# BLIND CHANNEL EQUALIZATION USING NEW STATISTICAL SIGNAL PROCESSING TECHNIQUES


**by**

**Aqdas Naveed Malik**

**A dissertation submitted to M.A.J.U in partial fulfillment of the requirements for the degree of**

## DOCTOR OF PHILOSOPHY

**Department of Electronics Engineering**

**Faculty of Engineering and Sciences**

**MOHAMMAD ALI JINNAH UNIVERSITY**

**2005**

**DEDICATED TO**

**THE HOLY PROPHET (P.B.U.H.)**

**THE GREATEST SOCIAL REFORMER**

**AND AFTERWARDS**

**MY PARENTS WHO HAD WISHED ME WITH THIS DISSERTATION**

# Abstract

In this dissertation we have given some new algorithms for blind equalization. In case of the equalization based on second order statistics (SOS), a sufficient condition for blind equalization is given and proved when the source symbols are coloured. Based on the given condition a new algorithm for coloured sources has been derived. The performance of the new algorithm as given by simulations is comparatively better than the previous works, especially at low signal to noise ratios. Then a new energy matching technique in SOS domain is introduced. It is very simple and computationally light scheme, which provides faster convergence. This scheme is valid only for equal energy source symbols. It outperforms the existing techniques of same category in simulations.

The work is then carried out for the domain of higher order statistics (HOS). In order to take advantage of both the SOS and HOS based techniques the hybrid approach of HOS-SOS is developed. The cost function of HOS based algorithms is modified by imposing the conditions based on SOS, which resulted into another new algorithm for binary source symbols. The algorithm gives excellent results and converges in three to five epochs only. This work is then further extended to cut down the computational complexity. Some over–stressing conditions with little contribution to the performance but possessing higher level of complexity have been removed. Energy constraint as given in SOS domain is also applied in the hybrid domain. Further modification is done to deal with complex source symbols e.g. 4–QAM. All this resulted in a much faster algorithm, which converges in one epoch only. Simulation results of each algorithm demonstrate the validity and superior performance of the new algorithms over the previous works. Some concluding remarks and future directions in this field have been proposed at the end.

# List of Publications and Submissions

1. A. Hussain, **A. Naveed**, I.M. Qureshi, "New Hybrid HOS-SOS Approach for Blind Equalization of Communication Channels," *IEE Electronics Letters*, vol. 41, no. 6, pp. 376-377, 2005.
2. **A. Naveed**, A. Hussain, I.M. Qureshi, T.S. Durrani, "Blind Restoration of Symbols from Binary and QAM Coloured Sources using Hybrid HOS-SOS Approach," Revised and Submitted to *IEEE Trans. Circuits and Systems–I.*
3. **A. Naveed**, A. Hussain, I.M. Qureshi, S. Fiori, "Blind Equalization of Channel for Equal Energy Sources: Energy Matching Approach," *IEE Electronics Letters*, vol. 42, Issue. 4, pp. 247-248, 2006.
4. **A. Naveed**, I.M. Qureshi, T.A. Cheema, M.A.S. Choudhri, " A New Approach to Blind Channel Equalization using Second Order Statistics," Accepted for publication in *Circuits, Systems & Signal Processing* journal.
5. **A. Naveed**, I.M. Qureshi, T.A. Cheema, A. Hussain, "Blind Equalization of Communication Channels with Equal Energy Sources Using a Combined HOS-SOS Approach," Accepted for publication in *IEEE International Conference on Engineering of Intelligent Systems 2006*.
6. T.A. Cheema, I.M. Qureshi, A. Jalil, and **A. Naveed**, "Artificial Neural Networks for Blur Identification and Restoration of Nonlinearly Degraded Images," *International Journal of Neural Systems*, vol. 11, vol. 5, pp. 455–461, 2001.
7. T.A. Cheema, I.M. Qureshi, A. Jalil, **A. Naveed**, "Blurred Image Restoration of Nonlinearly Degraded Images using ANN and Nonlinear ARMA Model,"

*Journal of Intelligent Systems*, vol. 11, no. 5, pp. 299–312, 2001.

8. A. Jalil, I.M. Qureshi, T.A. Cheema and **A. Naveed**, "Feature Extraction of Hand Written Characters by using Non-linear and Unsupervised Neural Networks," *Journal of Information Science and Engineering (JISE)*, vol. 21, no. 2, pp. 453-473, 2005.
9. **A. Naveed**, I.M. Qureshi, A. Jalil and T.A. Cheema**,** "Blind Equalization and Estimation of Channel using Artificial Neural Networks," IEEE 8th International Multitopic Conference INMIC 2004, Lahore, Pakistan, pp. 184-190, 2004.
10. T.A. Cheema, I.M. Qureshi, A. Jalil, and **A. Naveed**, "Blur and Image Restoration of Nonlinearly Degraded Images using Neural Networks Based on Modified ARMA Model," IEEE 8th International Multitopic Conference INMIC 2004, Lahore, Pakistan, pp. 102-107, 2004.
11. I.M. Qureshi, T.A. Cheema, **A. Naveed** and A. Jalil, "Genetic Algorithms Based Artificial Neural Networks for Blur Identification and Restoration of Degraded Images," *Pakistan Journal of Information and Technology(PJIT)*, vol. 2, no. 1, pp. 21-24, 2003.
12. A. Jalil, I.M. Qureshi, **A. Naveed** and T.A. Cheema, "Feature Extraction by using Non-linear and Unsupervised Neural Networks," *Pakistan Journal of Information and Technology (PJIT)*, vol. 2, no.1, pp. 40-43, 2003.
13. I.M. Qureshi, **A. Naveed**, T.A. Cheema and A. Jalil, "Artificial Neural Networks for Microstructure Analysis of Rolling Process," *Pakistan Journal of Information and Technology (PJIT)*, vol. 2, no.1, pp. 65-68, 2003.

# Acknowledgement

*Praise and thanks be to ALMIGHTY ALLAH, The compassionate and The merciful. Who gave us what we deserved and even more, what we did not deserve. All the knowledge emanates from the Almighty Lord.*

*Peace and prayers for His Prophet HAZRAT MUHAMMAD (S.A.W.) Whose incomparable life is the perfect model and mercy for all the creation.*

*I offer my sincere thanks to my supervisor, Dr. Ijaz Mansoor Qureshi, for his able guidance, useful suggestions and dynamic supervision throughout the research work. His personal interest and constructive criticism resulted in the completion of this dissertation.*

*I like to express my gratitude for scholarly guidance, co-operation and useful advices of my co-supervisor, Dr. Amir Hussain.*

*I would also like to forward my thanks to all my friends who encouraged me in the successful completion of this dissertation. Among my friends are Prof. Abdul Jalil, Dr. Tanveer Ahmed Cheema and Ch. Amir Saleem.*

*I also wish to express my feeling of gratitude to my brothers, and relatives, who prayed for my health and brilliance. Mr. Ghias Malik deserves my thanks for his support in the typing of this manuscript.*

*Last but not least, I offer my profound gratitude to my wife, Sumaira and son, Anas, for being patient and co-operative throughout my research period.*

(AQDAS NAVEED MALIK)

# Contents

# List of Figures

# Table of Abbreviations

| | |
|---|---|
| AAF | Adaptive Activation Function |
| ANN | Artificial Neural Networks |
| AWGN | Additive White Gaussian Noise |
| BE | Bayesian Estimator |
| BER | Bit Error Rate |
| CMA | Constant Modulus Algorithm |
| CR | Cross Relation |
| CRB | Cramer–Rao Bound |
| DMLE | Deterministic Maximum Likelihood Estimation |
| DSM | Deterministic Subspace Methods |
| FIR | Finite Impulse Response |
| GSM | Group Special Mobile |
| HF | High Frequency |
| HOC | Higher Order Cumulants |
| HOS | Higher Order Statistics |
| iid | independently identically distributed |
| ISI | Inter Symbol Interference |
| LSS | Least Squares Smoothing |
| MAI | Multiple Access Interference |
| MIMO | Multiple Input and Multiple Output |
| ML | Maximum Likelihood |

| | |
|---|---|
| MSE | Mean Square Error |
| NRMSISI | Normalized Root Mean Square Inter Symbols Interference |
| SER | Symbol Error Rate |
| SGSD | Stochastic Gradient Steepest Descent |
| SIMO | Single Input and Multiple Output |
| SISO | Single Input and Single Output |
| SMLE | Statistical Maximum Likelihood Estimation |
| SNR | Signal to Noise Ratio |
| SOS | Second Order Statistics |
| SSM | Statistical Subspace Methods |

# Table of Notations

| | |
|---|---|
| $s$ | Source symbols |
| $x$ | Received symbols at the channel output |
| $y$ | Equalizer output symbols |
| $\mathbf{h}$ | Channel vector |
| $\mathbf{s}$ | Source symbols vetor |
| $\mathbf{w}$ | Equalizer vector |
| $\mathbf{x}$ | Received symbols vector |
| $\mathbf{y}$ | Equalized symbols vector |
| $\mathbf{H}$ | Channel matrix |
| $\mathbf{R}$ | Correlation matrix |
| $\mathbf{W}$ | Equalizer matrix |
| $\alpha$ | Complex constant with unit norm |
| $\eta$, $\mu$ | Step sizes |

We shall use the notation $B^*$, $B^{\mathrm{T}}$, $B^{\mathrm{H}}$, $\hat{B}$ and $E[B]$ for complex conjugate, transpose, hermitian, estimated and expected value of $B$, respectively. Bold and capital symbols, such as $\mathbf{B}$, stand for matrix, bold and small symbols, such as $\mathbf{b}$, stand for vector and unbold symbols, such as $b$, stand for scalar quantities.

# Chapter 1

# Introduction

## 1.1 Introduction

There is an ever growing demand for high quality and high speed wireless communication. One of its major limiting factors is intersymbol interference (ISI). ISI may be due to one or more of the following factors. These factors could be the frequency selective characteristics of the channel, time varying multipath propagation which is prominent in mobile communication, carrier phase jitter, symbol clock residual jitter and limited channel bandwidth.

ISI is the major problem for single input and single output (SISO) and single input and multiple output (SIMO) channels. In case of multiple input and multiple output (MIMO) channels, multiple access interference (MAI) is also equally important problem to reckon with. Since in this dissertation, MIMO channels are not being considered, therefore ISI and its solutions are of major concern.

If the *a priori* knowledge of the channel is available, one can use this information for carrying out the equalization. However, in practical world most of the communication is through the unguided wireless medium for which the *a priori* knowledge of the channel does not exist. Hence simple equalization is practically of no use. The standard adaptive approach has been through the use of training sequences. This does waste a fraction of

transmission time and hence decrease the communication throughput, but the effect may be fairly insignificant in case of time invariant channels. However, for time-varying channels, the loss of throughput is fairly significant. For example, for high frequency (HF) communication, it can be as high as fifty percent of the total possible throughput.

The group special mobile (GSM) system, which has considerable overhead for training needs blind equalization. Some of the computer networks in which terminals need to be linked asynchronously with the central computer need blind equalization because training is almost impossible. Some other areas for blind equalization are high definition television, image restoration, geosciences, acoustic, speech etc.

In order to save the resource bandwidth and avoid the need of training phase, extensive research has gone in the techniques of unsupervised adaptive or blind equalization, which does not require any training sequences. Instead, the statistics of the transmitted symbols is being used to carry out equalization at the receiver without access to the symbols and the knowledge of the channel. While in case of the standard adaptive equalizer, the equalized output sequence is made close to the source symbol sequence in the mean square sense, in case of the blind equalizers, the statistics of the equalized output sequence are made to approach the statistics of the transmitted input sequence.

The channel dealt with in this problem is modeled as a finite impulse response (FIR) filter having finite length. Its coefficients formulate the finite impulse response of the channel. It would be easy to carry out the inverse filtering using zero forcing algorithm provided the channel is minimum phase, that is, all the zeros of the transfer function of the channel lie within the unit circle. Practical channels considered in this dissertation are nonminimum phase, resulting in an inverse filter that is no more stable.

Henceforth, one has to come up with more sophisticated techniques which are known as blind equalization techniques. The approaches and algorithms developed for blind equalization can be put broadly into two categories firstly, those based on higher order statistics (HOS) and secondly, those based on second order statistics (SOS) called SOS techniques.

## 1.2 HOS Based Approaches

The HOS based approaches have been the first in fashion for blind equalization of channels. They fall in two categories based on how the HOS of the received signal is exploited. First is the explicit HOS based category in which higher order cumulants (HOC) or their discrete Fourier transform called polyspectra, have been used. The first ones to propose in this direction were Pan et al. [1] and Hatzinakos et al. [2]. A moving average process models the received signal. The trispectrum of the received signal is used to identify the multipath channels. The algorithm provides exact identification of a nonminimum phase channels, whenever the HOS and trispectrum of the observed signal can be estimated accurately. This explicit HOS based algorithm, however, suffers from computational complexity and slow convergence as compared to other techniques. Since the received signal is sampled at baud rate, it is sensitive to timing recovery, unknown phase jitter and frequency offset. Lastly, if the noise is non–Gaussian, it may affect the performance of this HOC based approaches. Some of the works based on explicit HOS were [1]–[6]. The explicit SISO methods using HOS include the inverse filter criteria (IFC)-based algorithm [7]–[14], the super–exponential algorithm [15]–[20] and polyspectra –based algorithms [21]–[22].

Second is the implicit HOS based approach. HOS of the received signal is exploited indirectly. It is also called Bussgang type algorithm. When this algorithm converges in the mean square sense, the deconvolved signal exhibits Bussgang statistics, hence the name. The first major work in this direction was that of Sato [25] in which he used a nonconvex cost function that was minimized. This idea was implemented on the blind equalization of M–ary pulse amplitude modulation (PAM) systems. This idea was further developed and generalized by Godard [26] to give a class of constant modulus algorithm (CMA) for blind equalization [27]–[28]. In digital communication, the CMA has been a widely used approach to alleviate the ISI effect induced by telephone, cable or radio channels [29]–[30]. Treichler and Agee [31], Benveniste and Goursat [32], Picci and Prati [33] and [34]–[40] are some of the further developments in this very direction.

All the above are special cases of Bussgang algorithm. They are based on implicit HOS and are derived via some optimization criterion that involves HOS of the observation indirectly. They usually have a nonconvex cost function which is being minimized by stochastic gradient based algorithm.

The major disadvantage of these implicit HOS based algorithms is their potential for converging to local minimum [41], [42]. It has been pointed out in [42] that global convergence can be jeopardized if the channel has finite impulse response. Shalvi and Weinstein [43] have proposed an optimization criterion that ensures global optimization wherever ideal equalizers exist. However, these techniques are still sensitive to timing jitter and exhibit slow rate of convergence.

## 1.3 SOS Based Approaches

In the fast changing world of cellular communication, the methods with slow rate of convergence and greater computational complexities will be less popular comparatively. There has been a trend of shifting from HOS based algorithms to SOS based algorithms after the seminal work of Gardener [44] and Tong et al.[45], [46].

They explored the cyclostationary properties of an oversampled communication signal to allow the blind channel estimation by using SOS of the channel output. Ever since, people have come up with promising algorithms and techniques for their research problems, based on SOS [47]–[57]. Many developments have taken place in identifying nonminimum phase channels without using HOS. Following is a brief discussion on the subspace and maximum likelihood methods.

### 1.3.1 Subspace Methods

These methods have closed forms. The channel vector is represented in one-dimensional subspace of observation statistics. Then with $\mathbf{Q}$ being the matrix representing the noise structure, we optimize the quadratic cost function [44]

$$\hat{\mathbf{h}} = \arg\min \mathbf{h}^H \mathbf{Q} \mathbf{h} \qquad \mathbf{h} \in S$$

where S is a set that specifies the domain of the channel vector $\mathbf{h}$. Subspace methods are mainly of two types, deterministic subspace methods (DSM) and statistical subspace methods (SSM).

The DSM has no assumption about the structure of the input source. They assume noise to be white with zero mean covariance $\sigma^2$. They also assume the knowledge of the

channel order. Both these assumptions are impractical. However, the noise variance and channel order detection may be done jointly using singular values of the estimated covariance matrix [47]. Xu et al. [59] and Hua et al. [60] have given necessary and sufficient condition for the identifiability of the channel which is mainly the coprime nature of the subchannels.

In DSM, there are three approaches being followed for identification. First one is the cross relation (CR) approach. [61]–[64], [90], [113]. This approach exploits the multichannel structure. Its adaptive approach with artificial neural networks (ANN) was given by [65] and [66]. CR method is very effective for small data sample at high signal to noise ratio (SNR). In [90] it has been shown that CR method combined with maximum likelihood (ML) approach gives performance close to Cramer–Rao bound (CRB) [91]. The drawback of this CR approach is that channel order $L$ cannot be overestimated. Moreover, this algorithm may also be biased for finite samples.

The second approach is the noise subspace approach given by Moulines [104], [105]. This approach forces the signal space to have a block Toeplitz structure. This approach is similar to CR method but slightly more complex. It also requires the knowledge of the channel order.

The third approach is the least squares smoothing (LSS) approach. Its key idea hinges on the relation between the input and the observation spaces. Apart from the fact that a channel estimation problem has been converted to LSS problem, the joint order detection and channel identification can be done in a fashion to minimize the smoothing error [67], [68]. This approach is the only one that enables channel identification with the knowledge of the upper bound of the channel order only.

The SSM does assume that the source is a random sequence with known SOS. It assumes source sequences as zero mean, white with unit variance. Similarly the noise sequence is zero mean, with variance $\sigma^2$ and uncorrelated to source sequence. Channel order is assumed to be known, but some of the SSM require only the upper bound of the channel order. Once SOS at the output of the channel are known, the channel can be uniquely determined provided the subchannels are coprime.

The first SSM approach using SOS for blind channel estimation was given by Tong et al. in [69], [70]. They gave the two step closed form algorithm for identification of the channel. It was shown that once the correlation matrices with lag 0 and lag 1 at the output of the channel are known, the channel parameter matrix is identifiable.

The second SSM approach using SOS was given by Slock [84], [85], [86] followed by Abed–Meriam [87] and Valcarce et al. [88]. They used the linear prediction formulation of the multichannel problem. The algorithm uses SOS of the received signal. It does not require the exact order of the channel. The main drawback of linear prediction approach is that it is not robust in the presence of noise, since its algorithm is derived from the noiseless model. However, using autocorrelation techniques for estimation of parameters, people have decreased the effect of noise by subtracting the term related to noise.

### 1.3.2 Maximum Likelihood Methods

Maximum likelihood (ML) method is also one of the most popular parameter estimation method. Asymptotically under certain regularity conditions, the performance of ML estimation approaches the CRB. Unlike subspace based approach, the ML

approach usually cannot be obtained in closed form. Moreover, this approach suffers from the existence of local minima. However, the approach has been made more efficient by using suboptimal approach as initial procedure. The ML–based blind channel estimation can be of two types. Statistical ML estimation (SMLE) and deterministic ML estimation (DMLE). In SMLE the input sequence is assumed to be random but with known statistics. The only unknown parameter is the channel vector. The dimension of this unknown vector is fixed with respect to the data size. The SMLE depends on the availability and the evaluation of the likelihood function. If the input is independently identically distributed (iid) and noise is zero mean and additive white Gaussian noise (AWGN), then a theorem states that channel is identifiable by likelihood function if, and only if, one of the two conditions is satisfied, i.e., Either input symbols are non Gaussian or subchannels are coprime [72]. References [72]–[78] are different endeavors in this direction.

In case of DMLE, the input sequence is also a part of the unknown parameters alongwith the channel. Thus the dimension of the parameters increases with the size of the observed data. The iterative quadratic ML approach proposed by Bresler and Macovski [79], and two-step ML approach proposed by Hua [79], [79] fall in this category. Slock [86], Harikuma and Bresler [79] and [80]–[83] are some of the works on DMLE type of algorithms.

Considering the computational intensity of the subspace methods and the ML methods, researchers have also considered neural network based direct signal estimation techniques [94]. Fang et al. [95]–[96] has used linear neural network for blind equalization. In [96] they have used statistics matching of the symbols at the input of the

transmitter and the output of the equalizer. The square of the elements of the difference matrix between the correlation matrices at both ends of the channel has been minimized using stochastic gradient algorithm. This gives the updating formula for the weights of the equalizer, which is like a linear neural network. This approach has been discussed in detail in chapter 2, and would be of our interest in contributions.

## 1.4 Contributions of the Dissertation

The contributions given in this dissertation are divided into two categories. First category deals with the SOS–based approaches, whereas the second category provides the combined HOS–SOS approaches for blind equalization.

In case of SOS domain it has been proved that even for coloured sources, if the correlation matrix of the received symbols with lag greater than or equal to 1 is equal to the corresponding correlation matrix of the source symbols, then there exists a unique equalizer. This equalizer will give us the equalized symbols when applied to the channel output vectors. The equalized symbols will be the estimates of the original source symbols multiplied by a complex constant of unit norm. Based on this condition a cost function has been developed. Minimization of this cost function gives us the learning algorithm. The new algorithm has shown robustness against different channels and different values of SNR. The results are much improved espatially at low SNR as compared to the previous work in the literature.

Another algorithm in the SOS–domain has been given for the equal energy source symbols. It has been proved that if the input symbols to the channel have equal energies, then the blind equalization can be achieved by making the energies of the output symbols

equal to that of input symbols. This scheme is valid for both the coloured and white source symbols constellations. Thus oversampling of channel output plus energy mayching of the source and equalizer symbols ended up in another very simple and useful algorithm providing excellent convergence in only one epoch. The bit error rate (BER) for the new algorithm is almost zero even at SNR=20db. The algorithm performs well for the minimum number of received symbols.

In case of the hybrid or combined HOS-SOS approaches we have mainly focused on the Bussgang algorithm which is basically implicit HOS based technique for blind equalization. We have employed the SOS based statistics matching in addition to the modifications given by the modified Bussgang algorithm [98]. No oversampling is used in this case. The SOS of the transmitted symbols at the input of the channel must be equal to that of the received symbols at the output of the channel. Initially we have used the correlation matrices with lag zero and lag one for SOS matching. The norm of the weights of the equalizer was constrained to be a fixed value. The result of this combined HOS-SOS approach was a faster convergence, achieved in about 3-5 epochs. The new algorithm proved to be equally valid for the channels with and without norm one. The limitation of the original and modified Bussgang algorithms is their assumption about channel norm equal to one.

The final modification was valid only for those source symbols constellations in which the symbols have equal energy. Another term was added to the cost function called as energy matching term for input and output symbols. Moreover, for statistics matching only one correlation matrix with lag zero was used. In case of modified Bussgang algorithm given in [98] a constraint on the norm of the weights of equalizer has been

used. Such type of constraint is required to avoid the ill – conditioning of equalizer. In case of the new algorithm this constraint has been removed. The removal of such constraint was not possible in the absence of the correlation matching term. This modification has further cut down the computational complexity. After these modifications the new algorithm is providing the convergence in one epoch only. The computational cost has also been reduced.

## 1.5 Organization of the Dissertation

The remaining portion of this dissertation is organized as follows: Chapter 2 gives the original Bussgang algorithm as dealt by Bellini [97] and then the modified algorithm given by Fiori [98]. It covers all the essential mathematical details and principles. This chapter also contains the detailed SOS approaches using statistics matching and linear neural networks as direct blind equalizers.

Chapter 3 contains two new algorithms in SOS domain. First one is the genlization of SOS based algorithms for colour and white sources, with less complexity. The second one provides equalization for equal energy sources with requirement of energy matching and oversampling only.

Chapter 4 contains the hybrid HOS-SOS approaches. This can be considered as an important improvement to the conventional Bussgang type algorithms. There are about three to four modifications done in a manner that we can be justified to call it a hybrid HOS–SOS approach. Enough simulations have been given to verify the validity of the new proposed algorithms. Results of the new algorithms has been compared with different algorithms available in the literature.

Chapter 5 concludes the thesis. It gives summary of the results in SOS and hybrid HOS-SOS approaches. Some applications are suggested and also some future directions of research have been proposed.

