ITERATIVE DECONVOLUTION USING SVD-KLT FOURTH ORDER CUMULANT

W.L.Woo and S.Sali, University of Newcastle Upon Tyne, UK

Abstract: This paper introduces improved blind deconvolution schemes for equalising a communication channel. Proposed algorithms have been developed and are in fact extensions of the existing one and some modifications have been made to speed up its convergence. Simulation results have been included in this paper to demonstrate the speed and effectiveness of the proposed algorithms.

1. INTRODUCTION

In blind deconvolution, the desired signal or input to the channel is unknown to the receiver. The only information available about the input signal is its probability distribution. Since both the channel and the input signal are unknown, the task of blind deconvolution is to recover the unknown input sequence based only on its probabilistic and statistical properties. This is illustrated in Fig. 1.

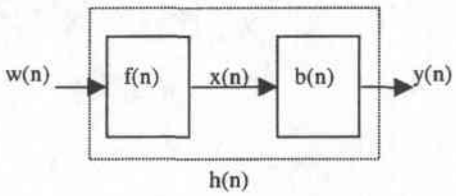

Fig. 1. Convolution-deconvolution model

In fig. 1, the input signal w(n) is transmitted and convolved with the channel impulse response f(n). The output of the channel denoted by x(n) is the signal received at the receiver. The channel f(n) may be a minimum or possibly a nonminimum phase system. The task of the equaliser is to deconvolve the sequence of x(n) by a filter denoted by b(n) so that the output is the recovered replica of the input w(n). It has already been proven that a sufficient condition for equalisation is that the probability distribution of the individual recovered sequence y(n) be equal to the probability distribution of the individual input sequence w(n).

Benveniste *et al.* [1] presents a comprehensive analytical study on blind equalisation. They establish that second order statistics of x(n) alone only provides the magnitude information of the linear channel and thus it is insufficient for blind equalisation of a nonminimum phase channel f(n) containing zeros both in the inside and outside of the unit circle. Hence, a nonminimum phase channel cannot be identified from its output when the input signal is i.i.d. Gaussian since second order moments completely characterises the input and output statistics of a linear system.

New criterion that is based on higher order statistics has been developed in [2,3]. New method requires the equalisation of only a few moments of the corresponding probability distributions. An important feature of the new criterion is that they are universal in the sense that they do not impose any restrictions on the probability distribution of the input sequence. In this paper, the proposed channel equalisation is based on the new criteria of higher order statistics.

2. PROBLEM FORMULATION

The following assumptions are made in the model developed in this paper:

1. The input sequence w(n) consists of zero mean independent and identically distributed (i.i.d.) real or complex random variables.
2. The unknown channel may be a nonminimum phase linear time invariant (LTI) system in which no zeros are placed on the unit circle.
3. Although the exact inverse of a nonminimum phase LTI channel is unstable, a truncated anticausal expansion can be delayed by a constant to allow a causal approximation of the inverse filter. This leads to our third assumption which states that the length of the equaliser be sufficiently long so as to avoid the anticausal truncation effect.

Considering fig. 1 , we can write for the output

$$y(n) = b(n) * f(n) * w(n) \tag{1}$$

where $*$ denotes the convolution operation. The convolution-deconvolution operator is defined as :

$$h(n) = b(n) * f(n) = f(n) * b(n) = \sum_k f(n-k)\, b(k) \tag{2a}$$

Therefore, (1) can be written as :

$$y(n) = h(n) * w(n) = \sum_k h(n-k)\, w(k) \quad \text{(2b)}$$

The deconvolution in z-domain can be written as:

$$Y(z) = B(z)\, F(z)\, W(z) \quad \text{(3a)}$$

The required convolution-deconvolution should satisfy the following criterion in general:

$$B(z) = \frac{\alpha\, e^{j\theta} z^{-m}}{F(z)} \quad \text{(3b)}$$

where α is any number, z^{-m} is the constant time delay as required for assumption 3 and $e^{j\theta}$ is the constant phase shift. Many researchers tackle the equalisation problem with $\alpha = 1$. The constant phase shift. is inherent when the probability distribution of the input sequence is symmetric under rotation. This phase ambiguity can be overcome by using differential encoding of the channel input.

