# OPTIMUM DELAY AND MEAN SQUARE ERROR USING CMA

*D J Brooks*, S Lambotharan and J A Chambers*

Signal Processing Section, Department of Electrical and Electronic Engineering
Imperial College, Exhibition Road, London, UK. SW7 2BT
Telephone +44 171 589 5111 ext. 56183. Email brooks@ee.ic.ac.uk

## ABSTRACT

The performance of the Constant Modulus Algorithm can suffer because of the existence of local minima with large Mean Squared Error(MSE). This paper presents a new way of obtaining the optimum MSE over all delays using a second equalizer under a mixed Constant Modulus and Cross Correlation Algorithm (CM-CCA). Proof of convergence is obtained for the noiseless case. Simulations demonstrate the potential of the method.

## 1. INTRODUCTION

The Constant Modulus Algorithm(CMA) [4, 9] has proved to be very successful for blind equalization. However, the performance of the CMA can suffer because of the existence of local minima with correspondingly large Mean Squared Error(MSE)[3]. It has been shown that for a Fractionally Spaced Equalizer (FSE) that satisfies the zeros and length constraints, CMA is guaranteed to reach an open-eye solution [2]. When noise is added, this is no longer the case, and the CMA may find a solution which has high MSE[8].

This leads to a desire to find appropriate ways to initialise the algorithm. Centre-spike initialisation does not guarantee convergence to a good solution[8] . More recently, the same authors suggest Channel Surfing Reinitialisation(CSR) as a way of finding good solutions. This paper describes a new method of reinitialisation which is computationally efficient.

The method relies on recent work in [6, 7] which demonstrates simultaneous blind multiple source reconstruction, and relies on a cross correlation addition to the standard CMA cost function to form the mixed Constant Modulus and Cross Correlation Algorithm (CM-CCA).

*This work was supported by EPSRC and Philips Paging, St. Andrews Road, Cambridge, UK.

## 2. THE MODEL

The problem is that of equalizing an FIR channel with zero mean Additive White Gaussian Noise (AWGN) using an FIR equalizer. The signals are oversampled at $L$ times the baud rate. The input symbols ($s(k)$) are assumed to be Independent and Identically Distributed (IID) and where $k$ is the discrete time index. Each of the $L$ sub-channels ($\mathbf{h}_l(k)$) is of order $M$ and has $M+1$ taps, and each sub-equalizer ($\mathbf{w}_l(k)$) is of order $N$, giving $(N+1)L$ taps for $\mathbf{w}(k)$. Defining the channel matrix $\mathbf{H}$ as

$$\begin{bmatrix} h_1(0) & \dots & h_1(M) & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ h_L(0) & \dots & h_L(M) & 0 & \dots & 0 \\ 0 & h_1(0) & \dots & h_1(M) & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & h_L(0) & \dots & h_L(M) & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & h_1(0) & \dots & h_1(M) \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & h_L(0) & \dots & h_L(M) \end{bmatrix}$$

and the symbol vector $\mathbf{s}(k) = [s(k), ..., s(k-N-M)]^t$, the input to the equalizer can be expressed as $\mathbf{x}(k) = \mathbf{H}\mathbf{s}(k) + \mathbf{u}(k)$, where $\mathbf{u}(k)$ is the noise vector, independent of $\mathbf{s}(k)$, $\mathbf{x}(k) = [x(k), ..., x(k-L(N+1)+1)]^t$. Defining all the equalizer tap weights in one vector as $\mathbf{w}(k) = \left[w_0, w_1, ..., w_{L(N+1)-1}\right]^t$, the output of the equalizer becomes $y(k) = \mathbf{w}^H\mathbf{x}(k)$, where $(.)^H$ denotes Hermitian transpose.

## 3. MIXED CM-CC ALGORITHM

The standard CMA cost function is given by $J_1 = E\left[\left(\gamma - |y(k)|^2\right)^2\right]$, where $\gamma = \frac{E\left[|s(k)|^4\right]}{E\left[|s(k)|^2\right]} = \sigma_s^2 k_s$

and $k_s = \frac{E\left[|s(k)|^4\right]}{E\left[|s(k)|^2\right]^2}$ is the kurtosis of $s(k)$. Using this criterion, and a sensible initialisation scheme, and in some cases certain [2], that the equalizer will converge to an open eye solution and hence $y(k) \approx s(k-d)e^{j\theta}$, where $\theta$ is an arbitrary but constant phase term. However, this does not guarantee a low MSE and some values of $d$ may produce large MSE. The idea behind this new method is to use a second equalizer with a different cost function to search for the value of $d$ which gives the minimum MSE.