# Chapter 2

# Mathematical Formulation of Implicit HOS and SOS Based Algorithms

In this chapter, we have given the actual Bussgang algorithm which is based on implicit HOS. This is the work of Bellini [97]. Then we have presented the modifications to the Bussgang algorithm put forward by Fiori [98]. Its simulation and results are given in chapter 4 for comparison with the new proposed algorithms for hybrid HOS–SOS approach.

The second half of this chapter deals with the algorithm based on SOS. We have focussed on neural network approach whose learning algorithm is based on the matching of statistics of the input symbols and the symbols at the output of the equalizer. This approach is given by Fang [95]–[96]. Since our new algorithm in SOS domain is related to this work, we have given its complete formulation along with relevant details.

## 2.1 Implicit HOS Based Algorithms

In this section first we discuss the conventional Bussgang algorithm as given by Bellini [97]. Then we discuss the modification of this Bussgang algorithm as given by

Fiori [98]. Since our present work is also in this direction therefore we have also given its sufficient mathematical details.

### 2.1.1 Conventional Bussgang algorithm

We first present the basic theme of blind equalization with the Bussgang algorithm. The discrete impulse response $\{h_n\}$ of Additive White Gaussian Noise (AWGN) channel is unknown and is assumed to have finite support of length $L_h$ i.e. $\mathbf{h} = \{h_1 \quad h_2 \quad \cdots \quad h_{L_h-1}\}$. It is assumed that $\sum_k h_k^2 = 1$, which implies the use of automatic gain control. This keeps the variance of the channel output, $x(n)$, constant. In general, the channel is noncausal $h_k \neq 0$ for $k < 0$ [101]. The input sequence, $\{s(n)\}$, to the channel comes from zero mean, stationary and independent and identically distributed (*ii*d) source such that $E\{s(k)s(l)\} = \delta_{kl}$.

The output symbols, $x(n)$, from the channel as depicted in Fig 2.1 is given as

$$x(n) = \sum_{k=0}^{L_h-1} h_k s(n-k) + \omega(n) \tag{2.1.1}$$

where $\omega(n)$ is zero mean, AWGN. It is neglected in the mathematical analysis because ISI dominates the degradation due to additive noise. However, we shall use AWGN noise in simulation and results. Figure 2.2 represents the block diagram for the conventional Bussgang algorithm. The transversal filter has $m$ number of tap weights. The output of the transversal filter is given as

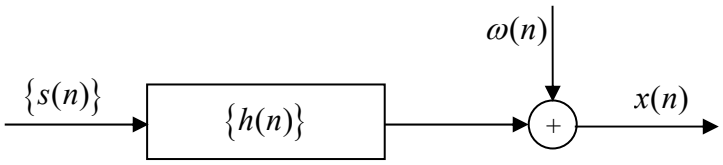



**Fig. 2.1 AWGN channel model.**

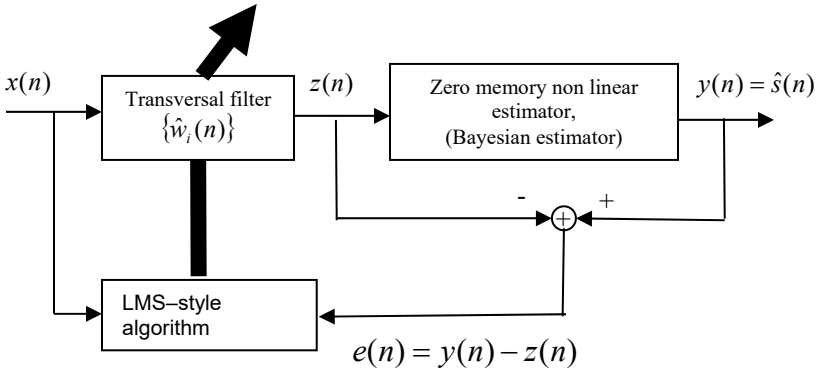



**Fig. 2.2 Block diagram for the Bussgang algorithm.**

$$\begin{aligned} z(k) &= \sum_i w_i x(i-k) + \sum_i \left(\hat{w}_i - w_i\right) x(i-k) \\ &= s(k) + v(k) \end{aligned} \tag{2.1.2}$$

where $w_i$ is the $i^{\text{th}}$ optimal tap weights of the transversal filter and $\hat{w}_i$ are their estimate. $v(k)$ is residual ISI, called the convolution noise. It is AWGN with zero mean. Moreover, it is statistically independent of the source sequence and its time-varying variance, $\sigma^2(n)$, decreases as the adaptive algorithm converges [101]

The zero memory nonlinear estimator produces $y(n)$ , which is the estimate of $s(n)$ ( $y(n) = \hat{s}(n)$ ). It is given as

$$\hat{s}(n) = g\left(z(n)\right) \tag{2.1.3}$$

The estimation error is taken as $e(n)$ such that

$$\begin{aligned} e(n) &= \hat{s}(n) - z(n) \\ &= g\left(z(n)\right) - z(n) \end{aligned} \tag{2.1.4}$$

The cost function is given as

$$\begin{aligned} J(n) &= \frac{1}{2} E\left[e^2(n)\right] \\ &= \frac{1}{2} E\left[\left(g\left(z(n)\right) - z(n)\right)^2\right] \end{aligned} \tag{2.1.5}$$

The updating formula for the weights of the transversal filter is given as:

$$\hat{w}_i(n+1) = \hat{w}_i(n) + \mu x(n-i) e(n) \qquad i = 0, \pm 1, \cdots, \pm L \tag{2.1.6}$$

where $\mu$ is the step size parameter and $x(n-i)$ is the $i^{th}$ tap input of the transversal filter at iteration $n$. It has been assumed that the input symbols, $\{s(n)\}$, have uniform distribution with zero mean and unit variance. That is, the PDF for $\{s_n\}$ is

$$f_s(s) = \begin{cases} \left(2\sqrt{3}\right)^{-1} & -\sqrt{3} \le s \le \sqrt{3} \\ 0 & \text{otherwise} \end{cases}$$

In the above scenario, result given by Bellini [97] for output symbols as estimate of input symbols, $\hat{s}$, is given as

$$\begin{aligned} \hat{s} &= E\left[s|z\right] \\ &= \int_{-\infty}^{+\infty} s f_s\left(s|z\right) ds \end{aligned}$$

Using the Bayes rule

$$f_s\left(s|z\right) = \frac{f_z\left(z|s\right) f_s\left(s\right)}{f_z\left(z\right)}$$

therefore,

$$\begin{aligned} \hat{s} &= \frac{1}{f_z\left(z\right)} \int_{-\infty}^{+\infty} s f_z\left(z|s\right) f_s\left(s\right) ds \\ &= \frac{z}{c_0} - \frac{\sigma}{c_0} \frac{Z(z_1) - Z(z_2)}{Q(z_1) - Q(z_2)} \end{aligned} \tag{2.1.7}$$

where the scaling factor $c_0$ is slightly smaller than unity. Its purpose is to keep the mean square value, $E\left[z^2\right]$, equal to unity. The variables

$$z_1 = \frac{1}{\sigma}\left(z + \sqrt{3}c_0\right)$$

and

$$z_2 = \frac{1}{\sigma}\left(z - \sqrt{3}c_0\right)$$

The functions $$Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

and

$$Q(z) = \frac{1}{\sqrt{2\pi}} \int_z^{\infty} e^{-q^2/2}\, dq$$

## 2.1.2 Modified Bussgang Algorithm

Fiori [98] has given the modification to the Bussgang algorithm. First modification is the constraint on the weights of the transversal filter $(\mathbf{w}^T\mathbf{w} = k^2)$. Second modification is the replacement of zero memory nonlinear estimator (Bayesian estimator)

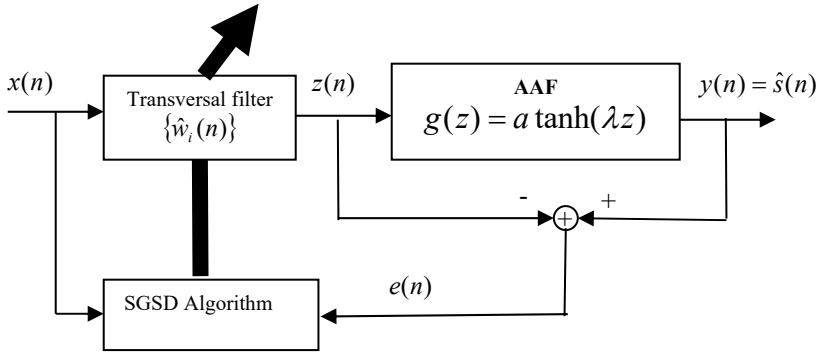



**Fig. 2.3 Block diagram for the modified Bussgang algorithm.**

by an adaptive activation function (AAF). Due to the first modification the cost function becomes

$$J_{old}(n) = \frac{1}{2} E\left[\left(g(z) - z\right)^2\right] + \lambda\left(\mathbf{w}^T\mathbf{w} - \kappa^2\right) \tag{2.1.8}$$

where $\boldsymbol{w}$ is the weight vector and $\lambda$ is a Lagrange multiplier, combined with the constant $\kappa^2$ it incorporates the constraint of norm on the weights. The Bayesian estimator is replaced by the AAF given by

$$g(z) \triangleq a\tanh(\lambda z) \tag{2.1.9}$$

where both $a$ and $\lambda$ are adaptive. Using LMS algorithm on $J_{old}(n)$ Fiori gets updating formulae for weights $\mathbf{w}$, $a$ and $\lambda$ [100]. Fiori has reported the fact that the product $a\lambda$ converges asymptotically to some constant $\Lambda$. Thus the updating of $a$ and $\lambda$ has been incorporated in one single updating formula for $\lambda$ given as

$$(\Delta\lambda)_{old} = -\eta\left[g(z)-z\right]\left[-\frac{g(z)}{\lambda}+\frac{\Lambda}{\lambda}z-\frac{\lambda z}{\Lambda}g^2(z)\right] \tag{2.1.10}$$

where $\eta$ is step size.

The updating of weights using SGSD algorithm is given as,

$$(\Delta\mathbf{w})_{old} = -\eta_w\left[\left(\Lambda-1-\frac{\lambda^2}{\Lambda}g^2(z)\right)\left(g(z)-z\right)\mathbf{x}\right] \tag{2.1.11}$$

The constraint on the norms is not implemented automatically through the Lagrangian, but instead after every iteration the weights $\frac{\kappa\mathbf{w}}{\|\mathbf{w}\|} \to \mathbf{w}$. The equations (2.1.9) to (2.1.11) formulate the gradient-based blind equalization method given by Fiori [98]. The simulations and results for the work of Fiori [98] have been done in chapter 4. Comparisons have been done with the proposed hybrid HOS-SOS approach.

## 2.2 SOS Based Algorithm

This category exploits only Second Order Statistics (SOS) of the received signal for recovery of the source symbols. This approach has gained importance due to the seminal work by Gardner [44] and Tong [45]–[46]. Most of the algorithms based on SOS use either subspace methods or ML methods [45]–[88]. However, they are computationally heavy as they are like batch processing algorithms. An alternative approach uses linear ANN for direct blind equalization, [95], [96], [106]–[110]. In this case, online learning is performed such that the output symbols not only produce the statistics of the original source symbols but they are in fact the original symbols multiplied by a unit norm constant.

### 2.2.1 Problem Statement

A digital communication model is shown in Fig. 2.4. In this case, $\{s_j\}$ presents the sequence of source symbols applied at the input of the channel. The source symbols are transmitted into the channel at baud rate $1/T$, i.e. $T$ is the symbol interval, $t$ denotes the continuous time, $x(t)$ is the received symbol, $v(t)$ is zero mean additive white Gaussian noise (AWGN). The overall channel response is given by $h(t)$. Though the transmitted symbols sequence at the input of the channel is discrete, the received channel output is continuous in time due to the convolution of the channel response with the input symbols. The remaining two blocks on right hand side of this figure are used for equalization. The most general equation for the received symbols at the output of the

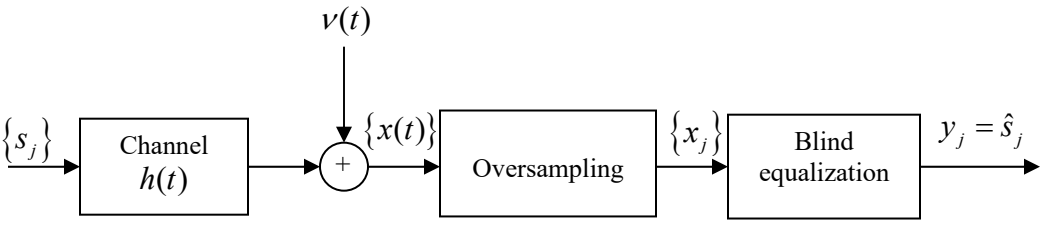



**Fig. 2.4 Block diagram for equalizer with SOS scheme.**

time-invariant channel in a digital communication system can be written in the form given by:

$$x(t) = \sum_{k=-\infty}^{\infty} s_k h(t - kT) + v(t) \tag{2.2.1}$$

In order to solve the problem of blind equalization, following assumptions are made.

1. Symbol interval $T$ is known and is a multiple of the sampling period $\tau$, i.e. $T = d\tau$, where $d$ is a positive integer.
2. Channel is like an FIR filter with finite impulse response $h(t)$, such that, $h(t) = 0$ for $t < 0$ and $t > L_c$, where $L_c = nT$ .
3. $\{s_j\}$ is a sequence of stationary white source symbols with zero mean, such that $E\left[ s_k s_l^* \right] = \delta_{kl}$ .

   $v(t)$ is white with zero mean and uncorrelated to $\{s_j\}$ . It is neglected at first stage because of its negligible contribution as compared to ISI [101]. However, its effect is considered on the performance in the simulations.

   Taking into account the assumptions 2 and 4, (2.2.1) becomes

$$x(t) = \sum_{k=\left\lceil \frac{t-L_c}{T} \right\rceil + 1}^{\left\lceil \frac{t}{T} \right\rceil} s_k h(t - kT) \tag{2.2.2}$$

where $\lceil u \rceil$ means highest integer less than or equal to $u$. For $x(t)$ which is sampled in any general interval, $(q-1)T$ to $qT$, we get

$$x(i\tau + (q-1)T) = \sum_{k=q-n}^{q-1} s_k h((\frac{i}{d} - k + q - 1)T) \qquad i = 0,..,d-1 \tag{2.2.3}$$

Now we consider an observation interval of length $mT$, where $m$ is a positive integer. The received signal in this observation interval can be expressed in vector form as

$$\mathbf{x}(i) = \mathbf{H}\mathbf{s}(i) \qquad i = 0,1,..,md-1 \tag{2.2.4}$$

where

$$\mathbf{x}(i) = \begin{bmatrix} x(iT) & x(iT+\tau) & \cdots & x(iT+(md-1)\tau) \end{bmatrix}^{\mathrm{T}} \tag{2.2.5}$$

$$\mathbf{s}(i) = \begin{bmatrix} s_{i-n+1} & s_{i-n+2} & \cdots & s_{i+m-1} \end{bmatrix}^{\mathrm{T}} \tag{2.2.6}$$

and

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 & \cdots & \mathbf{h}_n & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{h}_1 & \mathbf{h}_2 & \cdots & \mathbf{h}_n & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{h}_1 & \mathbf{h}_2 & \cdots & \mathbf{h}_n \end{bmatrix} \tag{2.2.7}$$

where

$$\mathbf{h}_l = \begin{bmatrix} h\left[(n-l)T\right] & h\left[(n-l+\frac{1}{d})T\right] & \cdots & h\left[(n-l+\frac{d-1}{d})T\right] \end{bmatrix}^T \tag{2.2.8}$$

and $\mathbf{0}$ is a vector of zeros.

The channel diversity, which is the requirement for blind equalization based on techniques using only SOS [45], has been exploited here. The duration of observation interval and over-sampling is selected in such a way that (2.2.4) becomes over-determined, i.e. the number of rows of matrix $\mathbf{H}$ becomes greater than or equal to the number of columns, $md \geq m+n-1$. This makes $\mathbf{H}$ a full-column rank matrix to assure the solvability of the blind equalization problem in accordance with the condition given by Tong et al. [45]. At this stage if $\mathbf{H}$ is known, then the problem is quite simple. The estimated output can be taken as

$$\hat{\mathbf{s}}(i) = (\mathbf{H}^{\mathrm{H}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{H}}\mathbf{x}(i) \tag{2.2.9}$$

where $(\mathbf{H}^{\mathrm{H}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{H}}$ is the pseudo-inverse of $\mathbf{H}$. However, in case of blind equalization, $\mathbf{H}$ is unknown.

The autocorrelation function of the source symbols, $\mathbf{s}(i)$ is defined by

$$\mathbf{R}_s(k) = E\left[\mathbf{s}(i)\mathbf{s}^H(i+k)\right] \tag{2.2.10}$$

Hence

$$\mathbf{R}_s(0) = \mathbf{I}$$
$$\mathbf{R}_s(n) = \mathbf{J}^n \quad \text{for} \quad n > 0$$

where $\mathbf{I}$ is the identity matrix, and

$$\mathbf{J} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}$$

is a shifting matrix.

Similarly the autocorrelation function of the equalizer output $\mathbf{y}(i)$, is defined as

$$\mathbf{R}_y(k) = E\left[\mathbf{y}(j)\mathbf{y}^{\mathrm{H}}(j+k)\right] \tag{2.2.11}$$

The following theorem gives a necessary and sufficient condition for the blind equalization when source is white [96].

***Theorem 2.1:***

Suppose $\mathbf{H}$ and $\mathbf{s}(i)$ satisfy the linear model given by (2.2.4). $\mathbf{s}(i)$ are from white source and $\mathbf{H}$ is a full column rank matrix. There exists a linear transformation $\mathbf{W}$ i.e.

$$\mathbf{y}(i) = \mathbf{W}\mathbf{x}(i) \tag{2.2.12}$$

such that for $|\alpha| = 1$

$$\mathbf{WH} = \alpha I \tag{2.2.13}$$

if, and only if, the following two conditions are satisfied

$$\mathbf{R}_{\mathrm{y}}(0) = \mathbf{I} \tag{2.2.14}$$

$$\mathbf{R}_{\mathrm{y}}(1) = \mathbf{J} \tag{2.2.15}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{J}$ is the shifting matrix given above. Please see Appendix–A for the proof.

## 2.2.2 Neural Network and Learning Algorithm

The block diagram of the linear ANN as proposed by Fang et al. [96] is shown in Fig. 2.5. The received symbols sequence $\{x_j\}$ at the output of the channel is applied to the input layer $L_1$ of the proposed ANN as sequence of vectors $\mathbf{x}(j)$, each having length $md$. The network layer $L_2$ provides the equalized output $\mathbf{y}(j)$, which is an estimate of

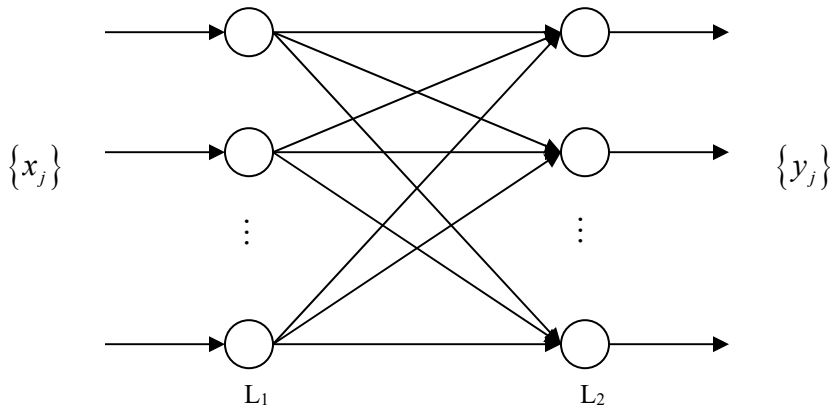



**Fig. 2.5 Proposed ANN for equalization.**

the source symbols. In order to establish the learning algorithm for updating the weights of the matrix between layers $L_1$ and $L_2$ of the network, the following cost function, $J(W)$, is to be minimized in the mean square sense. The bigradient algorithm for updating of the weights can be derived by minimizing the cost function.

$$J(W) = \sum_{m=1}^{d} \sum_{n=1}^{d} \left\{ \left| \mathbf{A}_{nm} \right|^2 + \left| \mathbf{B}_{nm} \right|^2 \right\} \tag{2.2.16}$$

where $\mathbf{A}_{nm}$ and $\mathbf{B}_{nm}$ are the $(n,m)$ elements of matrices $\mathbf{A}$ and $\mathbf{B}$, respectively. These matrices are given as

$$\mathbf{A}_{nm} = \left\{ \mathbf{R}_y(1) - \mathbf{J} \right\}_{nm} \tag{2.2.17}$$

$$\mathbf{B}_{nm} = \left\{ \mathbf{R}_y(0) - \mathbf{I} \right\}_{nm} \tag{2.2.18}$$

Both the terms of the cost function are based on the requirement, that the statistics at the destination must approach the statistics at the source, i.e., $\mathbf{R}_y(n) \to \mathbf{R}_s(n)$ for $n = 0,1,...$ The elements of weight matrix are updated according to SGSD algorithm

$$\Delta \mathbf{W} = -\eta \frac{\partial J(W)}{\partial \mathbf{W}} \tag{2.2.19}$$

where

$$\frac{\partial J(W)}{\partial \mathbf{W}_{kl}} = \sum_n \sum_m \left\{ \frac{\partial \mathbf{A}_{nm}}{\partial \mathbf{W}_{kl}} \mathbf{A}^*_{nm} + \mathbf{A}_{nm} \frac{\partial \mathbf{A}^*_{nm}}{\partial \mathbf{W}_{kl}} + \frac{\partial \mathbf{B}_{nm}}{\partial \mathbf{W}_{kl}} \mathbf{B}^*_{nm} + \mathbf{B}_{nm} \frac{\partial \mathbf{B}^*_{nm}}{\partial \mathbf{W}_{kl}} \right\}$$

where $\mathbf{W}_{kl}$ is the $(k,l)$ element of matrix $\mathbf{W}$.

If we wish to carry out the online learning process i.e. updating of weights for every new vector $\mathbf{x}(j)$ coming in, we remove the expectation operator

$$(\Delta \mathbf{W})_{ol} = -\eta \begin{bmatrix} \left\{ \mathbf{y}(j)\mathbf{y}^{\mathrm{H}}(j+1) - \mathbf{J} \right\}^{\mathrm{H}} \mathbf{W}\mathbf{R}_x(1) \\ + \left\{ \mathbf{y}(j)\mathbf{y}^{\mathrm{H}}(j+1) - \mathbf{J} \right\} \mathbf{W}\mathbf{R}_x^{\mathrm{H}}(1) \\ +2 \left\{ \mathbf{y}(j)\mathbf{y}^{\mathrm{H}}(j) - \mathbf{I} \right\} \mathbf{W}\mathbf{R}_x(0) \end{bmatrix} \tag{2.2.20}$$

where the subscript '*ol*' stands for online. Please see Appendix–B for detailed derivations. If weights are to be updated by batch processing then we shall carry on with the expectation operator, which gives the following formula for updating the weights:

$$\begin{aligned} (\Delta \mathbf{W})_{bp} = -\eta_A & \begin{bmatrix} \left\{ \mathbf{W}\mathbf{R}_x^{\mathrm{H}}(1)\mathbf{W}^{\mathrm{H}} - \mathbf{J}^{\mathrm{H}} \right\} \mathbf{W}\mathbf{R}_x(1) \\ + \left\{ \mathbf{W}\mathbf{R}_x(1)\mathbf{W}^{\mathrm{H}} - \mathbf{J} \right\} \mathbf{W}\mathbf{R}_x^{\mathrm{H}}(1) \end{bmatrix} \\ & -\eta_B \left[ \left\{ \mathbf{W}\mathbf{R}_x(0)\mathbf{W}^{\mathrm{H}} - \mathbf{I} \right\} \mathbf{W}\mathbf{R}_x(0) \right] \end{aligned} \tag{2.2.21}$$

where the subscript '*bp*' stands for batch processing. Please see Appendix–B for details. The simulations and results for the work of Fang [95] have been carried out in chapter 3 and 4 and the comparisons have been done with our proposed algorithm.