From (2a), the convolution-deconvolution operation in z-domain can be represented as:

$$H(z) = F(z)\, B(z) = \alpha\, e^{j\theta} z^{-m} \quad \text{(4a)}$$

The inverse z-transform of (4a) can be defined as

$$h(n) = \alpha\, e^{j\theta}\, \delta(n-m) \quad \text{(4b)}$$

Eq. (1) can now be simplified into a more compact form using the above result in (4b). It can be proven that when (4b) is satisfied, the transfer function of the equaliser is a scaled and rotated inverse of the channel transfer function f(n). Hence the output of the equaliser is:

$$y(n) = \sum_k \alpha\, e^{j\theta} \delta(n-m-k)\, w(k) = \alpha\, e^{j\theta} w(n-m) \quad \text{(5)}$$

4. ALGORITHM FORMULATION

This section presents the criterion used for the equalisation based on the Higher Order Statistics.

For complex case and assuming i.i.d. input w(n), kurtosis of the deconvolved (recovered) sequence is defined as :

$$c_{4y}(0,0,0) = \mathrm{cum}[y^*(n), y(n), y(n), y^*(n)] = \sum_k \{h(k)\, h^*(k)\, h^*(k)\, h(k) \times \mathrm{cum}[w^*(n-k), w(n-k), w(n-k), w^*(n-k)]\} \quad \text{(6a)}$$

where cum denotes the cumulant operation and superscript (*) denotes the complex conjugate. Assuming fourth order stationarity of the input sequence, the whole equation can be rewrite as:

$$\gamma_{4y} = \gamma_{4w} \sum_k |h(k)|^4 \quad \text{(6b)}$$

where γ_{4y} and γ_{4w} are defined as the kurtosis of the deconvolved (recovered) and transmitted sequence respectively. Following the same line of derivation, the variance of the deconvolved sequence (for complex case) is defined as:

$$\gamma_{2y} = \gamma_{2w} \sum_k |h(k)|^2 \quad \text{(6c)}$$

where γ_{2w} is defined as the variance of the transmitted sequence. Having defined the kurtosis and variance, the normalised kurtosis is defined as:

$$k_{4y} = \frac{\gamma_{4y}}{|\gamma_{2y}|^2} = \frac{\gamma_{4w} \sum_k |h(k)|^4}{|\gamma_{2w} \sum_k \{|h(k)|^2\}|^2} = \frac{\gamma_{4w}}{|\gamma_{2w}|^2} \times \frac{\sum_k |h(k)|^4}{|\sum_k \{|h(k)|^2\}|^2} = k_{4w} \times \frac{\sum_k |h(k)|^4}{|\sum_k \{|h(k)|^2\}|^2} \quad \text{(6d)}$$

Eq. (6a) to (6d) form the basis of the new criterion for the equalisation. Since it is always true that $\sum_k |h(k)|^4 \le \left(\sum_k |h(k)|^2\right)^2$ the inequality of $|k_{4y}| \le |k_{4w}|$ will always be obeyed and equal if and only if the vector $\tilde{h}$ is in the form defined by (4b). Therefore, the task of equalisation based on this criterion is to equalise the absolute value of the recovered sequence normalised kurtosis with that of the transmitted sequence normalised kurtosis. One of the techniques used to satisfy the criteria is to maximise the absolute value of the recovered sequence normalised kurtosis i.e.

$$\max |k_{4y}| = \max [k_{4y} \cdot \text{sign}(k_{4y})] \tag{6e}$$

and use of gradient algorithm to direct the hunting of the maximum point. Defining the cost function J to be the absolute value of the recovered sequence normalised kurtosis, this cost function is always nonnegative i.e.