The cost function of the second equalizer is specified as $J_2 = E\left[\left(\gamma - |y_2(k)|^2\right)^2\right]$ $+ \kappa \sum_{\delta=-(M+N),\delta\neq d_2}^{\delta=M+N} \left|E\left[y_2(k)y_1^*(k-\delta)\right]\right|^2$, where $\kappa \in \Re^+$ and $d_2 \in -(M+N), ... -1, 1, ..., M+N$. We can assume with high probability that the first equalizer has achieved convergence to an open-eye solution, by simply using CMA. Then the output $y_1(k) \approx s(k-d_1)e^{j\theta_1}$. Since $\delta = d_2$ is excluded from the cross-correlation, the second cost function is then minimised when $y_2(k) = s(k-d_2-d_1)e^{j\theta_2}$. By selecting values of $d_2 \neq 0$, we can obtain different symbol delays, that is $y_2(k) \neq y_1(k) = s(k-d_1)e^{j\theta_1}$. For fractionally spaced channels satisfying the zeros and length criterion, with no noise, the cost function has only minima corresponding to $s(k-d_2-d_1)e^{j\theta_2}$ if $0 < d1+d2 < M+N$, provided $\kappa > 2\sigma_s^2 k_s$. If the constraint on $d_1+d_2$ is not satisfied, then for $\kappa > 2\sigma_s^2 k_s$, there is only one minimum which is at the origin.

The standard CMA cost function can be expressed as a generalised version of that in [5] $J_1 = k_s\sigma_s^4 \sum_{i=0}^{p} |g_i|^4 + 2\sigma_s^4 \sum_{i=0}^{p}\sum_{l=0,l\neq i}^{p} |g_i|^2 |g_l|^2$ $+\left|E\left[s(k)^2\right]\right|^2 \sum_{i=0}^{p}\sum_{l=0,l\neq i}^{p} g_i^2 (g_l^*)^2 - 2\sigma_s^4 k_s ||\mathbf{g}||_2^2 + \sigma_s^4 k_s^2$, where $\mathbf{g}$ is the baud rate impulse response of the whole channel plus the equalizer of length $p+1 = M+N+1$and $(.)^*$ denotes complex conjugate. We can also write $J_2 = k_s\sigma_s^4 \sum_{i=0}^{p} |g_i|^4 + 2\sigma_s^4 \sum_{i=0}^{p}\sum_{l=0,l\neq i}^{p} |g_i|^2 |g_l|^2 +$ $\left|E\left[s(k)^2\right]\right|^2 \sum_{i=0}^{p}\sum_{l=0,l\neq i}^{p} g_i^2 (g_l^*)^2 - 2\sigma_s^4 k_s ||\mathbf{g}||_2^2$ $+ \sigma_s^4 k_s^2 + \kappa\sigma_s^2\left(||\mathbf{g}||_2^2 - |g_{d_1+d_2}|^2\right)$. Provided that the channel matrix, $H$, has full rank then $\nabla_{\mathbf{g}} J_2 = 0 \leftrightarrow \nabla_{\mathbf{w}} J_2 = 0$, so working with the gradient with respect to the channel and equalizer convolution is equivalent to using the gradient with respect to the equalizer alone.

If we define the complex gradient operator as $\nabla\mathbf{g} = \frac{1}{2}\left[\frac{\delta}{\delta\Re\{g_0\}} + \frac{j\delta}{\delta\Im\{g_0\}}, \cdots, \frac{\delta}{\delta\Re\{g_p\}} + \frac{j\delta}{\delta\Im\{g_p\}}\right]^t$, the gradient of $J_2$ becomes $\nabla_{\mathbf{g}}(J_2) = 2\Lambda g + 2\Phi g^*$ where $\Lambda = diag[\Lambda_0 + \kappa\sigma_s^2/2 \ldots \Lambda_{d_1+d_2} \ldots \Lambda_p + \kappa\sigma_s^2/2]$, $\Lambda_l = \sigma_s^4\left(|g_l|^2(k_s-2) + 2||\mathbf{g}||_2^2 - k_s\right)$ and $\Phi = diag[\Phi_0 \ldots \Phi_p]$, $\Phi_l = \left|E\left[s(k)^2\right]\right|^2 \sum_{i=0,i\neq l}^{p} g_i^2$. Defining the Hessian as $\nabla_{\mathbf{g}}\nabla_{\mathbf{g}}^H(J_2) = \Psi$, we have diagonal terms $\Psi_{l,l} = 2\sigma_s^4\left[|g_l|^2 k_s + 2||\mathbf{g}||_2^2 - k_s\right]$ $+\left\{\begin{array}{ll}\kappa\sigma_s^2 & l\neq d_1+d_2\\ 0 & l = d_1+d_2\end{array}\right\}$ and off diagonal terms, of the $lth$ row and $mth$ column, $\Psi_{l,m} = 8\sigma_s^4 g_l g_m^*$