# Chapter 3

# Blind Equalization of Channel Using SOS With Coloured Sources

The contributions in the SOS domain are given in this chapter. In the first part we used an equalizer which looks like the linear ANN as given before in Fig. 2.5, but the original symbols are from coloured source. We have proved that even in case of coloured sources if the statistics of the symbols is kept the same at the source and the destination by using only one correlation matrix with lag equal to or greater than one, then there exists a unique equalization matrix which provides us the original symbols at the output scaled by a complex constant of unit norm. Based on the condition for blind equalization, we develop a cost function, minimization of which gives us a learning algorithm. We have simulated two different 2-ray multipath channels, one real and the other complex, which shows the validity and superior performance of the new algorithm. It has been compared with the work of Afkhamie [111] and Valcarce [88] for coloured sources. Our results excel by far at low SNR and with less number of symbols, but competitive at high SNR and large number of symbols.

In the second part the energy matching approach for the over sampled case is done. It has been proved that the energy matching of the received symbols with the source symbols provide the equalization. For this very proposed algorithm, we have achieved

convergence in one epoch. The computational cost per epoch is also minimum in comparison with the other algorithms of SOS domain.

## 3.1 Theoretical Framework

As mentioned earlier the effect of ISI as compared to AWGN is so dominant that for the development of theoretical framework and its analysis, we neglect this noise. But in the case of simulations the noise effect is considered as well. The oversampled output of the channel is given as

$$\mathbf{x}(j) = \mathbf{H}\mathbf{s}(j) \qquad j = 0,1,.. \tag{3.1.1}$$

where $\mathbf{x}(j)$, $\mathbf{s}(j)$ and $\mathbf{H}$ are defined in (2.2.5), (2.2.6) and (2.2.7), respectively. $\mathbf{s}(j)$ is drawn from 4-QAM coloured constellation with identical distribution. The auto-correlation of the elements of vector $\mathbf{s}(j)$ is given as

$$E[s_i s_{i-k}^*] = \begin{cases} 2 & k = 0 \\ \pm j & k = \pm 1 \\ 0, & \text{otherwise} \end{cases} \tag{3.1.2}$$

Therefore considering the vector $\mathbf{s}(j)$, its correlation matrix with lag 0, $\mathbf{R}_s(0)$ is given as

$$a_{ik} = \begin{cases} 2 & k = i \\ \pm j & k = i \pm 1 \\ 0 & \text{otherwise} \end{cases} \tag{3.1.3}$$

Similarly the elements of $\mathbf{R}_s(1)$ are given as

$$a_{ik} = \begin{cases} -j & k = i \\ 2 & k = i+1 \\ +j & k = i+2 \\ 0 & \text{otherwise.} \end{cases} \tag{3.1.4}$$

### 3.1.1 Necessary and Sufficient Condition for Blind Equalization with Coloured Sources

If there exists an equalization matrix $\mathbf{W}$, then the estimate of the original symbols is given as $\mathbf{y}(j) = \hat{\mathbf{s}}(j)$, where

$$\begin{aligned} \mathbf{y}(j) &= \mathbf{W}\mathbf{x}(j) \\ &= \mathbf{W}\mathbf{H}\mathbf{s}(j) \end{aligned} \tag{3.1.5}$$

$\mathbf{R}_y(n)$ is the autocorrelation matrix of $\mathbf{y}(j)$, which was defined in (2.2.11). Imposing a certain condition on $\mathbf{R}_y(n)$, we can achieve a unique equalization matrix $\mathbf{W}$ by the following theorem.

***Theorem 3.1***:

Let

$$\mathbf{R}_y(n) = \mathbf{R}_s(n) \qquad n = 0,1,2,.. \tag{3.1.6}$$

$\mathbf{H}$ and $\mathbf{s}(j)$ satisfy (3.1.1) and its constraints. Then there exists a unique equalization matrix $\mathbf{W}$, within a phase constant such that,

$$\mathbf{W}\mathbf{H} = \alpha\mathbf{I} \tag{3.1.7}$$

for $n \geq 1$, where $\alpha\alpha^* = 1$ and $\mathbf{I}$ is an identity matrix. For $n = 0$, $\mathbf{W}\mathbf{H} = \alpha\mathbf{I}$ is not a unique solution.

***Proof***:

***Step 1***---We first prove that $\mathbf{WH} = \alpha\mathbf{I}$ is one of the solutions for $n = 0,1,2,...$

$$\mathbf{R}_y(n) = E\left[\mathbf{y}(j)\mathbf{y}^{\mathrm{H}}(j+n)\right] \tag{3.1.8}$$

Using eq. (3.1.5) we get

$$\begin{aligned}\mathbf{R}_y(n) &= E\left[\mathbf{WHs}(j)\mathbf{s}^{\mathrm{H}}(j+n)\mathbf{H}^{\mathrm{H}}\mathbf{W}^{\mathrm{H}}\right] \\ &= \mathbf{WHR}_s(n)(\mathbf{WH})^{\mathrm{H}}\end{aligned} \tag{3.1.9}$$

Since $\mathbf{R}_y(n) = \mathbf{R}_s(n)$, therefore

$$\mathbf{R}_y(n) = \mathbf{WHR}_y(n)(\mathbf{WH})^{\mathrm{H}}$$

$$\mathbf{R}_y(n) = \mathbf{AR}_y(n)\mathbf{A}^{\mathrm{H}} \tag{3.1.10}$$

where

$$\mathbf{A} = \mathbf{WH} \tag{3.1.11}$$

Now

$$\mathbf{A} = \alpha\mathbf{I} \tag{3.1.12}$$

is obviously one of the solutions satisfying eq. (3.1.10) for $n = 0,1,2,...$ with $\alpha\alpha^* = 1$.

***Step2***---Now we prove that $\mathbf{WH} = \alpha\mathbf{I}$ is the only solution in case of $n \geq 1$, to within a phase constant.

Suppose there exists another solution $\mathbf{B} \neq \alpha\mathbf{I}$ that satisfies (3.1.1). Therefore for lag $n = 1$

$$\mathbf{R}_y(1) = \mathbf{BR}_y(1)\mathbf{B}^{\mathrm{H}} \tag{3.1.13}$$

Pre-multiplying (3.1.13) by $\mathbf{B}$ and post-multiplying by $\mathbf{B}^{\mathrm{H}}$

$$\mathbf{B}\mathbf{R}_y(1)\mathbf{B}^{\mathrm{H}} = \mathbf{B}^2\mathbf{R}_y(1)\mathbf{B}^{2\mathrm{H}} \tag{3.1.14}$$

Similarly multiplying $m$ times, we get

$$\mathbf{B}\mathbf{R}_y(1)\mathbf{B}^{\mathrm{H}} = \mathbf{B}^m\mathbf{R}_y(1)\mathbf{B}^{m\mathrm{H}} \qquad m = 2,3,.. \tag{3.1.15}$$

This implies that $\mathbf{B}^n = \mathbf{B}\beta$ is the solution of (3.1.5), where $\beta\beta^* = 1$, i.e. $\beta = e^{j\varphi}$.

Let

$$\beta = \gamma^{-m+1} \tag{3.1.16}$$

where $\gamma$ is a complex constant with unit norm.

Hence

$$\begin{aligned} \gamma &= \beta^{m-1} \\ &= e^{j\varphi(m-1)} \end{aligned} \tag{3.1.17}$$

This implies

$$\mathbf{B}^m = \mathbf{B}\gamma^{-m+1}$$

$$\mathbf{B}^m\gamma^m = \mathbf{B}\gamma \tag{3.1.18}$$

Let

$$\mathbf{B}\gamma = \mathbf{C} \tag{3.1.19}$$

Therefore (3.1.18) becomes

$$\mathbf{C}^m - \mathbf{C} = 0 \qquad \forall \quad m \geq 2$$

which gives

$$\mathbf{C}\left(\mathbf{C}^{m-1} - \mathbf{I}\right) \qquad \forall \quad m \geq 2 \tag{3.1.20}$$

The above equation has two solutions, $\mathbf{C} = \mathbf{0}$ and $\mathbf{C}^{m-1} = \mathbf{I}$, $\forall \quad n \geq 2$

$\mathbf{C} = \mathbf{0}$ is ruled out because it gives $\mathbf{B} = \mathbf{0}$ which cannot satisfy (3.1.10). Therefore

$$\mathbf{C}^{m-1} = \mathbf{I} \qquad \forall \qquad m \geq 2 \tag{3.1.21}$$

For $m = 2$, we get $\mathbf{C} = \mathbf{B}\gamma = \mathbf{I}$ which gives us, using (3.1.17)

$$\begin{aligned} \mathbf{B} &= \gamma^{-1}\mathbf{I} \\ &= e^{j\varphi(1-m)}\mathbf{I} \end{aligned} \tag{3.1.22}$$

that is

$$\mathbf{B} = e^{j\vartheta}\mathbf{I} \tag{3.1.23}$$

which is the same type of solution as given in (3.1.12). Hence $\mathbf{W}$ is unique to within a phase constant.

***Step 3***---Finally we prove that $\mathbf{WH} = \alpha\mathbf{I}$ is not a unique solution in case of $n = 0$.

For lag $n = 0$ we have from (3.1.10)

$$\mathbf{R}_y(0) = \mathbf{A}\mathbf{R}_s(0)\mathbf{A}^{\mathrm{H}} \tag{3.1.24}$$

Since $\mathbf{R}_y(0)$ is Hermitian, it has Cholesky factorization

$$\mathbf{R}_y(0) = \mathbf{R}_y^{1/2}(0)\mathbf{R}_y^{\mathrm{H}/2}(0) = \mathbf{B}\mathbf{B}^{\mathrm{H}} \tag{3.1.25}$$

where

$$\mathbf{B} = \mathbf{R}_y^{1/2}(0) \tag{3.1.26}$$

From (3.1.24) and (3.1.26)

$$\begin{aligned} \mathbf{B}\mathbf{B}^{\mathrm{H}} &= (\mathbf{A}\mathbf{B})(\mathbf{A}\mathbf{B})^{\mathrm{H}} \\ &= \mathbf{C}\mathbf{C}^{\mathrm{H}} \end{aligned} \tag{3.1.27}$$

where

$$\mathbf{C} = \mathbf{A}\mathbf{B} \tag{3.1.28}$$

By the matrix factorization lemma [112]

$$\mathbf{B}\mathbf{B}^{\mathrm{H}} = \mathbf{C}\mathbf{C}^{\mathrm{H}} \tag{3.1.29}$$

if, and only if, there exists a unitary matrix $\boldsymbol{\theta}$ such that $\mathbf{C} = \mathbf{B}\boldsymbol{\theta}$.

Hence

$$\mathbf{AB} = \mathbf{B\theta}$$

This gives

$$\mathbf{A} = \mathbf{B\theta B}^{-1} \tag{3.1.30}$$

which is in general a solution to (3.1.10) for $n = 0$.

In this case $\mathbf{A} = \alpha\mathbf{I}$ is a solution only for a special form of unitary matrix, which is, $\boldsymbol{\theta} = \alpha\mathbf{I}$. In case $\boldsymbol{\theta}$ is kept as a general unitary matrix, then $\mathbf{A} = \mathbf{B\theta B}^{-1}$ is the solution to (3.1.10) for $n = 0$. This may not be equal to $\alpha\mathbf{I}$. ■

## 3.1.2 Neural Network and the new Learning Algorithm

The block diagram of the linear ANN is shown in Fig. 2.5. The received symbols sequence $\{x_j\}$ at the output of the channel is applied to the input layer $L_1$ of the neural network as sequence of vectors $\mathbf{x}(j)$, each having length $md$.

The network layer $L_2$ provides the equalized output $\mathbf{y}(j)$, which is an estimate of the source symbols. In order to establish the learning algorithm for updating the weights of the matrix between layers $L_1$ and $L_2$ of the network, the following cost function, $J(W)$, is to be minimized in the mean square sense.

$$J(W) = \sum_{m=1}^{d}\sum_{n=1}^{d}\left|\mathbf{A}_{nm}\right|^2 \tag{3.1.31}$$

where $\mathbf{A}_{nm}$ is the $(n,m)_{th}$ element of matrix $\mathbf{A}$. It is given as

$$\mathbf{A}_{nm} = \left\{\mathbf{R}_y(1) - \mathbf{R}_s(1)\right\}_{nm} \tag{3.1.32}$$

The cost function is based on the requirement, that the statistics at the destination must approach the statistics at the source, i.e., $\mathbf{R}_y(n) \to \mathbf{R}_s(n)$ for $n = 1, 2, \ldots$ The elements of the equalization matrix are updated according to SGSD algorithm

$$\Delta \mathbf{W} = -\eta \frac{\partial J(W)}{\partial \mathbf{W}} \tag{3.1.33}$$

where

$$\frac{\partial J(W)}{\partial \mathbf{W}_{kl}} = \sum_n \sum_m \left\{ \frac{\partial \mathbf{A}_{nm}}{\partial \mathbf{W}_{kl}} \mathbf{A}^*_{nm} + \mathbf{A}_{nm} \frac{\partial \mathbf{A}^*_{nm}}{\partial \mathbf{W}_{kl}} \right\}$$

where $\mathbf{W}_{kl}$ is the $(k,l)^{th}$ element of matrix $\mathbf{W}$.

$$\frac{\partial J(W)}{\partial \mathbf{W}_{kl}} = 2(\mathbf{A}^{\mathrm{H}})_k \mathbf{W} E\left[ \boldsymbol{x}(j) x_l^*(j+1) \right] + 2\mathbf{A}_k \mathbf{W} E\left[ \boldsymbol{x}(j+1) x_l^*(j) \right] \tag{3.1.34}$$

where $(\mathbf{A}^H)_k$ and $\mathbf{A}_k$ are the $k^{\text{th}}$ rows of matrices $\mathbf{A}^H$ and $\mathbf{A}$ respectively, and $x_l(j)$ is the $l^{\text{th}}$ element of vector $\mathbf{x}(j)$. The above equation can be written in the matrix form as

$$\frac{\partial J(W)}{\partial \mathbf{W}} = 2\mathbf{A}^{\mathrm{H}} \mathbf{W} E\left[ \boldsymbol{x}(j) \boldsymbol{x}^{\mathrm{H}}(j+1) \right] + 2\mathbf{A}\mathbf{W} E\left[ \boldsymbol{x}(j+1) \boldsymbol{x}^{\mathrm{H}}(j) \right] \tag{3.1.35}$$

Using (3.1.32) for matrix $\mathbf{A}$, the updating formula for the whole weight matrix becomes

$$\begin{aligned} \Delta \mathbf{W} &= -\eta \left[ \left\{ \mathbf{R}_y(1) - \mathbf{R}_s(1) \right\}^{\mathrm{H}} \mathbf{W} \mathbf{R}_x(1) + \left\{ \mathbf{R}_y(1) - \mathbf{R}_s(1) \right\} \mathbf{W} \mathbf{R}_x^{\mathrm{H}}(1) \right] \\ &= -\eta \begin{bmatrix} \left\{ E\left[ \mathbf{y}(j) \mathbf{y}^{\mathrm{H}}(j+1) \right] - \mathbf{R}_s(1) \right\}^{\mathrm{H}} \mathbf{W} \mathbf{R}_x(1) \\ + \left\{ E\left[ \mathbf{y}(j) \mathbf{y}^{\mathrm{H}}(j+1) \right] - \mathbf{R}_s(1) \right\} \mathbf{W} \mathbf{R}_x^{\mathrm{H}}(1) \end{bmatrix} \end{aligned} \tag{3.1.36}$$

If we wish to carry out the online learning process i.e. updating of weights for every new vector $\mathbf{x}(j)$ coming in, we remove the expectation operator

$$\left(\Delta \mathbf{W}\right)_{ol} = -\eta \begin{bmatrix} \left\{\mathbf{y}(j)\mathbf{y}^{\mathrm{H}}(j+1) - \mathbf{R}_s(1)\right\}^{\mathrm{H}} \mathbf{W}\mathbf{R}_x(1) \\ + \left\{\mathbf{y}(j)\mathbf{y}^{\mathrm{H}}(j+1) - \mathbf{R}_s(1)\right\} \mathbf{W}\mathbf{R}_x^{\mathrm{H}}(1) \end{bmatrix} \tag{3.1.37}$$

If weights are updated by batch processing, we carry out the expectation operator as follows

$$\begin{aligned} E\left[\mathbf{y}(j)\mathbf{y}^{\mathrm{H}}(j+1)\right] &= E\left[\mathbf{W}\mathbf{x}(j)\mathbf{x}^{\mathrm{H}}(j+1)\mathbf{W}^{\mathrm{H}}\right] \\ &= \mathbf{W}\mathbf{R}_x(1)\mathbf{W}^{\mathrm{H}} \end{aligned} \tag{3.1.38}$$

Thus updating of weights is given by

$$\left(\Delta \mathbf{W}\right)_{bp} = -\eta \begin{bmatrix} \left\{\mathbf{W}\mathbf{R}_x(1)\mathbf{W}^{\mathrm{H}} - \mathbf{R}_s(1)\right\}^{\mathrm{H}} \mathbf{W}\mathbf{R}_x(1) \\ + \left\{\mathbf{W}\mathbf{R}_x(1)\mathbf{W}^{\mathrm{H}} - \mathbf{R}_s(1)\right\} \mathbf{W}\mathbf{R}_x^{\mathrm{H}}(1) \end{bmatrix}$$

## 3.2 Phase Ambiguity in Equalized Symbols for Real Channel

We already know that

$$\begin{aligned} \mathbf{y}(i) &= \mathbf{W}\mathbf{x} \\ &= \alpha \mathbf{s}(i) \end{aligned} \tag{3.2.1}$$

where $\alpha$ is the unknown complex constant with unit norm.

Knowing $\mathbf{W}$ and $\mathbf{x}(i)$, we can calculate $\mathbf{y}'(i)$ defined as follows

$$\mathbf{y}'(i) = \mathbf{W}\mathbf{x}^*(i) \tag{3.2.2}$$

Now using (3.1.1), we can write

$$\mathbf{x}^*(i) = \mathbf{H}^*\mathbf{s}^*(i) \tag{3.2.3}$$

With the assumption of real channel i.e. $\mathbf{H}^* = \mathbf{H}$ the above equation becomes

$$\mathbf{x}^*(i) = \mathbf{H}\mathbf{s}^*(i) \tag{3.2.4}$$

and accordingly

$$\mathbf{y}'(i) = \alpha \mathbf{s}^*(i) \tag{3.2.5}$$

Adding (3.2.1) and (3.2.5) we get

$$\mathbf{y}(i) + \mathbf{y}'(i) = \mathbf{c}\alpha \tag{3.2.6}$$

where $\mathbf{c}$ is a vector whose elements are twice the real part of source symbols vector, $\mathbf{s}(i)$.

Writing the equation for $j^{\text{th}}$ component of this equation we get

$$\alpha \mathbf{c}_j = \left(\mathbf{y}(i) + \mathbf{y}'(i)\right)_j$$

Taking magnitude of both sides and exploiting the fact that $|\alpha| = 1$ and $\mathbf{c}_j$ is real,

$$\mathbf{c}_j = \left|\mathbf{y}(i) + \mathbf{y}'(i)\right|_j$$

thus

$$\alpha = \frac{\left(\mathbf{y}(i) + \mathbf{y}'(i)\right)_j}{\left|\mathbf{y}(i) + \mathbf{y}'(i)\right|_j} \tag{3.2.7}$$

The above equation provides us the unity norm constant $\alpha$, with an ambiguity of $\pi$.

## 3.3 Simulations and Results

In this section simulations are done to demonstrate the validity of the proposed algorithm. The algorithm was tested for two different channels, one is real and the other is complex. The algorithms of Valcarce [88] and Afkhamie [111] were also implemented in the same environment for comparison purpose.

The real channel has been taken from Valcarce [88]. Its impulse response with 2 sub-channels is truncated up to six symbols period i.e. $n = 6$, and has the following coefficients

$$[\boldsymbol{h_0 \ldots h_5}] = \begin{bmatrix} 0.1 & -0.12 & 0.43 & 0.87 & -0.12 & 0.04 \\ 0.15 & 0.45 & -0.76 & 0.21 & -0.15 & 0.11 \end{bmatrix}$$

where $\boldsymbol{h_i}$ is $i^{\text{th}}$ column vector. The observation interval in this case was taken as $5T$, that is, $m = 5$. The complex channel taken from Fang [95] has impulse response truncated up to four symbol periods. It is another 2-ray channel with the observation interval of $3T$, that is, $m = 3$. The coefficients of impulse response for complex channel are given as

$$[\boldsymbol{h_0 \ldots h_3}] = \begin{bmatrix} 0.0554 + 0.0165i & -1.3449 - 0.4523i & 1.0067 + 1.1524i & 0.3476 + 0.3153i \\ -0.8077 - 0.3183i & 0.4307 + 0.2612i & 1.2823 + 1.1456i & -0.3610 - 0.2743i \end{bmatrix}$$

For performance evaluation the Normalized Root Mean Square Inter-Symbol-Interference (NRMSISI) was evaluated as defined in [95], given by the following equation.

$$\text{NRMSISI} = \frac{1}{\|s\|} \sqrt{\frac{1}{N_m} \sum_{i=1}^{N_m} \|s - \alpha_i \hat{s}(i)\|^2} \tag{3.3.1}$$

where $\hat{s}(i)$ are the estimates of input from $i^{\text{th}}$ trial. $N_m$ is the number of Monte Carlo trials and $\alpha$ is given as

$$\alpha_i = \frac{1}{N} \sum_{k=1}^{N} \frac{s_k}{\hat{s}_k(i)} \tag{3.3.2}$$

where $N$ is the number of symbols used.