$$J = |k_{4y}| = k_{4y} \cdot \text{sign}(k_{4y}) \geq 0 \tag{7}$$

Since this cost function is always nonnegative and the required deconvolution is achieved only when maximisation is reached, we can therefore employ the steepest ascent algorithm to locate the maxima in the cost function. Cadzow [5] has studied several algorithmic approaches for solving the blind deconvolution problem and presented a general nonrecursive deconvolution algorithm based on maximising the normalised kurtosis using the steepest ascent algorithm. However, the algorithm converges slowly when the power spectral density of the convolved sequence has a large variation in itself, thereby increases the eigenvalue spread of the convolved sequence. Besides that, the algorithm also suffers from direction crash. Direction crash occurs quite often when the required deconvolved normalised kurtosis is in the vicinity close to zero and results in the algorithm moving in the wrong direction. In this paper, we propose a new algorithm based on the extension of the existing algorithm to complex case and some modifications to speed up its convergence and avoid direction crash. The cost function has been modified and is defined as:

$$J = |k_{4w} - k_{4y}|^2 \geq 0 \tag{8}$$

This cost function is also always nonnegative but the difference is that the minimum point here corresponds to the maximum point defined in (7). Cost function defined in this way avoids direction crash and the algorithm will always move in the same direction. The task of deconvolution is now to hunt for the minimum point so as to satisfy the criteria defined above. In particular, the new cost function requires a priori knowledge of the transmitted sequence normalised kurtosis. The latter parameter can be estimated as soon as when the probability density function of the transmitted input sequence is known since higher order cumulant is related to its moments. For example, if the transmitted sequence is of binary nature and has equal probability of transmitting 1's and 0's , then the normalised kurtosis as defined in (6d) is equal to −2. For a 4-QAM signal, it is approximately equal to −1.

Fast convergence has been observed when the convolved sequence is white, however, it is relatively slow when the convolved sequence has a large variation in the power spectral density. Therefore, in order to speed up the convergence, we proposed to use spectral prewhitening prior to the deconvolution algorithm. This can be achieved by using the Karhunen-Loéve Transform(KLT) which transforms a set of correlated random variables into a set of uncorrelated random variables. This in turn can be obtained via Singular Value Decomposition (SVD) a set of orthonormal eigenvectors.

A less stringent method to speed up its convergence speed without having to calculate the normalised kurtosis of the transmitted sequence but only requires the a priori knowledge of its sign is to orthogonalise the convolved sequence and use the cost function defined in (7) with a slight modification as follow:

$$J = k_{4y} \cdot \text{sign}(k_{4w}) \tag{9}$$

Algorithms based on this cost function will unlikely exhibit direction crash.

4. RESULTS

The proposed two algorithms are tested together with the existing algorithm based on (7). The channel response is described by

$$F(z) = \frac{-0.4 + z^{-1}}{1 - 0.4z^{-1}} \tag{10}$$

which is an allpass system. The input to the channel is a 4-QAM signal with equal probability of distribution. Results are shown in fig.2. (a) to 2(d). Fig.2. (a) shows the output of the channel. Fig.2. (b) shows the result of the conventional algorithm based on (7) after proper scaling and removing the constant time delay. Fig.2(c) and 2(d) show the results from the two new algorithms based on (8) and (9) respectively with KLT orthogonalisation after proper scaling and delay.

Results from fig.2 (b) to 2(d) show that all the algorithms produce similar results. This is expected

Initialisation

$\mathbf{b}(0) = [b_1(0)\ b_2(0)\ b_3(0)\ \cdots\cdots\ b_M(0)]^T = \mathbf{0}$ except for the centre tap - weight = 1 or the centre, left and right adjacent taps = 1

Perform Singular Value Decomposition on the convolved sequence:

$\mathbf{Q}^H\mathbf{Y}^H\mathbf{Y}\mathbf{Q} = \begin{bmatrix} \Sigma^2 & 0 \\ 0 & 0 \end{bmatrix}$ where **Y** is the overdetermined temporal data matrix of the convolved sequence

i.e. $\mathbf{Y} = [\mathbf{y}(n)\ \mathbf{y}(n+1)\ \mathbf{y}(n+2)\ \cdots\cdots\cdots\ \mathbf{y}(N)\,]^T$ and $\mathbf{y}(n) = [y(n)\ y(n-1)\ \cdots\ y(n-M+1)]^T$

Q is the orthonormal matrix containing the right singular vectors of data matrix **Y**

$\Sigma^2 = \mathrm{diag}[\sigma_1^2\ \sigma_2^2\ \sigma_3^2 \cdots\cdots \sigma_M^2]$ and σ_i is the i^{th} singular value of data matrix **Y**

Orthogonalise the convolved sequence via the singular vectors:

$\mathbf{V}(n) = \mathbf{Q}^H\mathbf{X}(n)$ where $\mathbf{V}(n) = [\nu_1(n)\ \nu_2(n)\ \cdots\cdots \nu_M(n)]^T$ and $\mathbf{X}(n) = [x(n)\ x(n-1) \cdots\cdots x(n-M+1)]^T$

for loops

$y(n) = \mathbf{b}^H(n)\mathbf{V}(n)$: output of equaliser , $\hat{k}_{4y} = \dfrac{\hat{\gamma}_{4y}}{|\hat{\gamma}_{2y}|^2}$: estimate of kurtosis as in equation (6)

estimate of gradient of the cost function defined by equation 8:

$$\nabla_b\left[|\hat{k}_{4w} - \hat{k}_{4y}|^2\right] = \nabla_b\left[(\hat{k}_{4w} - \hat{k}_{4y})^2\right] = 2(\hat{k}_{4w} - \hat{k}_{4y})\nabla_b\left[\hat{k}_{4w} - \hat{k}_{4y}\right] = -2(\hat{k}_{4w} - \hat{k}_{4y})\nabla_b\left[\hat{k}_{4y}\right]$$

update equation:

$$b_i(k+1) = b_i(k) + 2\alpha\frac{(\hat{k}_w - \hat{k}_y)}{\sigma_i^2}\nabla_b\left[\hat{k}_{4y}\right] \quad \text{where } \alpha \text{ is the step size and } \nabla_b[\bullet] = \frac{\partial}{\partial b}[\bullet]$$

alternatively

estimate of gradient of the cost function defined by equation 9:

$$\nabla_b\left[|\hat{k}_{4y}|\right] = \nabla_b\left[\hat{k}_{4y}\right]\mathrm{sign}(\hat{k}_{4w})$$

update equation:

while $|\hat{k}_{4y}(n+1)| \le |\hat{k}_{4y}(n)|$, start with $m = 0$

$$b_i(k+1) = b_i(k) + \kappa\cdot\frac{\beta^m}{\sigma_i^2}\nabla_b\left[\hat{k}_{4y}\right]\mathrm{sign}(\hat{k}_{4w}) \quad \text{where } \kappa \text{ is the multiplicative step size and } 0 < \beta < 1$$

$m = m + 1$

end

repeat the loops until $|\hat{k}_{4w} - \hat{k}_{4y}|^2 < \delta$ where δ is the required mean square error

Summary of the proposed algorithms

as the output of the channel is spectrally white and prewhitening prior to filtering will achieve the same results. However, the existing conventional algorithm degrades when the power spectral density of the convolved sequence has a large variation in itself. This can be illustrated by:

$$F(z) = \frac{-0.4 + z^{-1}}{1 - 1.8z^{-1} + 0.9995z^{-2}} \qquad (11)$$

The transfer function has a pair of poles settle in close approximity to the unit circle. Fig. 2(e) shows the output of the channel, which is highly dispersive and fig. 2(f) shows the inability of the conventional steepest ascent algorithm to equalise the channel. Fig. 2(g) and 2(h) show the equalised results via the algorithm based on (8) and (9) respectively with KLT orthogonalisation after proper scaling and removing the constant delay.

All the figures plotted describe the constellation diagram of the convolved and deconvolved sequences. It is also highly informative to consider the error convergence of all the three algorithms with respect to the number of iterations. The error convergence for