There are five ways in which the gradient can become zero.

(i) $\mathbf{g} = \mathbf{0}$ where $\mathbf{0}$ is the zero vector. The Hessian then becomes $\sigma_s^2 diag[(-2\sigma_s^2 k_s + \kappa) \ldots -2\sigma_s^2 k_s \ldots (-2\sigma_s^2 k_s + \kappa)]$, which is negative definite if $\kappa < 2\sigma_s^2 k_s$ indicating a maximum at the origin and indefinite when $\kappa > 2\sigma_s^2 k_s$, because of the single negative value on the diagonal indicating a saddle point at the origin. Note that if we do not satisfy the constraints on $d_2$ and so $0 < d_1+d_2$ or $d_1+d_2 > M+N$ the single negative value in the diagonal disappears giving a positive definite Hessian and a minimum at the origin for $\kappa > 2\sigma_s^2 k_s$.

(ii) One $g_i \neq 0, i \neq d_1+d_2$ and all other $g_i = 0$

This corresponds to the selection of a particular delayed symbol $s(k-i)$. Setting the gradient equal to zero gives $|g_i|^2 = 1 - \frac{\kappa}{2\sigma_s^2 k_s}$, which gives a Hessain of $2\sigma_s^2 diag[(2\sigma_s^2 - \kappa/k_s) \ldots (2\sigma_s^2 - \kappa(2+k_s)/2k_s) \ldots (2\sigma_s^2 - \kappa/k_s)]$. For $\kappa < 4\sigma_s^2 k_s/(2+k_s)$, we have a positive definite Hessian, giving a minimum and meaning that the solution $s(k-i)$ is an achievable delay and undesirable minimum. When $4\sigma_s^2 k_s/(2+k_s) < \kappa < 2\sigma_s^2 k_s$, the Hessian is indefinite and is a saddle point. For values of $\kappa > 2\sigma_s^2 k_s$, the value of $|g_i|^2$ is negative and hence there is no stationary point.

(iii) $g_{d_1+d_2} \neq 0$ and all other $g_i = 0$

Setting the gradient to zero yields $|g_{d_1+d_2}|^2 = 1$, hence the output becomes $s(k-d_1-d_2)$ as desired. The Hessian becomes $\sigma_s^2 diag[(4\sigma_s^2+\kappa) \ldots 4\sigma_s^2 \ldots (4\sigma_s^2+\kappa)]$ which is positive definite for $\kappa > -4\sigma_s^2$. This corresponds to the global minimum.

(iv) $v > 1$ number of $g_i \neq 0, i \neq d_1 + d_2$ and all other $g_i = 0$

Here, we restrict the analysis to QAM constellations, which means $\left|E\left[s(k)^2\right]\right|^2 = 0$. General constellation analysis can be found in [1].

When the gradient is zero, this gives $|g_i|^2 = \frac{\sigma_s^2 k_s - 4\sigma_s^2 ||g||_2^2 - \kappa}{2\sigma_s^2(k_s - 2)}$. The right hand side of this equation holds for all non-zeros $|g_i|^2$, so all terms have the same modulus giving $||g||_2^2 = v\,|g_i|^2$. This gives $|g_i|^2 = \frac{\sigma_s^2 k_s - 4\sigma_s^2 ||g||_2^2 - \kappa}{2\sigma_s^2(k_s + 2(v-1))}$. Since $v > 1$, there are no solutions to this equation when $\kappa > 2\sigma_s^2 k_s$ because the left hand side is negative.