The algorithm was tested for 100 independent Monte Carlo trials. The results of equalized outputs are given in Fig.3.1 for different values of SNR. The outputs of

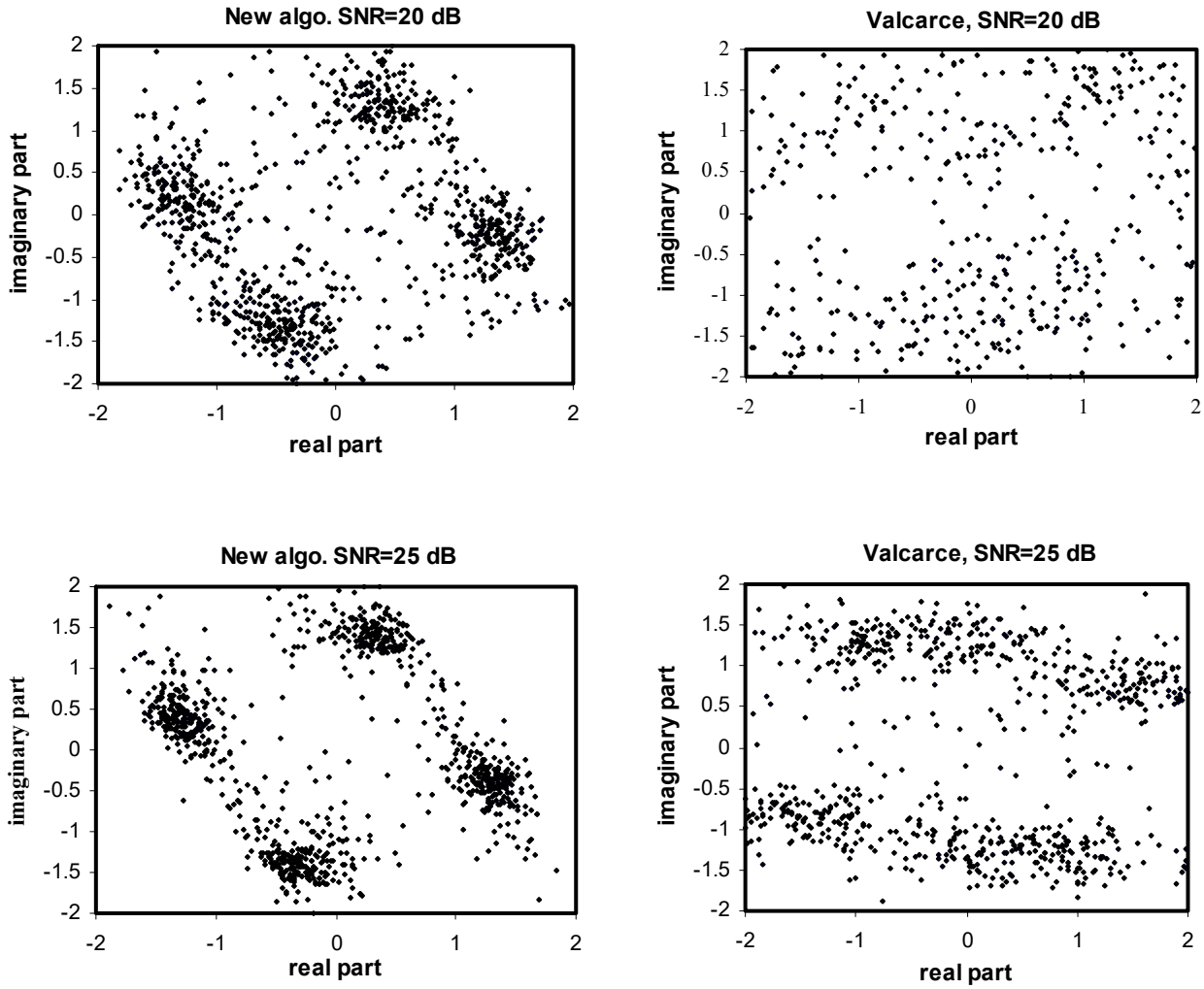



**Fig. 3.1 Symbols after equalization at different SNR.**

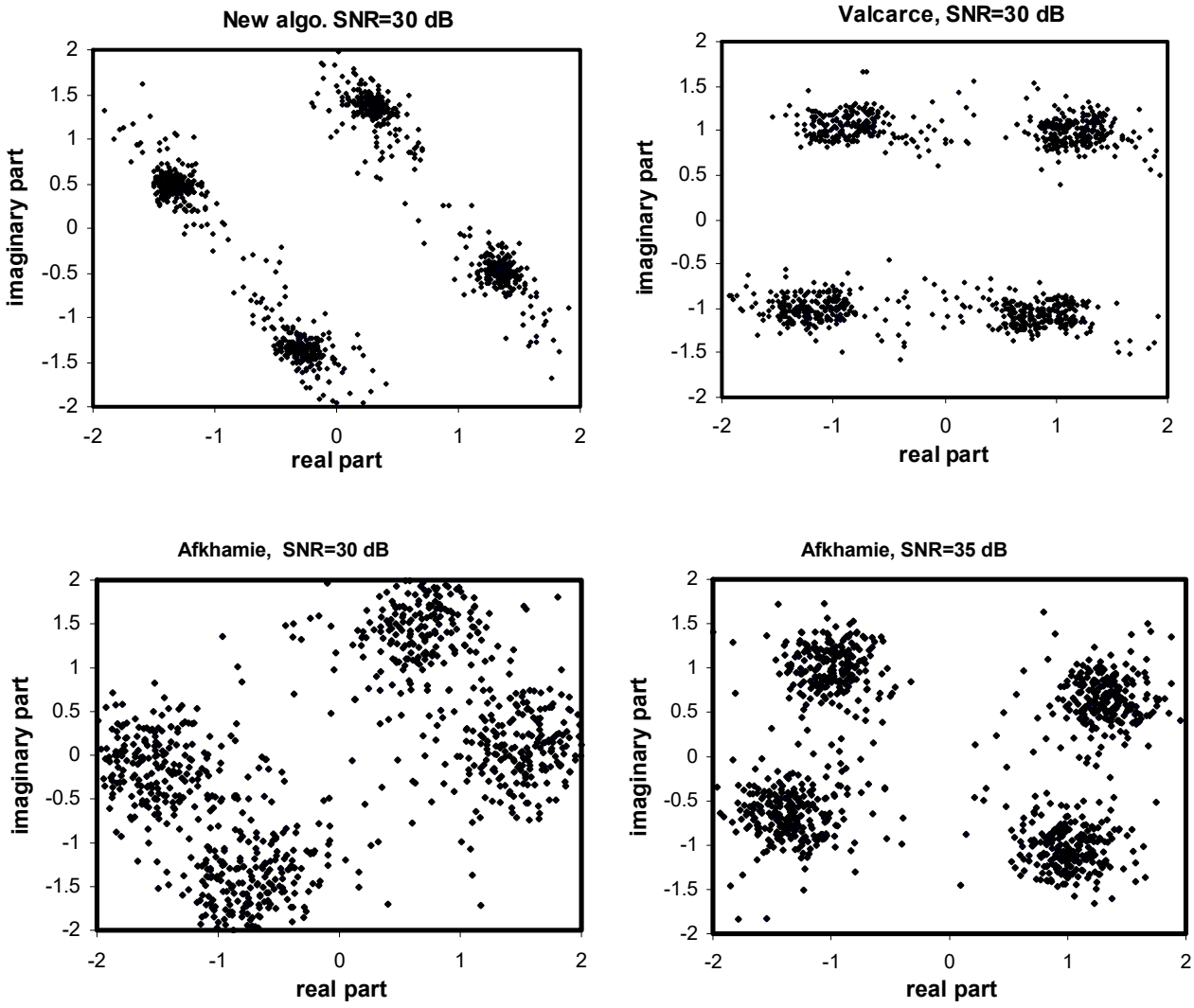



**Fig. 3.1 (Continued)**

Valcarce [88] and Afkhamie [111] is also given for comparison purpose. Number of source symbols were taken as 1000. Our results are better than that of Valcarce [88], especially at low SNR. Fig.3.2 indicates the variations of SER at different SNR. In this case the analysis was taken for real and complex channels using 1000 source symbols. The performance of the new algorithm is better than Valcarce [88] and Afkhamie [111] especially at low SNR and comparable at high SNR.

In Fig. 3.3 the plots of NRMSISI versus SNR are given for real and complex channels for 1000 symbols. For low SNR we see the new algorithm performing much better comparatively. Fig. 3.4 indicates the effect of changing the number of source symbols on SER. The set of observation is taken at SNR=25db. Again the new algorithm is outperforming the other two algorithms, especially for low number of symbols. In Fig. 3.5 the performance in terms of NRMSISI versus number of symbols is given for SNR=25db. The new algorithm again outperforms the other two.

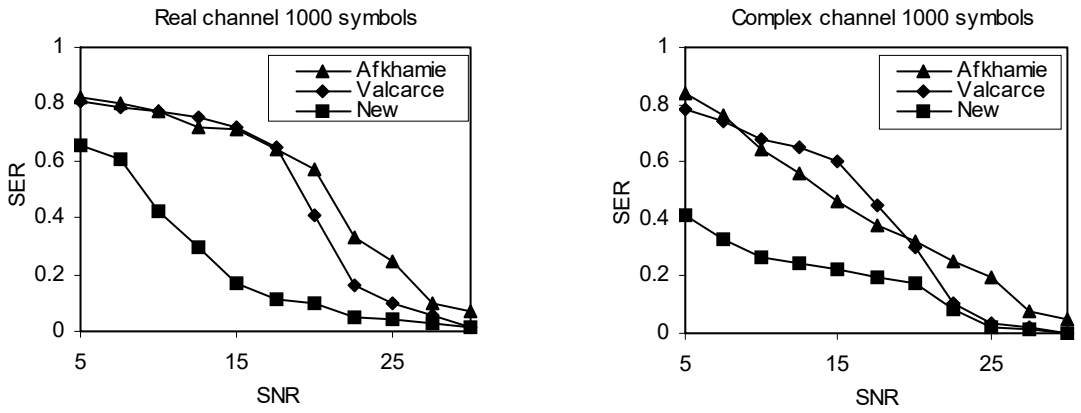



**Fig. 3.2 SER for real and complex channels versus SNR.**

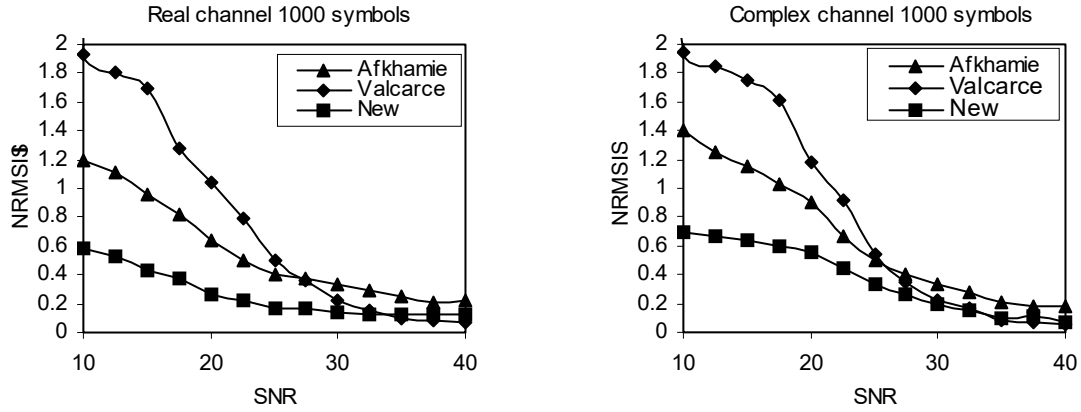



**Fig. 3.3 NRMSISI for real and complex channels versus SNR.**

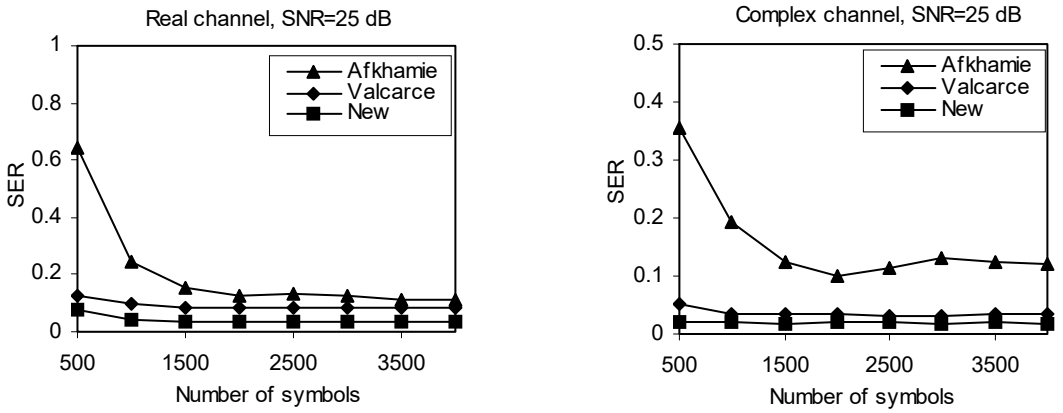



**Fig. 3.4 SER versus number of source symbols for real and complex channels.**

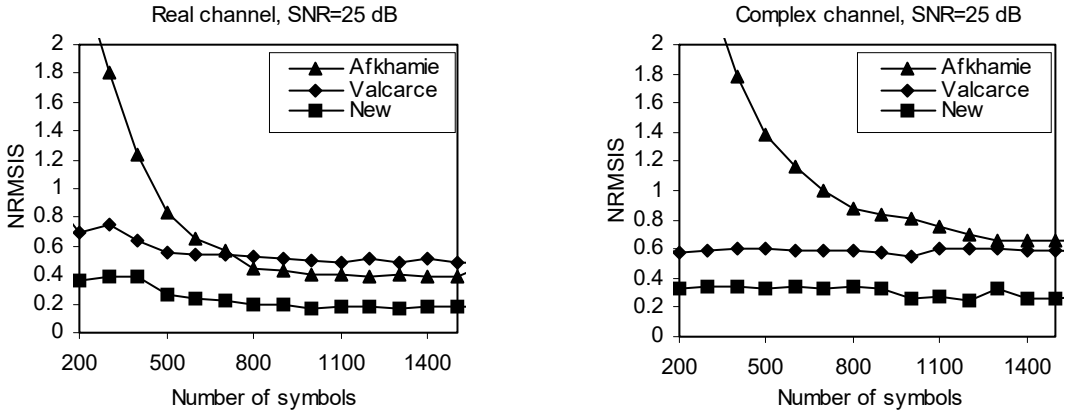



**Fig. 3.5 NRMSISI versus number of source symbols for real and complex channels.**

## 3.4 Energy Matching Approach for Blind Equalization of Channels With Equal Energy Sources

In this section a simple energy–matching approach for blind equalization of possibly nonminimum phase channels is presented. The new approach does exploit the concept of oversampling and hence cyclostationarity, but does not use the matching of statistics of the inputs with the outputs. All the source symbols must have equal energy. Thus an energy constraint is imposed on the outputs. A single FIR transversal filter is used as an equalizer, whose weights are updated by using the energy constraint. This scheme is computationally light and convergence is achieved in one epoch only.

It is well known that if there is only one channel, it is not possible to reconstruct the input sequence by using an FIR equalizer. Tong et al. [45] showed that oversampled output of the channel is cyclostationary in the wide sense, thus containing phase information of the channel. Later on [113] and [104] showed that multiple channel transmission is mathematically equivalent to oversampling, which makes oversampled case an FIR channel bank. Liu et al. [89] has shown that an FIR equalizer bank equalizes the FIR channel bank if, and only if, the composite output of the equalizer bank $y_k$ is temporally uncorrelated. They have also shown that an FIR equalizer bank exists for the zero–forcing condition if, and only if, multichannels do not share a common zero other than $z^{-L}$ for $L \geq 0$ which represents simply the delay. The block diagram of the proposed algorithm is shown in Fig. 3.6.

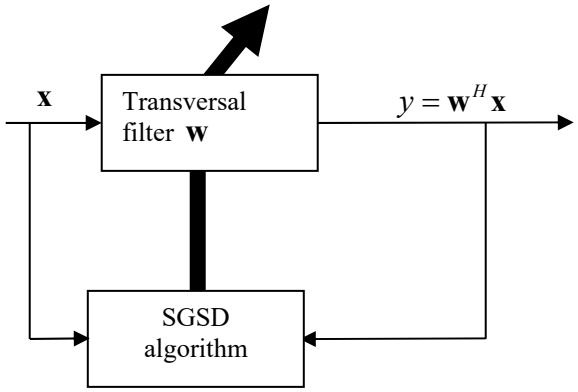



**Fig. 3.6 Block diagram of proposed algorithm.**

Our contention is that for oversampled output of the channel (equivalent to multi–channel case), a single FIR filter can equalize the output of the channel. There is no need of FIR equalizer bank as in [89], or a transformation matrix as an equalizer as in [95]–[96]. Moreover, a simple constraint on the energy of the output symbols of the equalizer is imposed which acts as its cost function for updating the weights of the FIR equalizer. The only drawback in this approach is that, only those constellations, whose symbols have equal energy, can be used. The advantage of this approach is that, the algorithm does not depend upon the knowledge of the statistics of the input symbols. Furthermore, the weights of the FIR equalizer converge in one epoch only. The results are relatively better as compared to some of the previous works in the literature, especially at low SNR.

### 3.4.1 Problem Formulation for Energy Matching Technique

In this case we have considered constellations in which all symbols have equal energy. We have not used statistics matching for input symbols $s_i$ and output symbols $y_i$

from the equalizer. The only constraint we have imposed on the output symbols from the FIR equalizer is the energy constraint. The output vector from the channel is given as

$$\mathbf{x}_j = \mathbf{H}\mathbf{s}_j \tag{3.4.1}$$

where $\mathbf{H}$ is the Sylvester matrix given in (2.2.7), $\mathbf{s}_j$ is the $j^{\text{th}}$ vector of input symbols given in (2.2.6) and $\mathbf{x}_j$ is the $j^{\text{th}}$ vector of channel output given in (2.2.5). The AWGN noise is omitted for theoretical analysis. The following theorem provides an equalization condition based on the matching of the energy of the output symbols with the energy of the input symbols to the channel.

### *Theorem 3.1:*

Suppose $\mathbf{H}$ and $\mathbf{s}_j$ satisfy the linear model (3.4.1) and $\mathbf{x}_j$ is the tap input vector. There exists an FIR equalizer with transversal weights $\mathbf{w}$ such that the output of the transversal filter

$$\begin{aligned} y_j &= \mathbf{w}^H \mathbf{x}_j \\ &= \alpha s_j \end{aligned} \tag{3.4.2}$$

with $\alpha$ a unimodular constant, if the following condition is satisfied:

$$\left|y_j\right|^2 = \left|s_j\right|^2 = E \tag{3.4.3}$$

### *Proof:*

$$\begin{aligned} \left|\mathrm{y_j}\right|^2 &= \mathrm{y_j y_j^H} \\ &= \mathbf{w}^H \mathbf{x}_j \left(\mathbf{w}^H \mathbf{x}_j\right)^H \\ &= \mathbf{w}^H \mathbf{H}\mathbf{s}_j \left(\mathbf{w}^H \mathbf{H}\mathbf{s}_j\right)^H \end{aligned}$$

$$= \mathbf{v}^H \mathbf{s}_j \left( \mathbf{v}^H \mathbf{s}_j \right)^H$$
$$= \left| \mathbf{v}^H \mathbf{s}_j \right|^2$$
$$= \left| s_j \right|^2$$

where energy constraint has been imposed on the last equality. In this case $\mathbf{v}^H = \mathbf{w}^H \mathbf{H}$ and $s_j$ is the leading symbol of the vector $\mathbf{s}_j$. The last equality is only possible if, and only if, $\mathbf{v}^H = [\alpha \ 0 \ 0 \ ..... \ 0]$, where $\alpha\alpha^*=1$. Thus an FIR equalizer does exist with weights $\mathbf{w}$ such that $\mathbf{w}^H \mathbf{H} = \alpha \mathbf{1}$, where $\mathbf{1} = [1 \ 0.... \ 0]$, if energy constraint is satisfied. ■

This condition of energy constraint, used in the theorem, has been utilized to construct the following cost function

$$J(n) = \frac{1}{2}\left( E - \left| y_n \right|^2 \right)^2 \tag{3.4.4}$$

where $y_n = \mathbf{w}^H \mathbf{x}_n$, while $\mathbf{x}_n$ is the tap–input vector at time $n$. The weights of the equalizer are updated by using the SGSD algorithm

$$\begin{aligned} \Delta \mathbf{w}_j &= -\eta \frac{\partial J}{\partial \mathbf{w}_j} \\ &= \eta \left( E - \left| y_n \right|^2 \right) \frac{\partial}{\partial \mathbf{w}_j} \left| y_n \right|^2 \end{aligned} \tag{3.4.5}$$

The final updating formula for $\mathbf{w}$ is given as

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \eta e_n \left( \mathbf{w}^T \mathbf{x}_n^* \right) \mathbf{x}_n \tag{3.4.6}$$

where

$$e_n = E - \left| y_n \right|^2 \tag{3.4.7}$$

### 3.4.2 Simulation and Results for Energy Matching Approach

Simulations are performed to demonstrate the validity and also the superior performance of new algorithm over the previous work. The figure of merit is taken as NRMSISI defined in (3.3.1) and (3.3.2). Symbol error rate (SER) is also considered as another performance evaluation criterion. Simulations are performed for coloured and white source symbols. The coloured sources used in these simulations have the correlation as given by Valcarce et al. [88].

$$E\left[s_i s_k^*\right] = \begin{cases} 2 & k = i \\ \pm j & k = i \pm 1 \\ 0 & \text{otherwise} \end{cases} \tag{3.4.8}$$

The comparison is done with Fang [95] for white case and with Valcarce [88] for coloured case. In both cases the source symbols were taken from 4-QAM constellation in which symbols $s_i \in \left[\pm 1 \pm j\right]$, where $j = \sqrt{-1}$. The energy of each transmitted symbol is $E = \left|s_i\right|^2$. The channel is taken from Tong et al. [45] and given as

$$\mathbf{h} = \begin{bmatrix} -0.02788 & 0.04142 & -0.07025 & 0.3874 & 0.3132 & -0.08374 \\ 0.009773 & -0.01959 & 0.08427 & 0.5167 & 0.01383 & -0.001258 \end{bmatrix}^T$$

The length of transversal equalizer was taken as 14 in our simulations and the learning constant was kept as 0.01.

Fig.3.7 shows the plots of NRMSISI versus SNR, both, for white as well as coloured sources, respectively. The Fang algorithm, [95], for white case was run over 60 epochs, whereas, the new algorithm in both cases was run for single epoch. Source symbols in both cases were kept as 1000. The performance of the new algorithm is clearly better than the performance of the other two algorithms at high, as well as, low SNR.

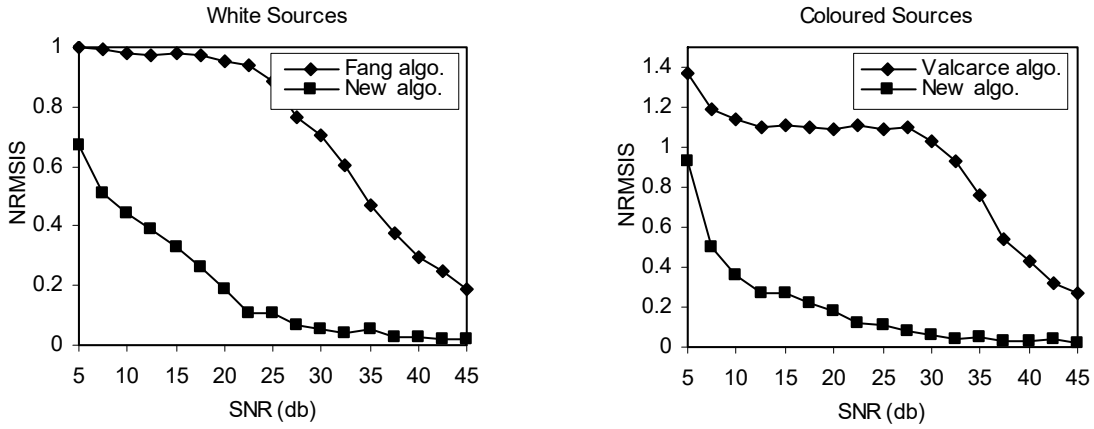



**Fig. 3.7 NRMSISI versus SNR (db), for white and coloured sources.**

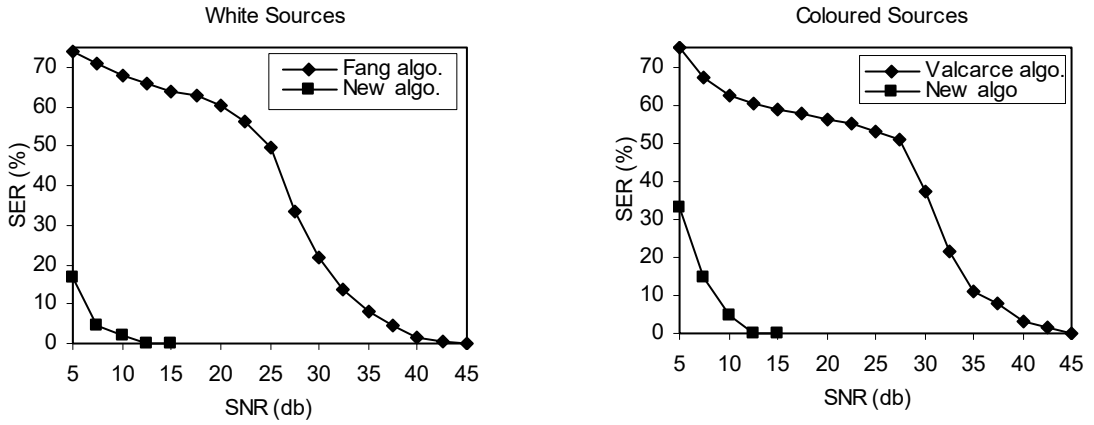



**Fig. 3.8 SER versus SNR (db) for white and coloured sources.**

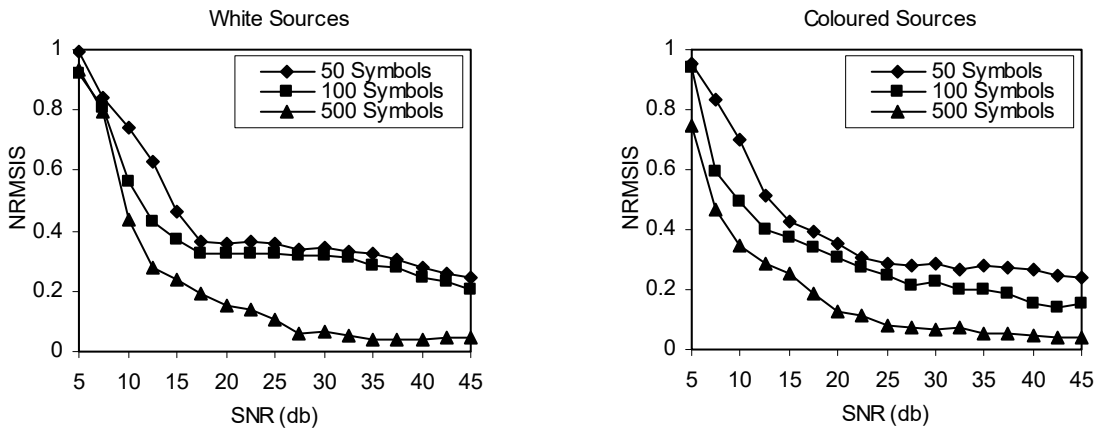



**Fig. 3.9 NRMSISI versus SNR (db) at different number of source symbols.**

Fig.3.8 gives plots for SER versus SNR for white and coloured cases. The new algorithm once again outperforms both the algorithms. SER drops to zero at SNR=15 db for the new algorithm, while for the others it drops to zero for SNR above 40 db. Moreover the new algorithm generates these results in one epoch only.