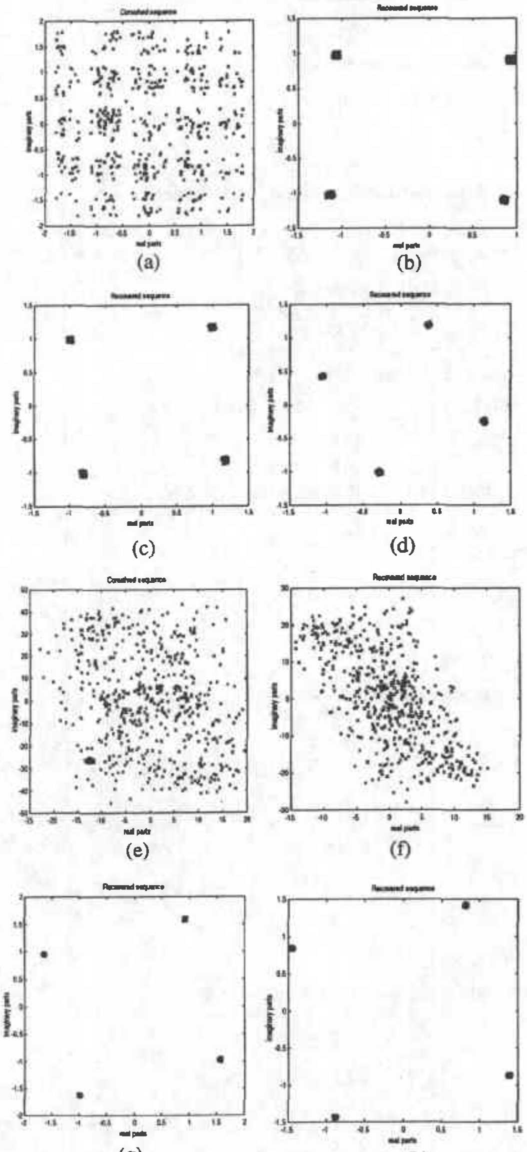

Figure 2(a) to 2(h)

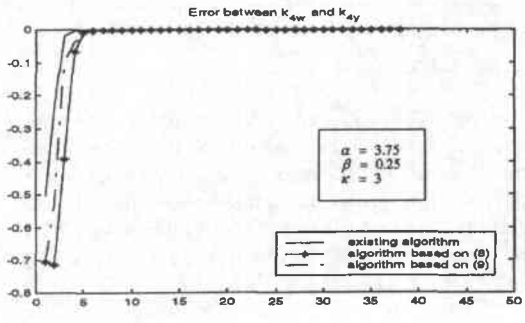

Figure 3: Error convergence between k_{4w} and k_{4y}

fig. 2(b) to 2(d) is plotted in fig. 3. A point worth noting is that the steady state error convergence of all the three algorithms occur almost simultaneously between the 5^{th} and the 7^{th} iteration. This is the direct implication of deconvolving a sequence which is spectrally white. No improvement is seen when using prewhitening prior to filtering.

The action of prewhitening can now be seen for the error convergence of the fig. 2(f) to 2(g). It can be seen

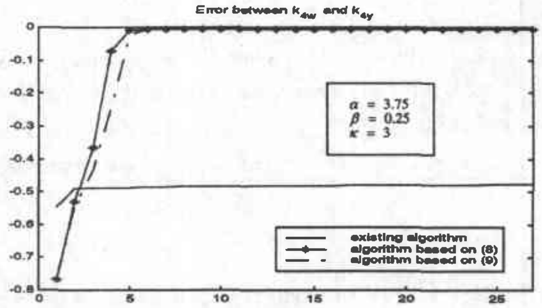

Figure 4 : Error convergence between k_{4w} and k_{4y}

from figure 4 that existing algorithm saturates after the 2^{nd} iteration. However, the other two algorithms converge at the 5^{th} and 6^{th} iteration leaving an unnoticeable error. This demonstrates the effectiveness of the two algorithms that employ SVD prewhitening.

5. CONCLUSION

In this paper, three algorithms have been presented which are nonrecursive. The optimisation route based on the cost function as defined in (7) has demonstrated fast convergence when the convolving sequence is white, however convergence speed is affected seriously when the convolved sequence has large variation in the power spectral density. Results have shown that the two algorithms introduced here achieved better results where the conventional one fails. Use of SVD-KLT prewhitening prior to filtering improves the convergence speed considerably albeit at more computational complexity.

6. REFERENCES

[1]. Benveniste, M. Goursat, and G. Rouget, "Robust Identification of a Nonminimum Phase System," *IEEE Transaction on Automatic Control*, vol.AC-25, no. 3, pp.385-399, June 1980.

[2]. D. L. Donoho, "On Minimum Entropy Deconvolution," Applied Time Series Analysis II, D.F.Findley, editor, New York, Academic Press, 1981.

[3]. O. Shalvi and E. Weinstein, "New Criteria for Blind Deconvolution of Nonmimimum Phase Systems," *IEEE Transaction on Information Theory*, vol.36, no.2, March 1990

[4]. J. A. Cadzow, "Blind Deconvolution via Cumulant Extrema," *IEEE SP Mag*, May 1996