(v) $v > 0$ number of $g_i \neq 0$, plus $g_{d1+d2} \neq 0$ and all other $g_i = 0$

We also restrict the analysis to QAM in this section. Zeroing the gradient and proceeding as in (iv) yields $|g_i|^2 = \frac{k_s(2\sigma_s^2(k_s-2)-\kappa)}{2\sigma_s^2(k_s-2)(k_s+2v)}$ and $|g_{d1+d2}|^2 = \frac{2\sigma_s^2(k_s-2)+2v\kappa}{2\sigma_s^2(k_s-2)(k_s+2v)}$. If $k_s - 2 > 0$, $|g_i|^2 < 0$ when $\kappa > 2\sigma_s^2(k_s - 2)$. If $k_s - 2 < 0$, $|g_{d1+d2}|^2 < 0$ when $\kappa > 2\sigma_s^2(2 - k_s)$. In both cases $\kappa > 2\sigma_s^2 k_s$ ensures that there are no stationary points on the error surface.

Even if there is noise or we do not satisfy the above constraints, provided the first equalizer converges to an open eye solution, simulations suggest that the second equalizer will still converge to the solution with the desired delay.

Once the second equalizer has converged for a given $d_2$, the second part of the cost function must be switched to zero ($\kappa = 0$) to facilitate a fair comparison of the MSE of the two equalizers. If an improvement is found then the taps in the first equalizer can be substituted with those of the second. Then the process can be repeated for a different delay $d_2$ until all possiblities are exhausted. This process will then find the lowest MSE over all delays. It is possible to use as many extra equalizers as are desired to speed up the search process. For instance, the use of one equalizer to search positive $d_2$ and another to search negative $d_2$ may be desirable, but this increases the complexity.

There are several advantages over the CSR scheme. Firstly, no estimate of the autocorrelation of the equalizer input is needed, and no matrix inversions are required with the mixed CM-CCA scheme, both of which are required for CSR. The CSR scheme approximates the channel matrix for different delay shifts and for large shifts, this approximation can lead to starting points which are remote from the desired minimum. Starting from one particular minimum, all the other minima are exactly obtainable using the mixed CM-CCA scheme.

## 4. ADAPTIVE SCHEME EQUATIONS

By making the usual assumptions about independence of tap weights and symbols and taking the instantaneous values instead of expectation operators, the instantaneous estimate of the gradient of the cost function becomes $\hat{\nabla}_{\mathbf{w}_2}(J_2) = \left(4y_2(k)(\gamma - |y_2(k)|^2)^2 + 2\kappa b(k)\right)\mathbf{x}(k)$, where $b(k) = \sum_{\delta=-(M+N),\delta\neq d_2}^{\delta=M+N} E\left[y_2(k)y_1^*(k-\delta)\right] y_1^*(k-\delta)$. Since the taps of both equalizers are evolving, it is necessary to use a windowed estimate of the expected cross-correlation. The expectation of $y_2(k)y_1^*(k-\delta)$ was therefore replaced by a sample estimator $\hat{p}(k)$, where $\hat{p}(k+1) = \lambda\hat{p}(k) + (1-\lambda)y_2(k)y_1^*(k-\delta)$. Similar windowed measures can be used to compare the MSE of the two equalizers. The method for selecting $d_2$ in these simulations was to initialise $d_2 = 1$ and then to increase the magnitude of $d_2$ trying first positive then negative delays to find a better MSE. The value of $d_2$ was reset to unity magnitude on finding a better MMSE.

## 5. SIMULATION RESULTS

The first example is the same example as given in [8]. The channel is a symbol spaced AR(n) channel of the form $H(z) = \frac{1}{1+\alpha z^{-n}}$. In this instance, $\alpha = 0.5$ and $n = 1$. A two tap symbol rate equalizer was used with BPSK symbols ($\pm 1$) and $SNR = 20dB$. The cost function for CMA is shown in Figure 1. This shows four minima, two of which correspond to $\pm s(k-1)$ and two to $\pm s(k)$. When initialised using a centre spike to $w_1 = [0, 1]$, the equalizer converges to the solution for $s(k-1)$ with MSE of 0.2620. The cost function for the second equalizer, with the first equalizer taps assumed to be fixed, is shown for $\kappa = 2$ in Figure 2. This Figure shows that only two minima are now present. With the first equalizer at one of the local minima,