Fig.3.9. shows the variation of NRMSISI against SNR with parameter as source symbols for the new algorithm. Different number of source symbols are taken in this case. The performance of new algorithm is quite acceptable even at 50 source symbols.

# Chapter 4

# Hybrid HOS-SOS Approach for Blind Equalization of Channels

## 4.1 Introduction

In this chapter a new approach to blind equalization is given which is the combination of approaches belonging to the two different categories. In this case the equalization techniques based on implicit HOS and SOS are combined together. The basic idea is to take advantage of the techniques of both categories by exploiting their positive features.

Generally, techniques based on SOS are supposed to be quick in performance and henceforth, providing faster convergence. This is due to the reason that the data requirement in this case is comparatively less than the case of HOS based algorithms. However, the channel in this case is assumed to exhibit the channel diversity, which is considered to be the backbone of the whole idea. To assure the solvability of the problem the diversified channels or sub-channels must not have a common zero. The mathematical complexity also increases in this case.

The techniques based on HOS implicitly provide Bussgang type algorithms, which are generally capable of providing good end results. These are relatively simple to

implement and there is no requirement of channel diversity. Drawbacks in this case are that these algorithms have slow rate of convergence and are generally limited to binary source symbols. Also the channel used in this case must have unit norm.

By combining the techniques of the above two categories, we are able to achieve a faster convergence without any need of channel diversity. Furthermore, there is no constraint on the norm of the channel in our case. Bussgang algorithm was originally proposed by Bellini [97] and recently modified by Fiori [98] in order to generate better results. Then there is an extra constraint applied on it, which is based on the category of SOS based algorithms. The new algorithm developed as a result was simulated and the results obtained demonstrate its good performance and validity.

The summary of earlier work has already been given in chapter 2. In this chapter the new algorithm is presented. Finally the extension of the new algorithm, modified new algorithm, has also been given and applied to white and coloured sources. Plots and discussion of simulation results are provided with each section.

## 4.2 New Proposal

We suggest two changes in the previous work – one major and the other minor. The proposed major change is the additional term in the cost function based upon the known SOS at the input and the output. We are justified to match the SOS for the input and the output symbols. The additional term in the cost function suggested by us is as follows,

$$J_{add}(n) = \sum_{l=0}^{1} \sum_{nm} \left| \mathbf{D}_{nm} \right|^2 \tag{4.2.1}$$

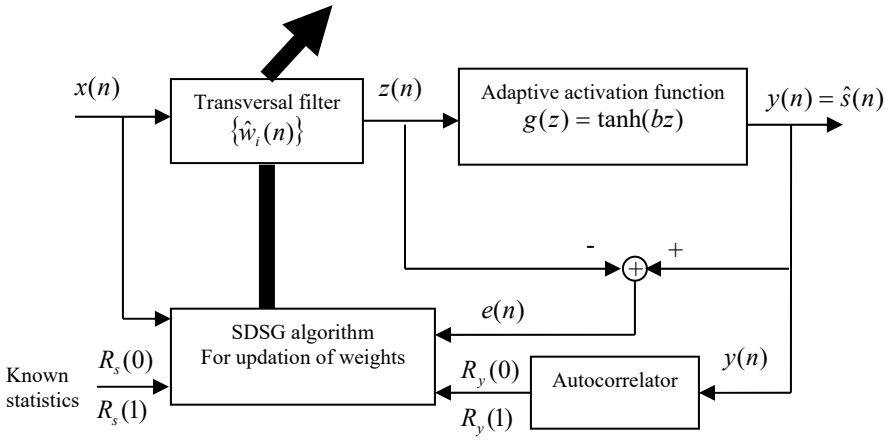



**Fig. 4.1 Schematic presentation of the proposed algorithm.**

where

$$\mathbf{D}_{nm} = \left(\mathbf{R}_y(l) - \mathbf{R}_s(l)\right)_{nm} \tag{4.2.2}$$

where $\mathbf{R}_s(l)$ is given by (2.2.10). $\mathbf{R}_y(l)$ is the correlation matrix of $M \times 1$ output vector symbols $\mathbf{y}(n)$ and is given as

$$\mathbf{R}_y(l) = E\left[\mathbf{y}(n)\mathbf{y}^H(n+l)\right] \quad \text{for} \quad n > 0 \tag{4.2.3}$$

Hence the total cost function becomes

$$J_{new}(n) = J_{old}(n) + J_{add}(n) \tag{4.2.4}$$

The minor change is related to the choice of adaptive activation function (AAF). Although, the same AAF, $g(z) = a\tanh(bz)$, is used, however, we keep $a = 1$ and make only $b$ as adaptive. The reason behind keeping $a$ constant is that, it is simply the amplitude of a sigmoid function and in the case of binary inputs its value plays little role in discriminating between the two inputs. As far as the parameter $b$ is concerned, its

small values will keep $g(z)$ fairly linear for wide range around the origin. As the value of $b$ rises, the sigmoid function tends to become more and more nonlinear and hence, more and more discriminatory between the two binary inputs. As $b \to \infty$, it tends to be a perfect signum function. Thus the updating of $b$ to minimize the cost function, and to play discriminatory role for the binary inputs, is more important and meaningful. Hence our AAF is taken as

$$g(z) = \tanh(bz) \tag{4.2.5}$$

Adaptivity of only one parameter also reduces the computational complexity.

Using SGSD algorithm, the cost function is minimized. This gives us the updating rule for weights of the transversal filter as well as the updating for adaptive coefficients of the activation function. We are considering real binary inputs and real channel.

$$\begin{aligned}(\Delta\mathbf{w})_{new} &= -\eta_w \frac{\partial J_{new}}{\partial \mathbf{w}_k} \\ &= (\Delta\mathbf{w})_{old} + (\Delta\mathbf{w})_{add} \\ &= -\eta_w \left[ \frac{\partial J_{old}(n)}{\partial \mathbf{w}} + \frac{\partial J_{add}(n)}{\partial \mathbf{w}} \right]\end{aligned} \tag{4.2.6}$$

Considering both the terms separately we get

$$\begin{aligned}(\Delta\mathbf{w})_{old} &= -\eta_w \frac{\partial J_{old}}{\partial \mathbf{w}} \\ &= -\eta_w \left(b - bg^2 - 1\right)(g - z)\left(\mathbf{x} - \frac{z}{\kappa^2}\mathbf{w}\right)\end{aligned} \tag{4.2.7}$$

The above result has been derived by Fiori [98]–[99] keeping $a = 1$.

The updating for the $k^{\text{th}}$ element of vector $(\Delta\mathbf{w}_k)_{add}$ is given as

$$
\begin{aligned}
(\Delta \mathbf{w}_k)_{add} &= -\eta_w \frac{\partial J_{add}}{\partial \mathbf{w}_k} \\
&= -\eta_w \sum_{l=0}^{1} \sum_{nm} \mathbf{D}_{nm}(l) \frac{\partial}{\partial \mathbf{w}_k} \mathbf{D}_{nm}(l)
\end{aligned}
\tag{4.2.8}
$$

where factor 2 has been absorbed in $\eta_w$.

$$
\begin{aligned}
\frac{\partial \mathbf{D}_{nm}(l)}{\partial \mathbf{w}_k} &= \frac{\partial}{\partial \mathbf{w}_k} \left\{ E\left[ \mathbf{y}(j) \mathbf{y}^H(j+l) \right] - \mathbf{R}_s(l) \right\}_{nm} \\
&= \frac{\partial}{\partial \mathbf{w}_k} E\left[ \mathbf{y}(j) \mathbf{y}^H(j+l) \right]_{nm} \\
&= \frac{\partial}{\partial \mathbf{w}_k} E\left[ g\left(\mathbf{z}(j)\right) g\left(\mathbf{z}^T(j+l)\right) \right]_{nm}
\end{aligned}
\tag{4.2.9}
$$

where $\mathbf{y}(j) = g(\mathbf{z}(j))$. Since we are considering the real case, the hermitian $H$ is replaced by transpose $T$. We will remove the expectation operator to consider the instantaneous values. The vector $\mathbf{z}(j)$ in the above equation has $M$ elements and is given as

$$
\begin{aligned}
\mathbf{z}(j) &= \left[ \mathbf{w}^T \mathbf{x}(j) \quad \mathbf{w}^T \mathbf{x}(j+1) \quad \cdots \quad \mathbf{w}^T \mathbf{x}(j+M-1) \right]^T \\
&= \left[ \mathbf{w}^T \mathbf{X}(j) \right]^T
\end{aligned}
\tag{4.2.10}
$$

where $\mathbf{x}(j)$ is a vector having length equal to the number of tap weights of the transversal filter, $N$, and the matrix $\mathbf{X}(j)$ is given as

$$
\mathbf{X}(j) = \left[ \mathbf{x}(j) \quad \mathbf{x}(j+1) \quad \cdots \quad \mathbf{x}(j+M-1) \right]_{N \times M}
\tag{4.2.11}
$$

Using the above terms, (4.2.9) can be written as

$$
\begin{aligned}
\frac{\partial \mathbf{D}_{nm}(l)}{\partial \mathbf{w}_k} &= \frac{\partial}{\partial \mathbf{w}_k} \left[ g\left(\mathbf{X}^T(j)\mathbf{w}\right)_n g\left(\mathbf{w}^T \mathbf{X}(j+l)\right)_m \right] \\
&= g\left(\mathbf{X}^T(j)\mathbf{w}\right)_n g'\left(\mathbf{w}^T \mathbf{X}(j+l)\right)_m \mathbf{X}_{km}(j+l) \\
&\quad + g\left(\mathbf{w}^T \mathbf{X}(j+l)\right)_m g'\left(\mathbf{X}^T(j)\mathbf{w}\right)_n \mathbf{X}_{kn}(j)
\end{aligned}
\tag{4.2.12}
$$

The updating for weights due to the additional cost function is given as follows

$$\begin{aligned}\left(\Delta \mathbf{w}_k\right)_{add} &= -\eta_w \sum_{l=0}^{1} \sum_{nm} \mathbf{D}_{nm}(l) \begin{Bmatrix} g\left(\mathbf{X}^T(j)\mathbf{w}\right)_n g'\left(\mathbf{w}^T\mathbf{X}(j+l)\right)_m \mathbf{X}_{km}(j+l) \\ +g\left(\mathbf{w}^T\mathbf{X}(j+l)\right)_m g'\left(\mathbf{X}^T(j)\mathbf{w}\right)_n \mathbf{X}_{kn}(j) \end{Bmatrix} \\ &= -\eta_w \sum_{l=0}^{1} \left\{\mathbf{X}(j+l)\mathbf{e}(l) + \mathbf{X}(j)\mathbf{f}(l)\right\}\end{aligned} \tag{4.2.13}$$

where

$$\mathbf{e}(l) = diag\left(\mathbf{V}_l^T \mathbf{D}(l)\right) \tag{4.2.14}$$

and

$$\mathbf{f}(l) = diag\left(\mathbf{D}(l)\mathbf{U}^T(l)\right) \tag{4.2.15}$$

$\mathbf{V}(l)$ and $\mathbf{U}(l)$ used in above equations are given as

$$\mathbf{V}(l) = g\left(\mathbf{X}^T(j)\mathbf{w}\right) g'\left(\mathbf{w}^T\mathbf{X}(j+l)\right) \tag{4.2.16}$$

$$\mathbf{U}(l) = g\left(\mathbf{w}^T\mathbf{X}(j+l)\right) g'\left(\mathbf{X}^T(j)\mathbf{w}\right) \tag{4.2.17}$$

It is straightforward to prove that $\mathbf{e}(0) = \mathbf{f}(0)$. Thus

$$\left(\Delta \mathbf{w}\right)_{add} = -\eta_w \left\{2\mathbf{X}(j)\mathbf{e}(0) + \mathbf{X}(j)\mathbf{f}(1) + \mathbf{X}(j+1)\mathbf{e}(1)\right\} \tag{4.2.18}$$

The overall updating of weights due to old and additional terms in the cost function as given in (4.2.6) will become

$$(\Delta \mathbf{w})_{new} = -\eta_w \begin{bmatrix} \left\{\left(b - bg^2 - 1\right)\left(g - z\right)\left(\mathbf{x} - \frac{z}{\kappa^2}\mathbf{w}\right)\right\} + \\ \left\{2\mathbf{X}(j)\mathbf{e}(0) + \mathbf{X}(j)\mathbf{f}(1) + \mathbf{X}(j+1)\mathbf{e}(1)\right\} \end{bmatrix} \tag{4.2.19}$$

Similarly the updating of $b$ is achieved as follows

$$\begin{aligned}\left(\Delta b\right)_{new} &= -\eta_b \left[\frac{\partial J_{old}(n)}{\partial b} + \frac{\partial J_{add}(n)}{\partial b}\right] \\ &= \left(\Delta b\right)_{old} + \left(\Delta b\right)_{add}\end{aligned} \tag{4.2.20}$$

Considering the terms separately

$$\begin{aligned}\left(\Delta b\right)_{add} &= -\eta_b \frac{\partial J_{add}}{\partial b} \\ &= -\eta_b \sum_{l=0}^{1} \sum_{nm} \mathbf{D}_{nm}(l) \frac{\partial}{\partial b} \mathbf{D}_{nm}(l)\end{aligned} \tag{4.2.21}$$

where

$$\begin{aligned}\frac{\partial \mathbf{D}_{nm}(l)}{\partial b} &= \left(1-g^2{}_0\right)\left(\mathbf{X}^T(j)\mathbf{w}\right)_n g\left(\mathbf{w}^T\mathbf{X}(j+l)\right)_m \\ &\quad +\left(1-g^2{}_l\right) g\left(\mathbf{w}^T\mathbf{X}(j)\right)_n \left(\mathbf{w}^T\mathbf{X}(j+l)\right)_m\end{aligned} \tag{4.2.22}$$

where

$$g_l^2 = g\left(\mathbf{w}^T\mathbf{X}(j+l)\right) g\left(\mathbf{X}^T(j+l)\mathbf{w}\right) \quad \text{for} \quad l = 0,1,.. \tag{4.2.23}$$

Finally

$$\left(\Delta b\right)_{add} = -\eta_b \left\{c_0 Tr\left[\mathbf{D}^T(0)\mathbf{Q}\right] + c_1 Tr\left[\mathbf{D}^T(1)\mathbf{R}(1)\right]\right\} \tag{4.2.24}$$

where

$$\mathbf{Q} = \mathbf{P}(0) + \mathbf{R}(0) + \mathbf{P}(1) \tag{4.2.25}$$

where

$$\mathbf{P}_{nm}(l) = \left(\mathbf{X}^T(j+l)\mathbf{w}\right)_n g\left(\mathbf{w}^T\mathbf{X}(j+l)\right)_m \tag{4.2.26}$$

and

$$\mathbf{R}_{nm}(l) = g\left(\mathbf{X}^T(j)\mathbf{w}\right)_n \left(\mathbf{w}^T\mathbf{X}(j+l)\right)_m \tag{4.2.27}$$

and $c_0 = 1-g_0^2$ , $c_l = 1-g_l^2$ .

The overall updating formula for $\left(\Delta b\right)_{new}$ is given as

$$\left(\Delta b\right)_{new} = -\eta_b \left\{\begin{aligned}&\left[(g-z)(1-g^2)z\right] \\ &+\left\{c_0 Tr\left[\mathbf{D}^T(0)\mathbf{Q}\right] + c_1 Tr\left[\mathbf{D}^T(1)\mathbf{R}(1)\right]\right\}\end{aligned}\right\} \tag{4.2.28}$$

The expectation operators have been removed for updating purpose from (4.2.19) and (4.2.28), which formulate our new algorithm.

## 4.3 Simulation Results for Binary Case

Simulations are performed on two different channels, Channel-1 and Channel-2. Channel-1 is taken from Tong et al. [45] and Channel-2 is taken from Valcarce et al. [88]. Both the channels have integer time delays, which are multiples of the symbol interval T. The channel parameter vectors for both are given below

Channel-1

$$\mathbf{h}_1 = \begin{bmatrix} -0.02788 & 0.04142 & -0.07025 & 0.3874 & 0.3132 & -0.08374 \end{bmatrix}^T \qquad (4.3.1)$$

Channel-2

$$\mathbf{h}_2 = \begin{bmatrix} 0.1 & -0.12 & 0.43 & 0.87 & -0.12 & 0.04 \end{bmatrix}^T \qquad (4.3.2)$$

Received symbols were obtained by convolving the binary source symbols with the channel parameters and adding noise subsequently. Equalization was then carried out using the proposed algorithm. The Fiori [98] algorithm was also implemented in the same environment for comparison purposes. Both the algorithms were tested with 100 independent Monte Carlo trials.

The algorithm is initialized with all weights set as zeros except the middle term, which is kept equal to 1, $b(0) = 1$, $\eta_w = 0.001$, $\eta_b = 0.002$ and $\kappa = 0.5$. The output symbols were translated in both cases to the mean level of +1 and -1 in accordance with the input statistics and MSE was calculated subsequently.

Fig.4.2 illustrates the equalized output symbols. In this case 2000 output symbols from channel with SNR 20db were fed into the equalizer and plotted after equalization. Their values after equalization are indicated along vertical axis. The outputs of the new algorithm were taken after 5 epochs. The results of Fiori's algorithm [98] are also given which were taken after 100 epochs in this case.

The convergence of proposed algorithm for different number of epochs with different SNR is plotted in Fig. 4.3. The proposed algorithm definitely gives faster convergence as compared to the other work. In this case the plots for Fiori's algorithm [98] are given separately since the new algorithm converges from 2 to 5 whereas Fiori's algorithm [98] takes 13 epochs for channel–2. However in case of channel–1 the Fiori's algorithm [98] suffers a lot and fails to reach its optimum values even after 150 epochs.

The variation of MSE against SNR is given in Fig. 4.4, the plots are given for both channels. The results of the new algorithm are taken after 5 epochs whereas, the results of Fiori [98] are given after 150 epochs for channel–1 and 30 epochs for channel–2 for comparison purpose. Again the performance of the new algorithm is far better for channel–1 and comparable with Fiori [98] for channel–2. The reason is that channel-2 is normalized to one ($\mathbf{h}^T\mathbf{h} = 1$), an assumption that is made for the conventional Bussgang algorithm [97], whereas channel-1 is not. Contrary to this, the performance of the new algorithm is not affected by any such assumption and hence, is more robust to the norm of the channel.

