computer experiments using CMA in various configurations. In each of the following experiments presented in Figures 2 - 5 the solid learning curve is for the direct (DR) LMS algorithm and the dotted learning curve is for the spread spectrum (SS) case. Shown first, in Figures 2 and 3, are the learning curves of two CMA 2-2 experiments. In these two cases the unknown system was a low order FIR system with an input that was colored by passing white pseudo-noise through a lowpass filter. The coloring filter for Case 1 was much narrower than that of Case 2, and hence the overall convergence rates shown in Figure 2 are slower than those of Figure 3. The rates of convergence are improved by the spread spectrum signal conditioning as implemented with PR modulation. For the example shown in Figure 3, the improvement in convergence rate is quite dramatic. Cases 3 and 4 presented in Figures 4 and 5 show similar experiments for the CMA 1-1 implementation.

The improvement achieved by the DS-SS signal conditioning is quite substantial for Case 3 where the input signal was filtered with a narrow band coloring filter. In Case 4 the coloring filter had a much wider bandwidth, resulting in similar convergence rates for both experiments. Case 3 (Figure 4) indicates that convergence acceleration is also possible for CMA 1-1. Note that Case 4 (Figure 5) shows dramatically improved steady state misadjustment for the CMA 1-1 algorithm using DS-SS conditioning, as predicted by the expression for misadjustment given in Eq. (13).

4. Summary

Promising results have been demonstrated by applying direct sequence spread spectrum signal conditioning to the constant modulus adaptive filtering algorithm. Further analysis will be needed to explain and confirm many interesting results, which up to this time, have been observed mainly through computer simulation.

**Acknowledgement**

This work is supported by the Joint Services Electronics Program under contract N00014-84-C-0149. The results and opinions expressed in this paper do not necessarily represent those of the sponsoring agency.
References