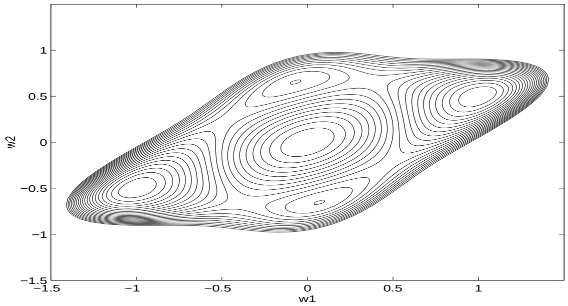


Figure 1: CMA for AR(1) channel

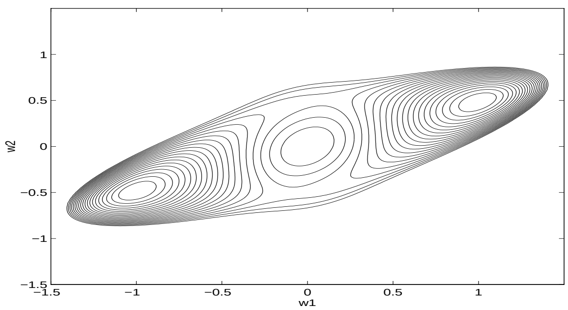


Figure 2: CM-CCA for AR(1) channel $\kappa = 2$

the second equalizer converged the global minima with MSE of 0.0205. Using a channel similar to the two ray model used in [8], the CM-CCA was tested in a fractionally spaced case. The impulse response was given by $h(t) = (0.1143 + 0.7740i)r(t) + (-0.4307 - 1.7330i)r(t - 3.3333T)$, where $T$ is the symbol period and $r(t)$ is the raised cosine function with roll-off 0.25 truncated to $6T$. Using 8PSK and $SNR = 15dB$, an 18 tap fractionally spaced equalizer was adapted using the mixed CM-CCA. After initialising the first equalizer by setting each tap to 1, while the rest remained at zero in different iterations, the final MSE was noted in each case. In all cases, the MSE was less than 0.0280, which is a very low value given the noise. All solutions eventually reached the optimum MSE over all delays. The speed at which this was achieved for different starting conditions depended on the way in which $d_2$ located the optimum delay from the starting point given by the first equalizer.

## 6. CONCLUSIONS

The mixed CM-CC method presented in this paper provides a method of finding the minimum MSE over all delays via a search technique using two or more equalizers. Provided that the CM algorithm can achieve an open eye solution to begin with, then the mixed CM-CC cost function appears to have only global minima. This is provable in the noiseless, fractionally spaced case with the usual zeros in common and length constraints. The use of the cross-correlation method in the DFE and in FSE-DFEs may also increase the probabililty of global convergence and is being investigated.

## 7. REFERENCES


[1] D. J. Brooks, S. Lambotharan, and J. A. Chambers. Optimum delay and Mean Square Error using a mixed Constant Modulus and Cross Correlation Algorithm. *In preparation for IEE Proceedings on Communications*, 1998.

[2] Z. Ding. Global Convergence of Fractionally Spaced Godard (CMA) Adaptive Equalizers. *IEEE Transactions on Signal Processing*, 44(4):818–826, April 1996.

[3] Z. Ding, R. A. Kennedy, B. D. O. Anderson, and C. R. Johnson, Jr. Ill Convergence of Godard Blind Equalizers in Data Communication Systems. *IEEE Transactions on Communications*, 39(4):1313–1327, April 1991.

[4] D. N. Godard. Self-Recovering Equalization and Carrier Tracking in Two-Dimensional Data Communication Systems. *IEEE Transactions on Communications*, 28(11):1867–1875, November 1980.

[5] C. R. Johnson, Jr. and B. D. O. Anderson. Godard Blind Equalizer Error Surface Characteristics: White, Zero-Mean, Binary Source Case. *International Journal on Adaptive Control and Signal Processing*, 9:301–324, 1995.

[6] S. Lambotharan and J. Chambers. On the Surface Characteristics of a Mixed Constant Modulus and Cross-correlation Criterion for the Blind Equalization of an MIMO Channel. *Accepted for European Signal Processing*, September 1997.

[7] C. Papadias and A. Paulraj. A Constant Modulus Algorithm for Multiuser Signal Separation in Presence of Delay Spread Using Antenna Arrays. *IEEE Signal Processing Letters*, 4(6):178–181, June 1997.

[8] L. Tong and H. H. Zeng. Channel Surfing Reinitialisation of CMA. *IEEE Signal Processing Advances in Wireless Communications Workshop*, pages 45–48, 1997.

[9] J. R. Treichler and B. G. Agee. A New Approach to Multipath Correction of Constant Modulus Signals. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 31:459–472, April 1983.