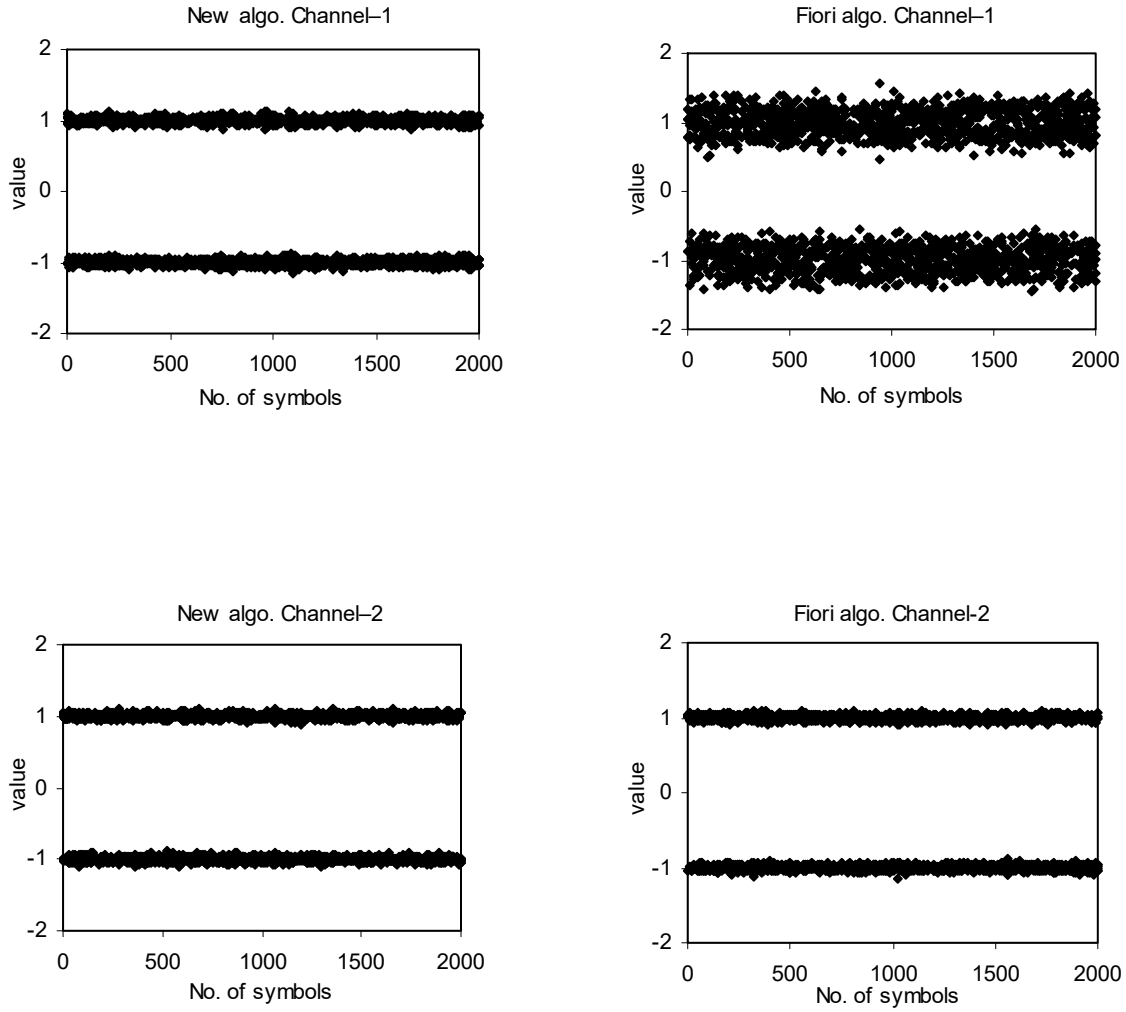



**Fig. 4.2 Values of output symbols vs. number of symbols for both the channels and both the algorithms.**

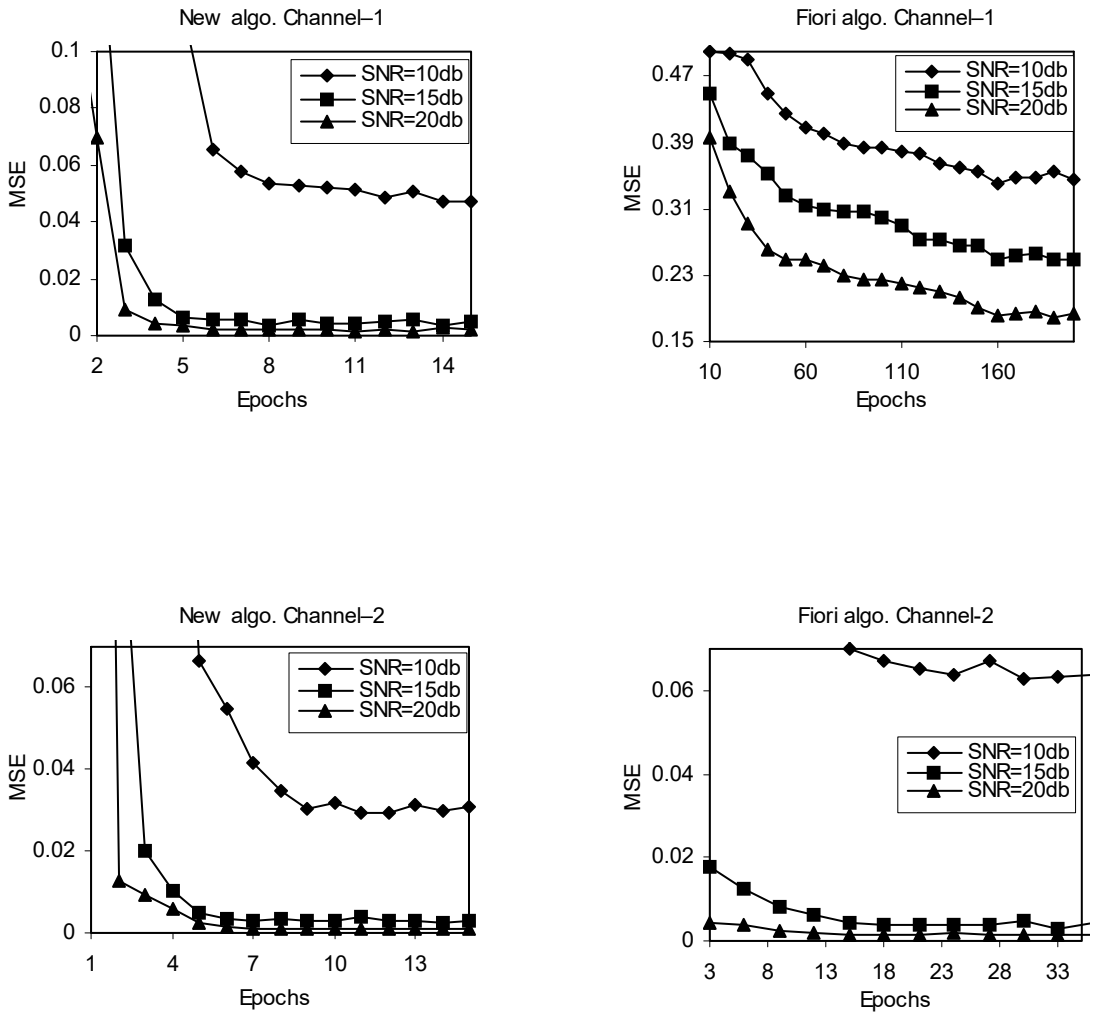



**Fig. 4.3 MSE vs. No. of epochs for both the channels and both the algorithms keeping SNR(db) as parameter.**

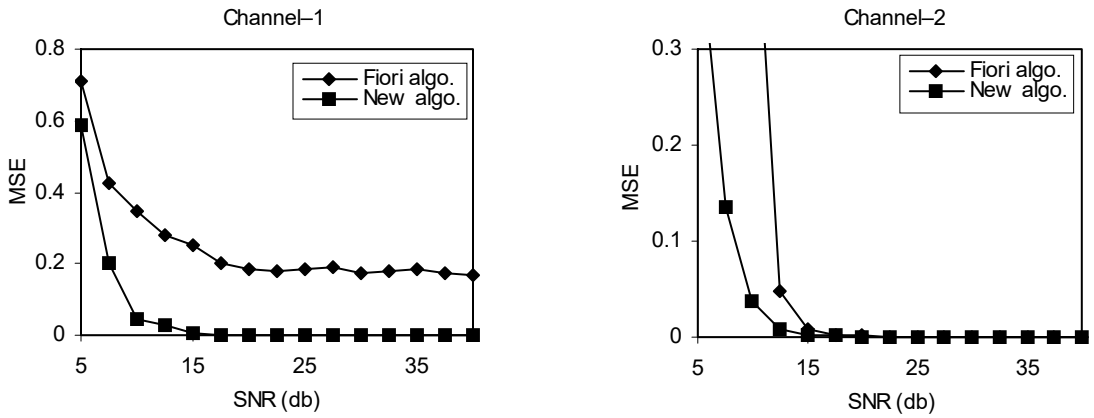



**Fig. 4.4 MSE vs. SNR for both the channels and both the algorithms.**

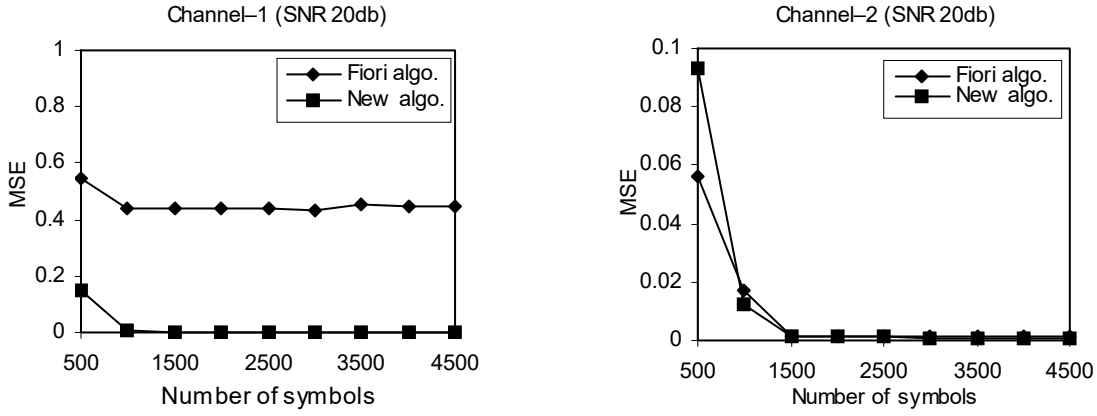



**Fig. 4.5 MSE vs. No. of symbols for both the channels and both the algorithms.**

The effect of SNR on BER is plotted in Fig. 4.5. The results for new algorithm are taken after 7 epochs, whereas those of Fiori [98] are taken after 150 and 30 epochs for channel–1 and channel–2 respectively. In case of Channel-2 both the algorithms perform well. However, in case of Channel-1 the performance of Fiori's algorithm [98] is suffering because of the same reason as mentioned above. The new algorithm is thus definitely more robust to different channels.

## 4.4 Extension to Complex Coloured 4-QAM and the use of Energy Matching Term

In this section an additional term in the cost function is suggested for the constellations having equal energy symbols. Moreover, only one correlation matrix out of the two, as required previously in [114], is now shown to be sufficient for SOS matching at both ends and achieving convergence in one epoch only. The constraint on tap-weights is also removed and the validity of the algorithm is tested using both white and coloured sources and results are compared with other algorithms of same domain.

### 4.4.1 Proposed Modification

In this case a new term in the previous cost function is introduced, which is effective for constellations having equal energy symbols. As test cases the sources used are binary, where $s_i \in \{1,-1\}$, and 4-QAM, where $s_i \in \{\pm 1 \pm j\}$, where $j = \sqrt{-1}$. Energy of every symbol in 4-QAM is $E_s = 2$. In Fig. 4.6 the output of the transversal filter $z$ is given as

$$z = z_R + jz_I \tag{4.4.1}$$

where the subscripts $R$ and $I$ stand for real and imaginary parts, respectively, throughout. The AAF is now complex for the QAM case, such that the output $y$ is given as

$$\begin{aligned} y &= \tanh\left(bz_R\right) + j\tanh\left(bz_I\right) \\ &= g\left(z_R\right) + jg\left(z_I\right) \\ &\square\ g_R + jg_I \end{aligned} \tag{4.4.2}$$

Our new term for the cost function, which is a restriction on the energy of the output symbols, is given as

$$J_{new} = \frac{1}{2}\left(E_s - \left|y\right|^2\right)^2 \tag{4.4.3}$$

where $y$ is the output symbol and $E_s$ is the energy of the symbol.

Apart from the above modification, two more changes are suggested to the previous work. Firstly, we use only one correlation matrix $\mathbf{R}_s(0)$ at the input, which is matched to $\mathbf{R}_y(0)$ at the output. This reduces computational complexity without compromising on the convergence and accuracy of the algorithm. Secondly, the restriction on the norm of the weight vector $\mathbf{w}^T\mathbf{w} = \kappa^2$ is removed to further cut down the computational load. In fact, $J_{new}$ is now a strong indirect constraint, both on the weights and the convergence. The total cost function becomes

$$\begin{aligned} J &= \frac{1}{2}\left|g\left(z\right) - z\right|^2 + \frac{1}{2}\sum_{nm}\left|\mathbf{D}(0)_{nm}\right|^2 + \frac{1}{2}\left(E_s - \left|y\right|^2\right)^2 \\ &= J_{old} + J_{add} + J_{new} \end{aligned} \tag{4.4.4}$$

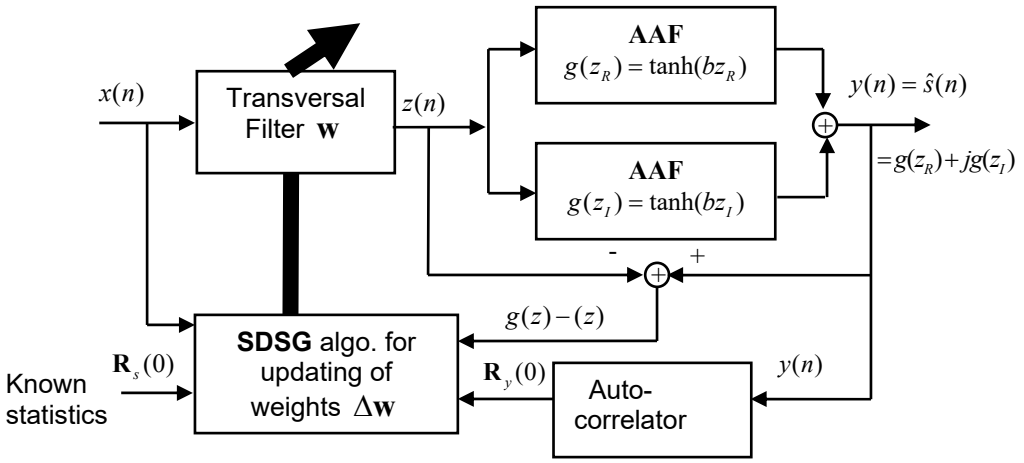



**Fig. 4.6 Modified equalization algorithm.**

where $\mathbf{D}(0) = \mathbf{R}_y(0) - \mathbf{R}_s(0)$. Using the Stochastic Gradient Steepest Descent (SGSD) algorithm, the updating of the weights is given as

$$\begin{aligned}\Delta\mathbf{w} &= -\eta_1\left(\frac{\partial J_{old}}{\partial \mathbf{w}} + \frac{\partial J_{new}}{\partial \mathbf{w}}\right) - \eta_2 \frac{\partial J_{add}}{\partial \mathbf{w}} \\ &= -\eta_1\left[\left(b - bg_R^2 - 1\right)\left(g_R - z_R\right) - j\left(b - bg_I^2\right)\left(g_I - z_I\right)\right]\mathbf{x} \\ &\quad -\eta_1\left[b\left(E_s - |y|^2\right)\left(g_R\left(1 - g_R^2\right) - j \;\; g_I\left(1 - g_I^2\right)\right)\right]\mathbf{x} \\ &\quad -\eta_2\left[\mathbf{X}(j)(\mathbf{p} + \mathbf{q})\right]\end{aligned} \tag{4.4.5}$$

where $\mathbf{X}(j)$ is defined by (4.2.11).

Since the updating of weights due to $J_{old}$ and $J_{new}$ is done on a symbol to symbol basis, therefore, their step size is kept the same, i.e. $\eta_1$. In case of updating due to $J_{add}$, first $M$ output symbols are required to give the instantaneous value of the output correlation matrix. Then the updating due to $J_{add}$ is carried out every $M$ symbols. For this reason its step size is kept as different, i.e., $\eta_2$. This allows more freedom for improving the convergence of the algorithm.

The updating for $b$ is given as follows

$$\begin{aligned}\Delta b &= -\eta_b \frac{\partial}{\partial b}\left(J_{old} + J_{new}\right) \\ &= \eta_b \left[\left(1-g_R^2\right) z_R \left(g_R - z_R - 2g_R e\right) + \left(1-g_I^2\right) z_I \left(g_I - z_I - 2g_I e\right)\right]\end{aligned} \tag{4.4.6}$$

where $e = E_s - \left|y\right|^2$ . As can be seen from (4.4.6), the updating of $b$ uses only $J_{old}$ and $J_{new}$, and involves only scalar multiplications and additions. The updating due to $J_{add}$, as used in our previous section, involved matrices which were computationally expensive and time consuming, with little contribution to $\Delta b$. Therefore, this has been neglected. Equations (4.4.5) and (4.4.6) represent our new modified algorithm.

## 4.4.2 Simulation and Results for Modified Algorithm

The source symbols for the white case are taken from binary and 4-QAM constellations. The correlation matrix for the binary case is $\mathbf{R}_s(0) = \mathbf{I}_{M\times M}$ while for 4-QAM it is $2\mathbf{I}_{M\times M}$, where $\mathbf{I}$ is the identity matrix. In case of binary source symbols, the imaginary term in eq. (4.4.5) and terms with subscript $I$ in eq. (4.4.6) are dropped. For the coloured case, only 4-QAM is used but elements of $\mathbf{R}_s(0)$, as used by Valcarce [88], are given as in (3.4.8). The simulations were performed on the same two channels as given by (4.3.1) and (4.3.2). The equalizer was implemented by choosing the values of constants as $\eta_1 = \eta_2 = \eta_b = 0.005$ initially. They are finally tuned independently for better and better results. All the weights were taken to be zero initially, except for the middle weight, which was kept as 1. Algorithms of Fang et al. [95] and Valcarce et al. [88]

were also implemented in the same environment for comparison purposes. All the algorithms were tested for 100 independent Monte-Carlo trials.

Fig. 4.7 shows the variation of SER and NRMSISI against SNR for both the channels with white and coloured 4-QAM sources. The number of source symbols in this case was set to 1000 for each algorithm. Fang's algorithm [95], for white sources, was run over 60 epochs while the new algorithm was run for a single epoch. As can be seen from Fig.4.7, the performance of the new algorithm is superior to Fang [95] for the case of white QAM sources, and to Valcarce [88] for the case of coloured QAM sources. The new algorithm converges at about SNR=15 db for both white and coloured sources, while the other algorithms converge after about SNR=35 db.

In case of Fig. 4.8 the same simulations are repeated but for channel–2. New algorithm once again outperforms the two competitors for both white and coloured cases. Fig. 4.9 shows the results for binary source symbols. Again we have utilized both the channels. As can be seen from Fig. 4.9, Fiori's algorithm [95] exhibits satisfactory performance for the case of channel with norm one (Fig. 4.9(b)), and relatively poor performance for the case of the other channel (Fig.4.9(a)) whose norm is not equal to one. On the other hand, the new modified algorithm is seen to perform consistently well for both the channels demonstrating comparatively enhanced robustness to the value of the channel norm. The reason for the poor performance of Fiori's algorithm [98] on channel without norm one (Fig.4.9 (a)) arises from the fact that being an HOS-only based algorithm, its performance is limited to the channels with norm one [101].

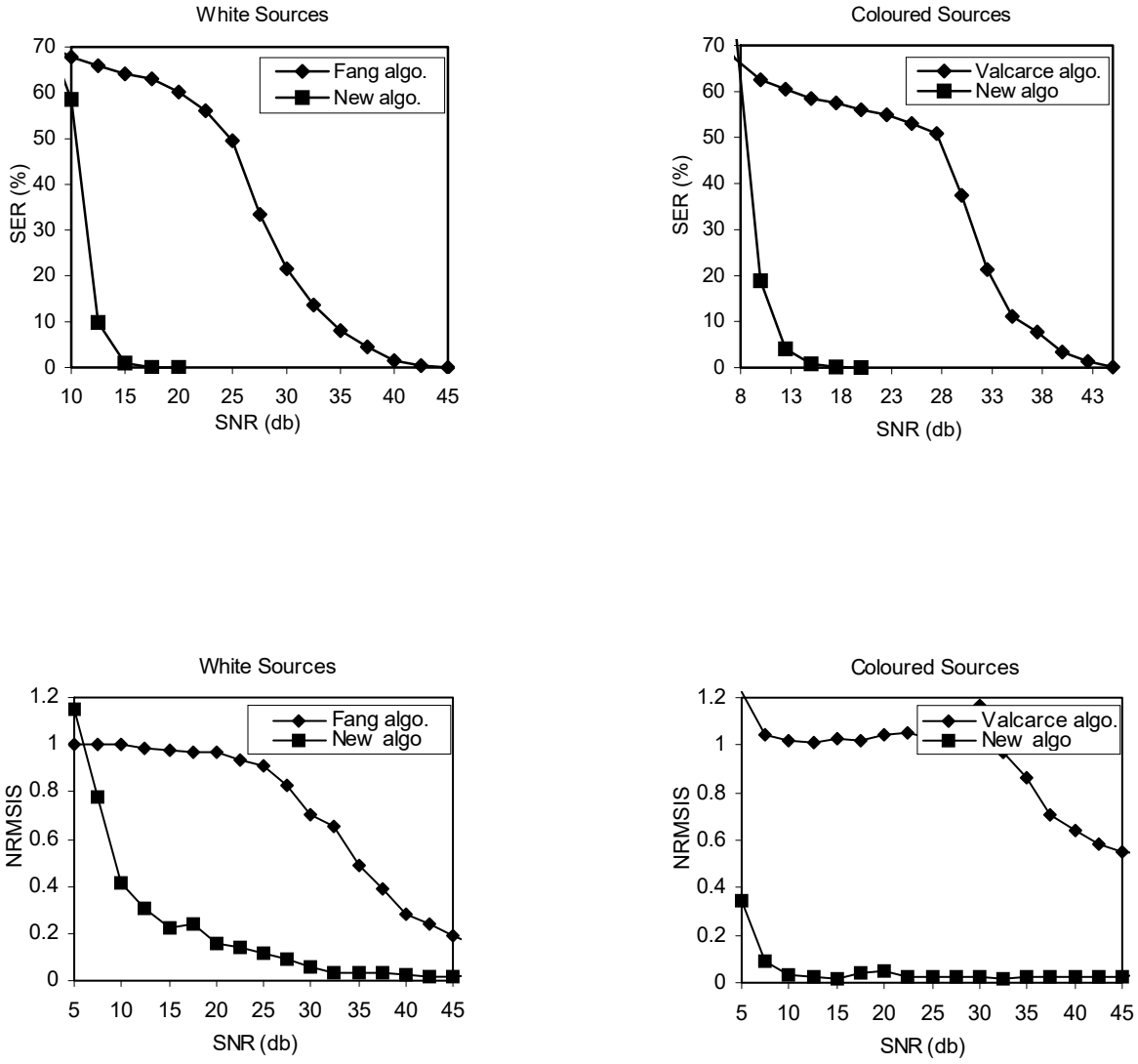



**Fig. 4.7 SER and NRMSISI vs SNR for channel–1 with white and coloured sources from 4-QAM.**

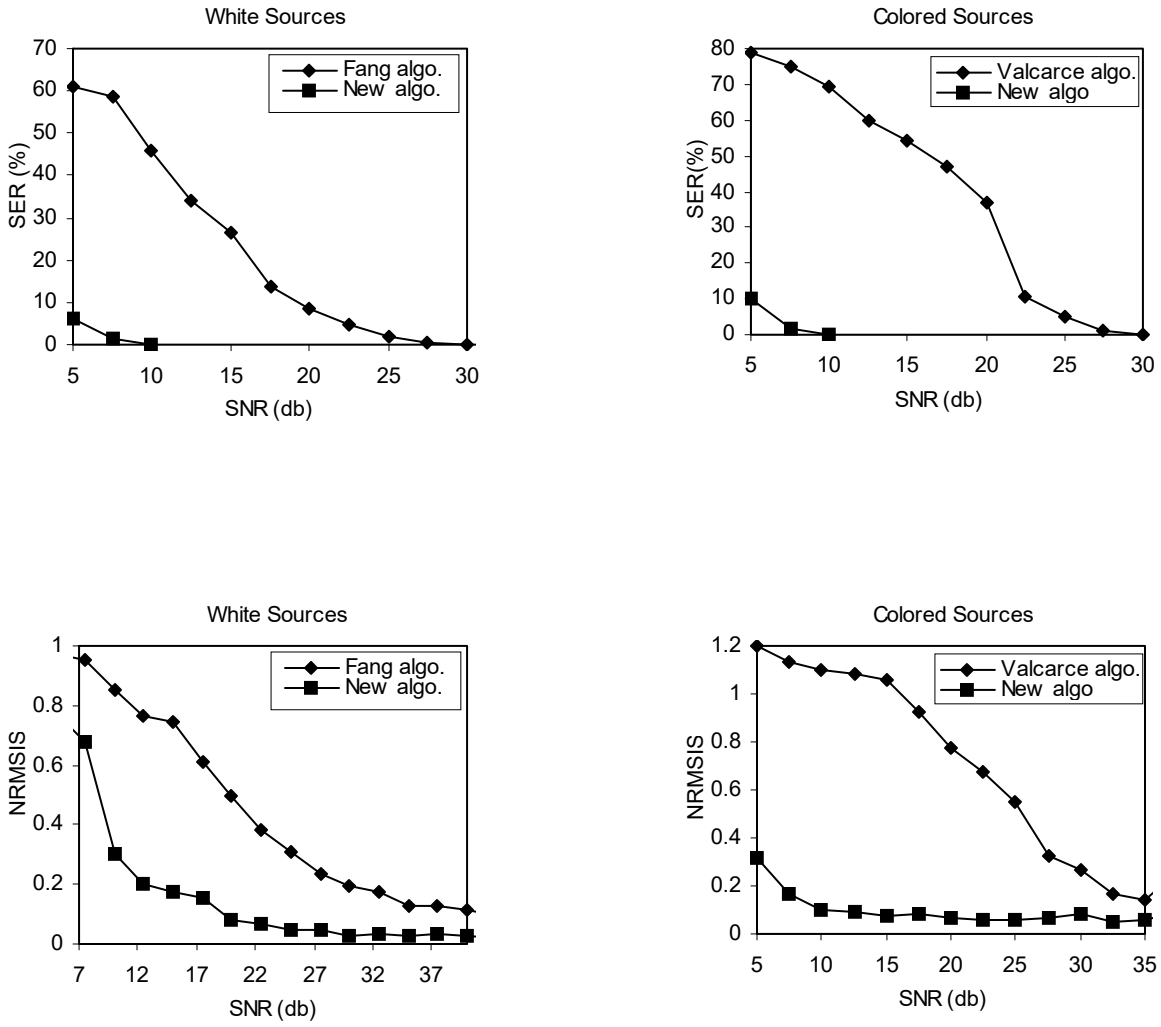
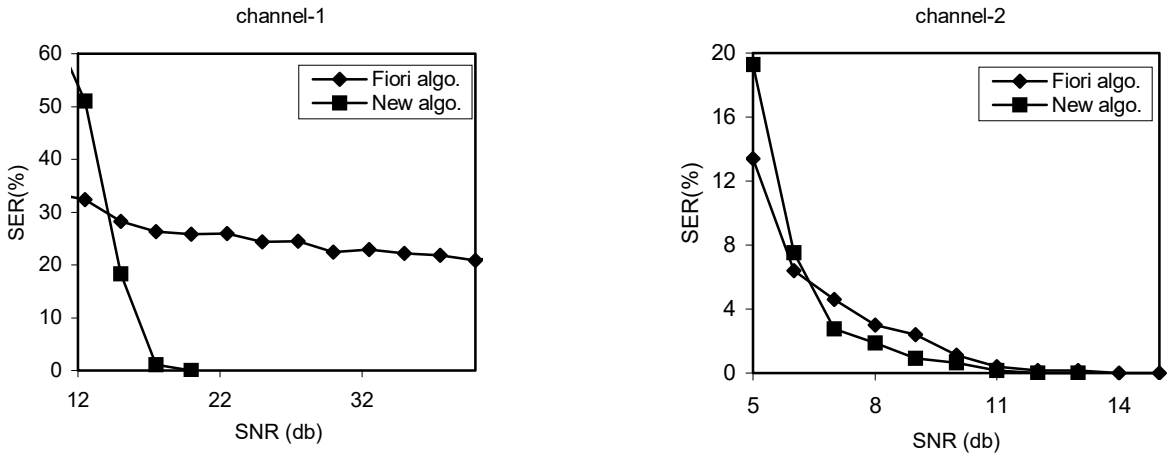



**Fig. 4.8 SER and NRMSISI vs SNR for channel–2 with white and coloured sources from 4-QAM.**




(a) (b)

**Fig. 4.9 SER vs SNR for binary source symbols for both the channels.**

# Chapter 5

# Conclusions

## 5.1 Conclusions

The problem of blind and semiblind equalization is an on going problem as the speed, accuracy and volume of the communication is increasing. Depending on the type of application, the blind equalization utilizes the structural information of the channel and some properties of the input to the channel. Due to the diversity in different applications there are new approaches cropping up. This dissertation deals with the problem of blind equalization of channels, which are SISO as well as SIMO. In the literature, the algorithms for blind equalization fall in HOS domain or in SOS domain. We have dealt with algorithms of both domains and, in fact, also introduced a hybrid approach with much faster convergence. .

In case of SOS domain we have tried to decrease the computational complexity and at the same time tried to achieve faster rate of convergence as compared to some of the competitors in the literature. To cut down the computational complexity, we have used the matching of single correlation matrix with lag one at the source and the destination while using the coloured source symbols as well. Its necessity and sufficiency was proved as a theorem. Based on this result we used a cost function to develop a learning algorithm for the weights of the equalizer An ANN was then implemented,

whose weights are updated using the proposed algorithm. Our results have shown robustness and convergence even at low SNR compared to the other work in the relevant area.

Our next contribution in SOS domain was the algorithm based on the use of the energy matching term. This is limited to those constellations, which have equal energy source symbols. We constrained the symbols at the output of the equalizer to have same energy as that of the input symbols. This resulted in convergence in one epoch only. The sources for this algorithm were taken from 4–QAM constellation which are white as well as coloured. These results excelled by far the results given by Valcarce [88] and Fang [95].

The second part of our work was more novel and more contributing. In this case we used a hybrid or combined HOS-SOS approach. Keeping intact the major formulation of implicit HOS or Bussgang algorithm, we carried out an additional simple SOS matching of the correlation matrices of lag zero and lag one of the source symbols with the equalized symbols. Note that no oversampling was done in this case, because it is fundamentally an HOS technique. With this additional constraint of statistics matching, the convergence took place in three to five epochs as compared to the conventional Bussgang algorithm by Bellini [97], which takes more than two hundred epochs to converge and Fiori [101] which takes about thirty epochs.

The combined HOS-SOS was further improved by using only one correlation matrix with lag zero to be matched at the source and the destination. At the same time, we added another constraint on output symbols, which is that, the energy of the output symbols must be equal to the energy of the input symbols. Although this modification

gave the same results as before, but the convergence took place in one epoch only. The limitation in this case is that we have to use constellations in which all symbols have equal energy.

With these modifications the new algorithms were equally valid for the channels having norm other than one, which was assumed to be the fundamental constraint on the channels used by the Bussgang type algorithms. All this effort was carried out to achieve faster convergence of our results (less number of epochs) with least possible SER and at the same time to enhancing the robustness of the proposed algorithm, especially at low SNR.

## 5.2 Future works

The work done in SOS and HOS domain for SISO and SIMO channels may also be looked into for multiple input and multiple output (MIMO) channels. Obviously apart from ISI, the multiple access interference (MAI) has also to be dealt with.

The energy matching technique in SOS domain may also be looked into as a strong candidate for radio channels because of its faster convergence.

For fast fading techniques one may use space-time codes along with these algorithms to face the fast fading channels.

Currently the energy matching technique is applicable to the constellations with equal energy symbols. In future work one could look into using the same technique for constellations with symbols having different energies.

The same algorithms could also be used for channels having coloured noise instead of white noise.

The algorithms developed above can be extended to nonlinear as well as bilinear channels.

Another important direction could be the hybrid computing, e.g. ANN and Genetic Algorithms, for fast convergence of the algorithm.

# References:


[1] R. Pan and C.L. Nikias, "The complex cepstrum of higher order cumulants and nonminimum phase identification," *IEEE Trans. Acoust. Speech Signal Process.,* vol. 36, pp.186–205, 1988.

[2] D. Hatzinakos and C.L. Nikias, "Estimation of multipath channel response in frequency selective channels," *IEEE J. Select. Areas Commun.,* vol. 7, pp. 12–19, Jan. 1989.

[3] D. Hatzinakos, "Blind equalization based on polyspectra," *Ph.D. dissertation,* Northeastern Univ., Boston, MA, 1990.

[4] D. Hatzinakos and C.L. Nikias, "Blind equalization using a tricepstrum-based algorithm," *IEEE Trans. Commun*., vol. 39, pp. 669-682, May 1991.

[5] D. Hatzinakos and C.L. Nikias, "Blind equalization based on higher-order statistics (H.O.S.)," in *Blind Deconvolution*, S. Haykin, Ed. Englewood Cliffs, NJ: Prentice-Hall, 1994.

[6] C.L. Nikias and A.P. Petropulu, "Higher-Order Spectra Analysis: A Nonlinear Signal Processing Framework," *Englewood Cliffs,* NJ: Prentice-Hall, 1993.

[7] J.A. Cadzow, "Blind deconvolution via cumulant extrema," *IEEE Signal Processing Mag*., vol. 13, pp. 24-42, May 1996.

[8] C. Y. Chi and M. C. Wu, "A unified class of inverse filter criteria using two cumulants for blind deconvolution and equalization," in Proc. *IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Detroit, MI, 9-12 May 1995, pp. 1960-1963.

[9] C.-Y. Chi and M.-C. Wu, "Inverse filter criteria for blind deconvolution and equalization using two cumulants," *IEEE Trans. Signal Process.*, vol. 43, no. 1, pp. 55-63, Apr. 1995.

[10] D.L. Donoho, "On minimum entropy deconvolution," in *Applied Time Series Analysis II,* D.F. Findly, Ed. New York: Academic, 1981.

[11] W. Gray, "Variable norm deconvolution," *Ph.D. dissertation*, Dept. Geophysics, Stanford Univ., Stanford, CA, 1979.

[12] O. Shalvi and E. Weinstein, "Universal methods for blind deconvolution," in *Blind Deconvolution*, S. Haykin, Ed. Englewood Cliffs, NJ: Prentice-Hall, 1994.

[13] J.K. Tugnait, "Estimation of linear parametric models using inverse filter criteria and higher order statistics," *IEEE Trans. Signal Processing*, vol. 41, pp. 3196-3199, Nov. 1993.

[14] R.A. Wiggins, "Minimum entropy deconvolution," *Geoexploration*, vol. 16, pp. 21-35, 1978.

[15] M.C. Campi, R. Leonardi, and L.A. Rossi, "Generalized super-exponential method for blind equalization using Kautz filters," in Proc. *1999 IEEE Signal Processing Workshop Higher-Order Statistics*, Caesarea, Israel, 14-16 June 1999, pp. 107-111.

[16] J. Gomes and V. Barroso, "A super-exponential algorithm for blind fractionally spaced equalization," *IEEE Signal Processing Lett.*, vol. 3, pp. 283-285, Oct. 1996.

[17] F. Herrmann and A.K. Nandi, "Reduced computation blind super-exponential equaliser," *Electron. Lett*., vol. 34, no. 23, pp. 2208-2210, 12 Nov. 1998.

[18] F. Herrmann and A.K. Nandi, "Super-exponential equalizer – a modified eigenvector algorithm (MEVA)," in Proc. *1999 IEEE Signal Processing Workshop Higher-Order Statistics*, Caesarea, Israel, 14-16 June 1999, pp. 93-96.

[19] G. Scarano and G. Jacovitti, "On the optimality of Bussgang and super exponential blind deconvolution methods," in Proc. *1997 IEEE Signal Processing Workshop Higher-Order Statistics*, Banff, Canada, 21-23 July 1997, pp. 244-247.

[20] O. Shalvi and E. Weinstein, "Super-exponential methods for blind deconvolution," *IEEE Trans. Inform. Theory*, vol. 39, pp. 504-519, Mar. 1993.

[21] A.G. Bessios and C.L. Nikias, "POTEA: The power cepstrum and tricoherence equalization algorithm," *IEEE Trans Commun*., vol. 43, pp. 2667-2671, Nov. 1995.

[22] D.H. Brooks and C.L. Nikias, "Cross-bicepstrum and cross-tricepstrum pproaches to multichannel deconvolution," in Proc. *Int. Signal Processing Workshop Higher-Order Statistics*, Chamrousse, France, July 1991, pp. 91-94.

[23] I-T. Lu and J. Kwak, "Adaptive blind MIMO channel estimation and multiuser detection in DS-CDMA systems," in Proc. *1999 IEEE Global Telecommunications Conf.,* Rio de Janeiro, Brazil, 5-9 Dec. 1999, vol. 4, pp. 2254-2258.

[24] T. Nguyen and Z. Ding, "CMA beamforming for multipath correlated sources," in Proc. *IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Munich, Germany, 21-24 Apr. 1997, vol. 3, pp. 2521-2524.

[25] Y. Sato, “A method of self-recovering equalization for multilevel amplitude-modulation,” *IEEE Trans. Commun*., vol. COM-23, pp. 679–682, June 1975.

[26] D. N. Godard, “Self-recovering equalization and carrier tracking in two-dimensional data communication systems,” *IEEE Trans. Commun.,* vol. COM-28, pp. 1867–1875, Nov. 1980.

[27] C.R. Johnson P. Schniter, T.J. Endres, J.D. Behm, D.R. Brown, and R.A. Casas, “Blind equalization using the constant modulus criterion: A review,” Proc. *IEEE*, vol. 86, pp. 1927-1950, Oct. 1998.

[28] L.R. Litwin, “Blind channel equalization,” *IEEE Potentials*, vol. 18, pp. 9-12, Oct.-Nov. 1999.

[29] J.R. Treichler, I. Fijalkow, and C.R. Johnson, Jr., “Fractionally spaced equalizers: How long should they really be?,” *IEEE Signal Processing Mag*., vol. 13, pp. 65-81, May 1996.

[30] J.R. Treichler and M.G. Larimore, “New processing techniques based on the constant modulus adaptive algorithm,” *IEEE Trans. Acoustics, Speech, Signal Processing*, vol. 33, pp. 420-431, Apr. 1985.

[31] J.R. Treichler and B.G. Agee, “A new approach to multipath correction of constant modulus signals,“ *IEEE Trans. Acoustics, Speech, Signal Processing,* vol. 31, pp. 349-472, Apr.1983.

[32] A. Benveniste and M. Goursat, “Blind Equalizers,” *IEEE Trans. Commun.*, vol. COMM–32, pp. 871–883, Aug. 1984.

[33] G. Picci and G. Prati, “Blind equalization and carrier recovery using a stop–and–go decision directed algorithm,” *IEEE Trans. Commun.,* vol. COMM–35, pp. 877–887, Sept. 1987.

[34] T. Nguyen and Z. Ding, “Blind CMA beamforming for narrowband signals with multipath arrivals,” *Int. J. Adapt. Contr. Signal Process*., vol. 12, no. 2, pp. 157-172, Mar. 1998.

[35] C.B. Papadias and A.J. Paulraj, “A constant modulus algorithm for multiuser signal separation in presence of delay spread using antenna arrays,” *IEEE Signal Processing Lett*., vol. 4, pp. 178-181, June 1997.

[36] J. Sheinvald, “On blind beamforming for multiple non-Gaussian signals and the constant-modulus algorithm,” *IEEE Trans. Signal Processing*, vol. 46, pp. 1878-1885, July 1998.

[37] J.J. Shynk and R.P. Gooch, “The constant modulus array for co–channel signal copy and direction finding,” *IEEE Trans. Signal Processing*, vol. 44, pp. 652-660, Mar. 1996.

[38] J.K. Tugnait, “Blind spatio-temporal equalization and impulse response estimation for MIMO channels using a Godard cost function,” *IEEE Trans. Signal Processing*, vol. 45, pp. 268-271, Jan. 1997.

[39] J.K. Tugnait and T. Li, “Blind asynchronous multiuser CDMA receivers for ISI channels using code-aided CMA,” *IEEE J. Select. Areas Commun*. vol. 19, pp. 1520-1530, Aug. 2001.

[40] C. Xu, G. Feng, and K.S. Kwak, “A modified constrained constant modulus approach to blind adaptive multiuser detection,” *IEEE Trans. Commun*., vol. 49, pp. 1642-1648, Sept. 2001.

[41] Z. Ding, R.A. Kennedy, B.D.O. Anderson, and C.R. Johnson, “Ill– convergance of Godard blind equalizers in data communication systems,” *IEEE Trans. Commun*., vol. 39, pp. 1313-1327, Sept. 1991.

[42] C. R. Johnson, “Admissibility in blind adaptive channel equalization,” *IEEE Contr. Syst. Mag.,* pp 3–15, Jan. 1991.

[43] O. Shalvi and Weinstein, “New criteria for blind deconvolution of nonminimum phase systems (channels),” *IEEE Inform. Theory*, vol. 36, no. 2, pp. 312–321, Mar. 1990.

[44] W.A. Gardner, “A New Method of Channel Identification”, *IEEE Trans. On Commun.,* vol. 39(6), pp. 813-817, June 1991.

[45] L. Tong, G. Xu and T. Kailath, “A new approach to blind identification and equalization of multipath channels,.” in Proc. *Asilomar Conf. Signals, Syst., Comput.*, pp. 856-860 Nov. 1991.

[46] L. Tong, G. Xu, and T. Kailath, “Blind Identification and Equalization using Spectral Measures, Part II: A Time Domain Approach," In *W.A. Gardner Ed.,* editor, *Cyclostationarity in Communications and Signal Processing*. IEEE Press, 1993.

[47] B. Xerri and B. Borloz, “An iterative method using conditional second order statistics applied to blind source separation problem,” *IEEE Trans. Signal Processing*, vol. 52, pp. 313 – 328, Feb. 2004.

[48] A. Farid, Z. Luo, Z. Ding, “Blind channel equalization based on second order statistics,” in Proc. *IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Mar. 2005, pp. III-557 – III-560.

[49] Z. Xu, P. Liu and X. Wang, ”Blind multiuser detection: From MOE to subspace methods ,” *IEEE Trans. Signal Processing*, vol. 52, pp. 510 – 524, Feb. 2004.

[50] F. A. Dietrich and W. Utschick, “Pilot assisted channel equalization based on second order statistics,” *IEEE Trans. Signal Processing*, vol. 53, pp. 1178 – 1193, Mar. 2005.

[51] K. Rahbar, J. P. Reilly and J. H. Manton, “Blind identification of MIMO FIR systems driven by quasistationary source using second order statistics: A frequency domain approach,” *IEEE Trans. Signal Processing*, vol. 52, pp. 406 – 417, Feb. 2004.

[52] J. K. Tugnait and W. Luo, “Blind identification of time varying channels , using multistep linear predictors,” *IEEE Trans. Signal Processing*, vol. 52, pp. 1739 – 1749, Jun. 2004.

[53] Y. Huang and P. M. Djuric, “A blind partical filtering detector of signals transmitted over flat fading channels,” *IEEE Trans. Signal Processing*, vol. 52, pp. 1891 – 1900, Jul. 2004.

[54] J. Liang and Z. Ding, “Blind MIMO system identification based on cumulant subspace decomposition,” *IEEE Trans. Signal Processing*, vol. 51, pp. 1457 – 1468, Jun. 2003.

[55] A. Napolitano and M. Tanda, “Doppler channel blind identification for non cicrcular transmission in multiple access systems,” IEEE Trans. Comm., vol. 52, pp. 2073 – 2078, Jan. 2004.

[56] M. Castella, J.C. Pesquet and A.P. Petropulu, "Family of Frequency and Time-Domain Contrasts for Blind Separation of Convolutive Mixtures of Temporally Dependent Signals," IEEE Trans. on Signal Processing, vol.53(1), pp. 107-120, January 2005.

[57] S. Buzzi, M. Lops and H. V. Poor, “Blind adaptive joint multiuser detection and equalization in dispersive differentially encoded CDMA channels,” IEEE Trans. Signal Processing, vol. 51, pp. 1880 – 1893, Jul. 2003.

[58] M. Wax and T. Kailath, “Detection of signals by information theoretic criteria,” *IEEE Trans. Audio, Speech, Signal Processing*, vol. ASSP-33, pp. 387–392, Apr. 1985.

[59] G. Xu, H. Liu, L. Tong, and T. Kailath, “A least-squares approach to blind channel identification,” *IEEE Trans. Signal Processing*, vol. 43, pp. 2982–2993, Dec. 1995.

[60] Y. Hua and M. Wax, “Strict identifibility of multiple FIR channels driven by an unknown arbitrary sequence,” *IEEE Trans. Signal Processing*, vol. SP-44, pp. 756–759, Mar. 1996.

[61] M. L. G¨urreli and C. L. Nikias, “EVAM: An eigenvector-based deconvolution of input coloured signals,” *IEEE Trans. Signal Processing*, vol. 43, pp. 134–149, Jan 1995.

[62] M. I. G¨urreli and C. L. Nikias, “A new eigenvector-based algorithm for multichannel blind deconvolution of input coloured signals,” in Proc. *IEEE Intl. Conf. Acoustics, Speech, Signal Proc.*, Minneapolis, MN, Apr. 1993, vol. 4, pp. 448–451.

[63] L. A. Baccala and S. Roy, “A new blind time-domain channel identificaiton method based on cyclostationarity,” presented at *26th Conf. Information Sciences and Systems*, Princeton, NJ, Mar. 1994.

[64] L. A. Baccala and S. Roy, “A new blind time-domain channel identification method based on cyclostationarity,” *IEEE Signal Processing Lett*., vol. 1, pp. 89–91, June 1994.

[65] G. Dong and R. Liu, “Adaptive blind channel identification,” in Proc. *1995 IEEE Int. Conf. Communications*, Dallas, TX, June 1996, pp. 828–831.

[66] G. Dong and R. Liu, “An orthogonal learning rule for null-space tracking with implementation to blind two-cahnnel identification,” *IEEE Trans. Circuits Syst. I*, vol. 45, pp. 26–33, Jan 1998.

[67] L. Tong and Q. Zhao, “Blind channel estimation by least squares smoothing,” in Proc. *1998 IEEE Intl. Conf. Acoustics, Speech, Signal Processing*, Seattle, WA, vol. 5, pp. 2121–2124.

[68] L. Tong and Q. Zhao, “Joint order detection and channel estimation by least squares smoothing,” presented at *30th Conf. Information Sciences and Syst*ems, Princeton, NJ, Mar. 1998.

[69] L. Tong, G. Xu, and T. Kailath, "A new approach to blind identification and equalization of multipath channels," presented at *25th Asilomar Conf.*, Pacific Grove, CA, Nov. 1991.

[70] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on secondorder statistics: A time domain approach," *IEEE Trans. Inform. Theory*, vol. 40, pp. 340–349, Mar. 1994.

[71] L. Tong and S. Perreau, "Multichannel blind identification: from Subspace to Maximum likelihood methods," *Proc. IEEE*, vol. 86, no. 10, October 1998.

[72] G. K. Kaleh and R. Vallet, "Joint parameter estimation and symbol detection for linear or nonlinear unknown dispersive channels," *IEEE Trans. Commun.*, vol. 42, pp. 2406–2413, July 1994.

[73] L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–285, Feb. 1989.

[74] M. Shao and C. L. Nikias, "An ML/MMSE estimation approach to blind equalization," in Proc. *ICASSP'94 Conf.*, Adelaide, Australia, vol. 4, pp. IV–569–572.

[75] D. M. Titterington, "Recursive parameter estimation using incomplete data," *J. Royal Statist. Soc. B*, vol. 46, no. 2, pp. 257–267, 1984.

[76] E. Weinstein, M. Feder, and A. Oppenheim, "Sequential algorithms for parameter estimation based on the Kullback–Leibler information measure," *IEEE Trans. Signal Processing*, vol. SP- 38, pp. 1652–1654, Sept. 1990.

[77] L. B. White and V. Krishnamurphy, “Adaptive blind equalization of FIR channels using hidden Markov Models,” presented at *Int. Conf. Communications*, Geneva, June 1993.

[78] L. B. White, S. Perreau, and P. Duhamel, “Reduced computation blind equalization for FIR channel input Markov models,” in Proc. *Int. Conf. Communications*, Seattle, WA, June 1995, pp. 993–997.

[79] G. Harikuma and Y. Bresler, “Analysis and comparative evaluation of techniques for multichannel blind equalization,” in Proc. *8th IEEE Signal Processing Workshop Statistical and Array Signal Processing*, Corfu, Greece, June 1996, pp. 332–335.

[80] C. B. Papadias, “Methods for blind equalization and identify– cation of linear channels,” *Ph.D. dissertation,* ´ Ecole Nationale Sup´erieure des T´el´ecommunications, Paris, France, Mar. 1995.

[81] M. Feder and J. A. Catipovic, “Algorithms for joint channel estimating and data recovery—Application to underwater communications,” *IEEE J. Oceanic Eng*, vol. 16, pp. 42–55, Jan. 1991.

[82] N. Seshadri, “Joint data and channel estimation using fast blind trellis search techniques,” in Proc. *Globecom’90*, 1991, pp. 1659–1663.

[83] M. Ghosh and C. L. Weber, “Maximum-likelihood blind equalization,” *Opt. Eng*., vol. 31, no. 6, pp. 1224–1228, June 1992.

[84] D. Slock, “Blind fractionally-spaced equalization, perfect reconstruction filter banks, and multilinear prediction,” presented at *ICASSP’94,* Adelaide, Australia.

[85] D. Slock, “Blind joint equalization of multiple synchronous mobile users using oversampling and/or multiple antennas,” presented at *28th Asilomar Conf. Signals, Systems, Computers*, Monterrey, CA, Oct. 1994.

[86] D. Slock and C. B. Papadias, “Further results on blind identification and equalization of multiple FIR channels,” in *IEEE Proc. Intl. Conf. Acoustics, Speech, Signal Processing,* Detroit, MI, Apr. 1995, pp. 1964–1967.

[87] K. Abed-Meraim, E. Moulines, and P. Loubaton, “Prediction error method for second-order blind identification,” *IEEE Trans. Signal Processing*, vol. 45, pp. 694–705, Mar. 1997.

[88] R. L. Valcarce and S. Dasgupta, “Blind channel equalization with coloured sources based on second order statistics: A linear prediction approach,” *IEEE Trans .Signal Processing*, vol. 49, pp. 2050-2059, Sep. 2001.

[89] R. Liu and G. Dong, “A fundamental theorem for multiple-channel blind equalization,” *IEEE Trans. Circuits and Systems*-I, pp. 472-473, May 1997.

[90] Y. Hua, “Fast maximum likelihood for blind identification of multiple FIR channels,” in Proc. *28th Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 1994.

[91] S. Haykin, “Adaptive Filter Theory, 4th edition,” *Pearson Education Asia*, pp. 803-804.

[92] Y. Hua, “Fast maximum likelihood for blind identification of multiple FIR channels,” *IEEE Trans. Signal Processing*, vol. 44, pp. 661–672, Mar. 1996.

[93] Y. Bresler and A. Macovski, “Exact maximum likelihood parameter estimation of supper imposed exponential signals in noise,” *IEEE Trans. Audio, Speech, Signal Processing*, vol. ASSP-34, pp. 1081–1089, Oct. 1986.

[94] S. Amari and A. Cichocki, “Adaptive blind signal processing –Neural network approaches,” *Proc. of the IEEE*, vol. 86, no. 10, pp. 2026–2048. CA, Oct. 1998.

[95] Y. Fang and T. W. S. Chow, “Blind equalization of noisy channel by linear neural networks,” *IEEE Trans. Neural Networks*, vol. 10, no. 4, pp. 918–924, July 1999.

[96] Y. Fang, T. W. S. Chow and K. T. Ng, “Linear neural network based blind equalization,” *Elsevier, Signal Processing 76*, pp. 37-42, 1999.

[97] S. Bellini, “Blind Equalization,” *Alta Frequenza*, vol. 57, pp. 445 – 450, 1988.

[98] S.Fiori, “Analysis of modified ‘Bussgang’ algorithm (MBA) for channel equalization,” *IEEE Trans. on Circuits and Systems - Part I*, vol. 51, no. 8, pp. 1552 - 1560, August 2004.

[99] S.Fiori, “Blind deconvolution by simple adaptive activation function neuron,” *Neurocomputing*, vol. 48, No.1-4, pp. 763 - 778, Oct.2002.

[100] S. Fiori, A. Uncini, F. Piazza, “Blind deconvolution by modified Bussgang algorithm,” ISCAS’99, International symp. on Circuits and Systems, Orlando, Florida, May 30–June 2, 1999.

[101] S. Haykin, “Adaptive Filter Theory, 4th edition,” *Pearson Education Asia*, pp. 684-735.

[102] K. Rahbar, J. P. Reilly, and J. H. Manton, “Blind Identification of MIMO FIR systems Driven by Quasi–stationary Sources Using Second-Order Statistics: A

Frequency Domain Approach," *IEEE Trans. Signal Processing*, vol. 52, pp. 406-417, Feb. 2004.

[103] G. B. Giannakis, "Linear cyclic correlation approach for blind identification of FIR channels," in *Proc. 28th Asilomar Conf. Signals, Systems, and Computers,* Pacific Grove, CA, pp. 420–423, Nov. 1994.

[104] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, "Subspace-methods for the blind identification of multichannel FIR filters," presented at *ICASSP'94 Conf.*, Adelaide, Australia, Apr. 1994.

[105] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, "Subspace-methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, pp. 516–525, Feb. 1995.

[106] P. F. Baldi and K. Hornik, "Learning neural networks: A survey," *IEEE Trans. Neural networks*, vol. 6, pp. 837-858, Jul. 1995.

[107] A. Hyvarinen, J. Karhunen and E. Oja, "Independent Component Analysis," *J. Wiley*, Chapter on Blind Deconvolution and Convolutive Mixtures, 2001.

[108] J. Karhunen, E. Oja, L. Wang, R. Vigario, J. Joutsensalo, "A class of neural networks for independent component analysis," *IEEE Trans. Neural Networks*, vol. 8, no. 3, pp. 486-504, 1997.

[109] A. J. Bell and T. J. Sejnowski, "An information maximization approach to blind separation and blind deconvolution," *Neural Comput*., vol. 7, pp. 1129-1159, 1995.

[110] S. Haykin, "Unsupervised Adaptive Filtering, vol. I: Blind Source Separation," *J. Wiley*, Review chapters (8-9) by Torkkola, and Lambert & Nikias, 2000.

[111] K. H. Afkhamie and Z. Q. Luo, "Blind equalization of FIR systems driven by Morkov-like input signals," *IEEE Trans. Signal Processing*, vol. 48, pp. 1726-1736, June 2000.

[112] S. Haykin, "Adaptive Filter Theory, 4th edition," *Pearson Education Asia*, pp. 506-534.

[113] H. Liu, G. Xu, and L. Tong, "A deterministic approach to blind equalization," In Proc. *27th Asilimar conference on Signals, Systems & Computers*, pages 751–755, Pacific Grove, California, November 1993.

[114] A. Hussain, A. Naveed, I. Qureshi, "New Hybrid HOS-SOS Approach for Blind Equalization of Communication Channels," *IEE Electronics Letters*, vol.41, no.6, pp. 376-377, 2005.

# Appendix–A

## Proof of Theorem

### A.1 Theorem:

Suppose $\mathbf{H}$ and $\mathbf{s}(i)$ satisfy the linear model given by (2.2.4). $\mathbf{s}(i)$ are from white source and $\mathbf{H}$ is a full column rank matrix. There exists a linear transformation $\mathbf{W}$ i.e.

$$\mathbf{y}(i) = \mathbf{W}\mathbf{x}(i) \tag{A.1}$$

such that for $|\alpha| = 1$

$$\mathbf{W}\mathbf{H} = \alpha I \tag{A.2}$$

if, and only if, the following two conditions are satisfied

$$\mathbf{R}_{\mathrm{y}}(0) = \mathbf{I} \tag{A.3}$$

$$\mathbf{R}_{\mathrm{y}}(1) = \mathbf{J} \tag{A.4}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{J}$ is the shifting matrix.

### Proof:

Necessary condition is straightforward.

For sufficiency, substitute (2.2.4) into (2.2.12)

$$\mathbf{y}(i) = \mathbf{W}\mathbf{H}\mathbf{s}(i) \tag{A.5}$$

Correlation matrix with lag zero for symbols $y(i)$ is given as

$$\begin{aligned}\mathbf{R}_y(0) &= E\left[\mathbf{y}(i)\mathbf{y}^{\mathbf{H}}(i)\right] \\ &= E\left[\mathbf{WHs}(i)\mathbf{s}^{\mathbf{H}}(i)(\mathbf{WH})^{\mathbf{H}}\right] \\ &= \mathbf{WH}\,E\left[\mathbf{s}(i)\mathbf{s}^{\mathbf{H}}(i)\right](\mathbf{WH})^{\mathbf{H}} \\ &= (\mathbf{WH})(\mathbf{WH})^{\mathbf{H}} \\ &= \mathbf{I}\end{aligned}$$

where we have used (2.2.14)

Letting $\mathbf{Q} = \mathbf{WH}$, we get

$$\mathbf{QQ}^{\mathbf{H}} = \mathbf{I} \tag{A.6}$$

Similarly using (2.2.14) to find correlation matrix with lag one for symbols $y(i)$ and using (2.2.13) we get

$$\mathbf{QJQ}^{\mathbf{H}} = \mathbf{J} \tag{A.7}$$

Let $\mathbf{Q}_i$ be the $i^{\text{th}}$ column of $\mathbf{Q}$. (A.7) gives a Jordan chain of length $p$ $(= m + n - 1)$.

$$\mathbf{JQ}_1 = \mathbf{Q}_2;\ \ \mathbf{JQ}_2 = \mathbf{Q}_3;...,\ \ \mathbf{JQ}_{p-1} = \mathbf{Q}_p;\ \ \mathbf{JQ}_p = 0$$

The last equality in (A.7) $\mathbf{Q}_P = [0,...,0,\alpha I]$

Consequently matrix $\mathbf{Q} = \alpha\mathbf{I}$ i.e $\mathbf{WH} = \alpha\mathbf{I}$ by using the above Jordan chain

This implies

$$\begin{aligned}\mathbf{y}(i) &= \mathbf{WHs}(i) \\ &= \alpha\mathbf{s}(i)\end{aligned}$$

Since $\mathbf{Q}$ is a unitary matrix, it implies $\alpha\alpha^* = 1$.

# Appendix–B

## Derivation of the terms

The detailed derivations of the terms used in section 2.2.2 are carried out in this appendix as follows:

Given the cost function.

$$J(W) = \sum_{m=1}^{d} \sum_{n=1}^{d} \left\{ \left| \mathbf{A}_{nm} \right|^2 + \left| \mathbf{B}_{nm} \right|^2 \right\} \tag{B.1}$$

where

$$\mathbf{A}_{nm} = \left\{ \mathbf{R}_y(1) - \mathbf{J} \right\}_{nm} \tag{B.2}$$

$$\mathbf{B}_{nm} = \left\{ \mathbf{R}_y(0) - \mathbf{I} \right\}_{nm} \tag{B.3}$$

Both the terms of the cost function are based on the requirement, that the statistics at the destination must approach the statistics at the source, i.e., $\mathbf{R}_y(n) \to \mathbf{R}_s(n)$ for $n = 0,1,...$

The elements of weight matrix are updated according to SGSD algorithm

$$\Delta \mathbf{W} = -\eta \frac{\partial J(W)}{\partial \mathbf{W}} \tag{B.4}$$

where

$$\begin{aligned} \frac{\partial J(W)}{\partial \mathbf{W}_{kl}} = {} & 2(\mathbf{A}^{\mathrm{H}})_k \mathbf{W} E\left[ \boldsymbol{x}(j) x_l^*(j+1) \right] + 2\mathbf{A}_k \mathbf{W} E\left[ \boldsymbol{x}(j+1) x_l^*(j) \right] \\ & + 2(\mathbf{B}^{\mathrm{H}})_k \mathbf{W} E\left[ \boldsymbol{x}(j) x_l^*(j) \right] + 2\mathbf{B}_k \mathbf{W} E\left[ \boldsymbol{x}(j) x_l^*(j) \right] \end{aligned} \tag{B.5}$$

where $(\mathbf{A}^H)_k$, $\mathbf{A}_k$ are the $k^{\text{th}}$ rows of matrices $\mathbf{A}^H$, $\mathbf{A}$ and $(\mathbf{B}^H)_k$, $\mathbf{B}_k$ are the $k^{\text{th}}$ rows of matrices $\mathbf{B}^H$, $\mathbf{B}$ respectively. $x_l(j)$ is the $l^{\text{th}}$ element of vector $\mathbf{x}(j)$. The above equation can be written in the matrix form as

$$\begin{aligned}\frac{\partial J(W)}{\partial \mathbf{W}} &= 2\mathbf{A}^{\mathrm{H}}\mathbf{W}E\left[\boldsymbol{x}(j)\boldsymbol{x}^{\mathrm{H}}(j+1)\right]+2\mathbf{A}\mathbf{W}E\left[\boldsymbol{x}(j+1)\boldsymbol{x}^{\mathrm{H}}(j)\right] \\ &\quad +2\mathbf{B}^{\mathrm{H}}\mathbf{W}E\left[\boldsymbol{x}(j)\boldsymbol{x}^{\mathrm{H}}(j)\right]+2\mathbf{B}\mathbf{W}E\left[\boldsymbol{x}(j)\boldsymbol{x}^{\mathrm{H}}(j)\right]\end{aligned} \tag{B.6}$$

Since third and fourth terms are identical, therefore their effect can be added up. Using (2.2.11) and (2.2.12) for matrices $\mathbf{A}$ and $\mathbf{B}$, the updating formula for the whole weight matrix becomes

$$\begin{aligned}\Delta\mathbf{W} &= -\eta\begin{bmatrix}\left\{\mathbf{R}_y(1)-\mathbf{J}\right\}^{\mathrm{H}}\mathbf{W}\mathbf{R}_x(1) \\ +\left\{\mathbf{R}_y(1)-\mathbf{J}\right\}\mathbf{W}\mathbf{R}_x^{\mathrm{H}}(1) \\ +2\left\{\mathbf{R}_y(0)-\mathbf{I}\right\}\mathbf{W}\mathbf{R}_x(0)\end{bmatrix} \\ &= -\eta\begin{bmatrix}\left\{E\left[\mathbf{y}(j)\mathbf{y}^{\mathrm{H}}(j+1)\right]-\mathbf{J}\right\}^{\mathrm{H}}\mathbf{W}\mathbf{R}_x(1) \\ +\left\{E\left[\mathbf{y}(j)\mathbf{y}^{\mathrm{H}}(j+1)\right]-\mathbf{J}\right\}\mathbf{W}\mathbf{R}_x^{\mathrm{H}}(1) \\ +2\left\{E\left[\mathbf{y}(j)\mathbf{y}^{\mathrm{H}}(j)\right]-\mathbf{I}\right\}\mathbf{W}\end{bmatrix}\end{aligned} \tag{B.7}$$

If we wish to carry out the online learning process i.e. updating of weights for every new vector $\mathbf{x}(j)$ coming in, we remove the expectation operator

$$\left(\Delta\mathbf{W}\right)_{ol} = -\eta\begin{bmatrix}\left\{\mathbf{y}(j)\mathbf{y}^{\mathrm{H}}(j+1)-\mathbf{J}\right\}^{\mathrm{H}}\mathbf{W}\mathbf{R}_x(1) \\ +\left\{\mathbf{y}(j)\mathbf{y}^{\mathrm{H}}(j+1)-\mathbf{J}\right\}\mathbf{W}\mathbf{R}_x^{\mathrm{H}}(1) \\ +2\left\{\mathbf{y}(j)\mathbf{y}^{\mathrm{H}}(j)-\mathbf{I}\right\}\mathbf{W}\mathbf{R}_x(0)\end{bmatrix} \tag{B.8}$$

where the subscript '*ol*' stands for online. If weights are to be updated by batch processing, we carry out the expectation operator as follows

$$E\left[\mathbf{y}(j)\mathbf{y}^{\mathrm{H}}(j)\right]=E\left[\mathbf{W}\mathbf{x}(j)\mathbf{x}^{\mathrm{H}}(j)\mathbf{W}^{\mathrm{H}}\right]\ =\mathbf{W}\mathbf{R}_x(0)\mathbf{W}^{\mathrm{H}} \tag{B.9}$$

$$E\left[\mathbf{y}(j)\mathbf{y}^{\mathrm{H}}(j+1)\right]=E\left[\mathbf{W}\mathbf{x}(j)\mathbf{x}^{\mathrm{H}}(j+1)\mathbf{W}^{\mathrm{H}}\right]\ =\mathbf{W}\mathbf{R}_x(1)\mathbf{W}^{\mathrm{H}} \tag{B.10}$$

thus

$$\left(\Delta\mathbf{W}\right)_{bp}=-\eta\begin{bmatrix}\left\{\mathbf{W}\mathbf{R}_x^{\mathrm{H}}(1)\mathbf{W}^{\mathrm{H}}-\mathbf{J}^{\mathrm{H}}\right\}\mathbf{W}\mathbf{R}_x(1)\\+\left\{\mathbf{W}\mathbf{R}_x^{\mathrm{H}}(1)\mathbf{W}^{\mathrm{H}}-\mathbf{J}^{\mathrm{H}}\right\}^{\mathrm{H}}\mathbf{W}\mathbf{R}_x^{\mathrm{H}}(1)\\+2\left\{\mathbf{W}\mathbf{R}_x(0)\mathbf{W}^{\mathrm{H}}-\mathbf{I}\right\}\mathbf{W}\mathbf{R}_x(0)\end{bmatrix} \tag{B.11}$$

where the subscript '*bp*' stands for batch processing. It will give more flexibility to the updating formula if we choose $\eta_A$ and $\eta_B$ as step sizes for the two separate terms as given below

$$\begin{aligned}\left(\Delta\mathbf{W}\right)_{bp}=-\eta_A&\begin{bmatrix}\left\{\mathbf{W}\mathbf{R}_x^{\mathrm{H}}(1)\mathbf{W}^{\mathrm{H}}-\mathbf{J}^{\mathrm{H}}\right\}\mathbf{W}\mathbf{R}_x(1)\\+\left\{\mathbf{W}\mathbf{R}_x(1)\mathbf{W}^{\mathrm{H}}-\mathbf{J}\right\}\mathbf{W}\mathbf{R}_x^{\mathrm{H}}(1)\end{bmatrix}\\&-\eta_B\left[\left\{\mathbf{W}\mathbf{R}_x(0)\mathbf{W}^{\mathrm{H}}-\mathbf{I}\right\}\mathbf{W}\mathbf{R}_x(0)\right]\end{aligned} \tag{B.12}$$

where the constant 2 in the last term has been absorbed in $\eta_B$. The simulations and results for the work of Fang [95] have been carried out in chapter 3 and 4 and the comparisons have been made with our proposed algorithm.

# Appendix–C

## Construction of channels

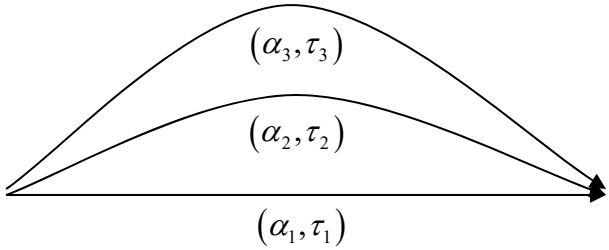



**Fig. C.1 Three ray multipath channel**

Consider a 3–ray multipath channel with attenuation constants given as $\alpha_1$, $\alpha_2$, $\alpha_3$ and propagation delays given as $\tau_1$, $\tau_2$, $\tau_3$. We take the propagation delay of direct path $\tau_1 = 0$ and hence $\tau_2$ and $\tau_3$ are considered as relative propagation delays. Hence the impulse response is given as:

$$h_1(t) = \alpha_1 \delta(t) + \alpha_2 \delta(t - \tau_2) + \alpha_3 \delta(t - \tau_3)$$

The transmitter waveform is a raised cosine pulse with roll off factor $\beta$ given as:

$$c(t, \beta) = \frac{\sin\left(\pi t / T\right)}{\left(\pi t / T\right)} \cdot \frac{\cos\left(\pi \beta t / T\right)}{\left(1 - 4\beta^2 t^2 / T^2\right)}$$

where $T$ is symbol interval and is an integer multiple of the sampling period.

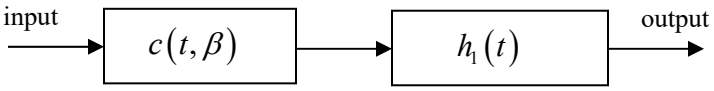



**Fig. C.2 Overall channel**

The overall impulse response of the channel becomes:

$$h(t) = \int_{-\infty}^{\infty} c(\tau,\beta)\, h_1(t-\tau)\, d\tau$$
$$= \alpha_1 c(t,\beta) + \alpha_2 c(t-\tau_2,\beta) + \alpha_3 c(t-\tau_3,\beta)$$

$h(t)$ is limited to some finite length given as $L_h = nT$, where $n$ is 4 or 6 in over case.

Taking $T$ as discrete sampling time which is an integer multiple of $\Delta$, sampling period, that is, $T = m\Delta$, we get discrete impulse response.

For Tong's [70] channel $\alpha_1 = 0.2$, $\alpha_2 = 0.4$, $\tau_2 = 2.5T$, $L_h = 6T$ and sampling at twice the baud rate i.e. $m = 2$, we get the following values:

$$\mathbf{h} = \begin{bmatrix} -0.02788 & 0.04142 & -0.07025 & 0.3874 & 0.3132 & -0.08374 \\ 0.009773 & -0.01959 & 0.08427 & 0.5167 & 0.01383 & -0.001258 \end{bmatrix}^T$$

Similarly for Valcarce channel

$$[\boldsymbol{h}_0 \ldots \boldsymbol{h}_3] = \begin{bmatrix} 0.0554 + 0.0165i & -1.3449 - 0.4523i & 1.0067 + 1.1524i & 0.3476 + 0.3153i \\ -0.8077 - 0.3183i & 0.4307 + 0.2612i & 1.2823 + 1.1456i & -0.3610 - 0.2743i \end{bmatrix}^T$$

Figure C.3 shows the zeros plots for both of channels.

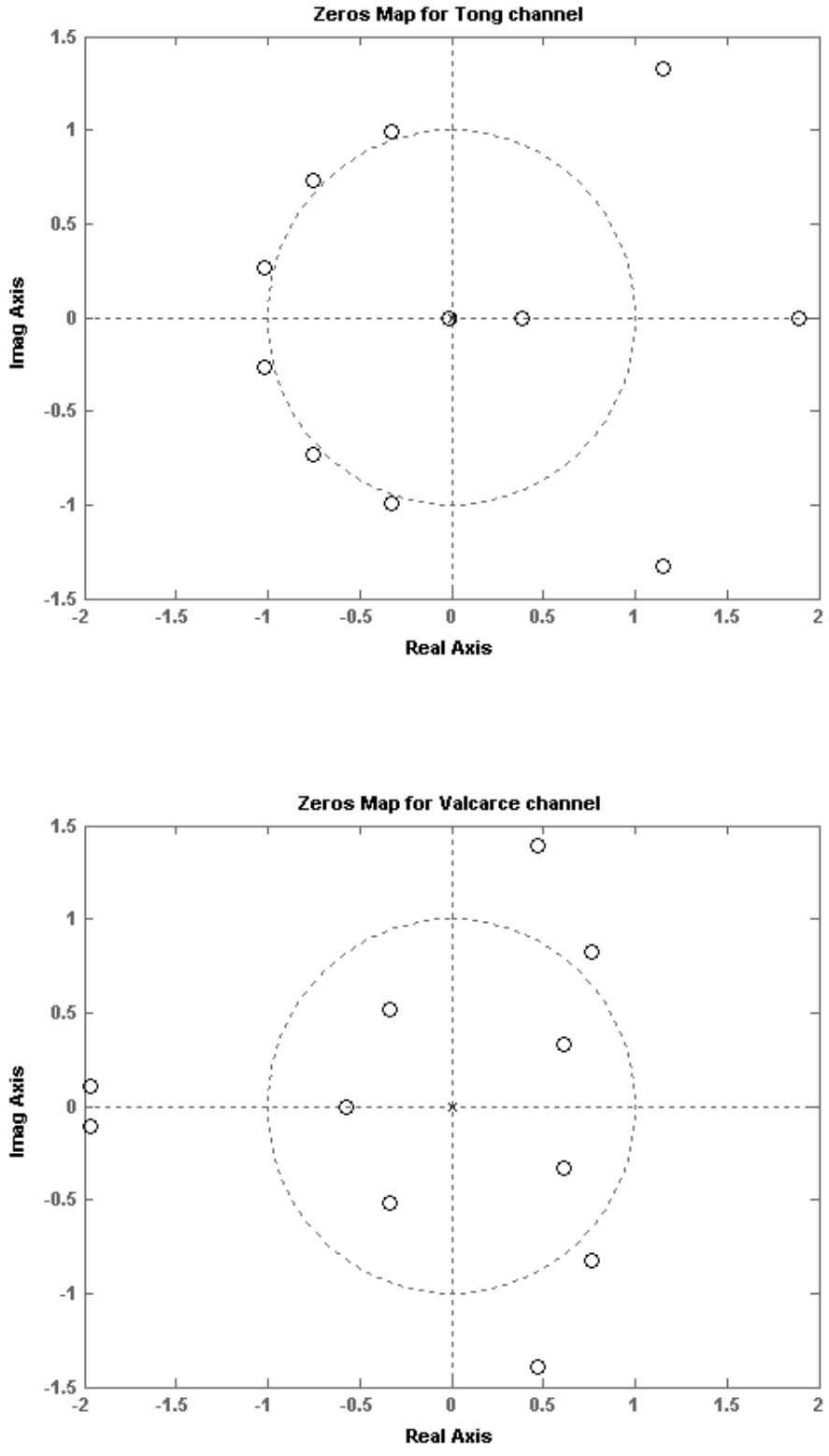



**Fig. C.3 Zeros plots for Tong and Valcarce channels**