

LEAKY CONSTANT MODULUS ALGORITHMS: SENSITIVITY OF LOCAL MINIMA

S. Lambotharan, J.A. Chambers and A.G. Constantinides

Signal Processing and Digital Systems Section
 Department of Electrical and Electronic Engineering
 Imperial College, London, SW7 2BT
 E-mail: s.lambotharan@ic.ac.uk

ABSTRACT

We propose a new family of mixed constant modulus algorithms for the elimination of local minima associated with the fractionally spaced constant modulus algorithm in the presence of channel noise. A special case of this family is the Leaky Constant Modulus Algorithm (L-CMA). We show that L-CMA aims to minimise jointly the intersymbol interference (ISI) and the noise gain introduced by the equalizer. Moreover, we derive a suitable range of leakage factors for which all local minima due to large noise amplification are eliminated.

1. INTRODUCTION

The Constant Modulus Algorithm (CMA) is a widely used blind equalisation algorithm due to its robustness to various physical imperfections such as loss of disparity, channel length under modelling and channel noise. For an invertible fractionally spaced channel with no noise, all minima of CMA perform equally well [6]. However, when there is channel noise, various minima behave differently as the noise is amplified by the squared l_2 norm of the equalizer impulse response. Various techniques have been proposed to obtain convergence to a global minimum in the presence of noise. Typically, a solution based upon cross-correlation based equations, [4], and a channel surfing and re-initialisation technique, [8], which unfortunately require inversion of a matrix, have been proposed. In [1], a mixture of constant modulus and cross correlation criterion ([7] and [3]) was employed to obtain global convergence, but this requires a computationally demanding search over all possible minima of CMA. We propose a family of mixed constant modulus algorithms which have the potential for global convergence by minimising the following cost

$$J = E\{(|y(k)|^2 - 1)^2\} + \kappa \sum_{i=0}^{M-1} |w_i|^p \quad p \in \mathbb{Z}^+ \quad (1)$$

where $y(k)$ is the equalizer output and $\{w_i\}$ is the set of M equalizer impulse response coefficients, and κ is a mixing constant. We have shown elsewhere that the

case for $p = 1$ and $p = 2$ eliminates some of the local minima associated with CMA [5], but we did not produce then an adequate proof. For $p = 2$, (1) is a mixture of constant modulus criterion and $\|\mathbf{w}\|_2^2$ which is the noise amplification factor of the equalizer; hence for a suitable value of κ , (1) is expected to eliminate local minima which have large noise amplification. The algorithm update equation is given by

$$\mathbf{w}(k+1) = \mathbf{w}(k)(1-2\mu\kappa) - 4\mu(|y(k)|^2 - 1)y(k)\mathbf{x}(k) \quad (2)$$

where $\mathbf{w}(k)$ and $\mathbf{x}(k)$ are respectively the equalizer parameter vector and the regressor vector at sample number k , and μ is a small step size. We refer to this case as the Leaky Constant Modulus Algorithm (L-CMA). Two sub-channels were chosen for our simulation experiment, sub-channel 1 = $[0.6098 \quad -0.4878]^T$ and sub-channel 2 = $[0.9615 \quad 0.1923]^T$, and the channel noise power is taken to be 0.1 (i.e. SNR is 10 dB). The performance surfaces associated with CMA and L-CMA are depicted in Figure 1 for $\kappa = 0.26$ (discussion of the choice of κ will be provided later). There are two global minima and two local minima for CMA, but the local minima have disappeared in L-CMA.

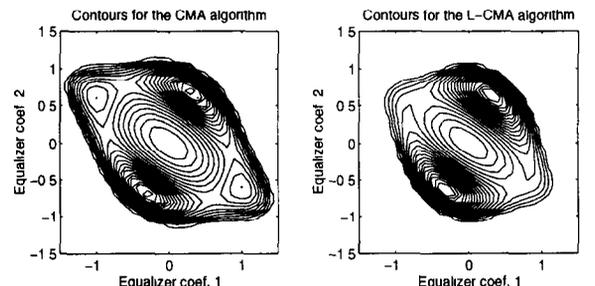


Figure 1: The contours of the performance surfaces for CMA and L-CMA

In the next section, we analyse the local convergence behaviour of L-CMA and provide suitable bounds of κ to avoid ill convergence to local minima.

2. MINIMA OF CMA AND L-CMA

The constant modulus cost function for a BPSK signal in terms of the combined channel and the equalizer impulse response $\{h_i\}$ of length N can be written as, [2],

$$\begin{aligned} J_{CMA} &= 1 - 2 \sum_{i=0}^{N-1} h_i^2 + 3 \left(\sum_{i=0}^{N-1} h_i^2 \right)^2 - 2 \sum_{i=0}^{N-1} h_i^4 (3) \\ &= (\|\mathbf{h}\|_2^2 - 1)^2 + 2I_h \end{aligned}$$

where $\|\mathbf{h}\|_2^2$ is the squared l_2 norm of $\{h_i\}$ and I_h characterises a measure of ISI as follows

$$I_h = \sum_{i \neq j} h_i^2 h_j^2 = \|\mathbf{h}\|_2^4 - \sum_{i=0}^{N-1} h_i^4 \quad (4)$$

The cost function associated with L-CMA can, therefore, be written as

$$J = (\|\mathbf{h}\|_2^2 - 1)^2 + 2I_h + \kappa \|\mathbf{w}\|_2^2 \quad (5)$$

where \mathbf{w} is the equalizer parameter vector. Assume that the channel convolution matrix \mathbf{C} to be invertible. Hence $\mathbf{w} = \mathbf{C}^{-1}\mathbf{h}$. Therefore, in principle, we have

$$J = (\|\mathbf{h}\|_2^2 - 1)^2 + 2I_h + \kappa \|\mathbf{C}^{-1}\mathbf{h}\|_2^2 \quad (6)$$

It is convenient to analyse the above cost function in terms of the radial and spherical components which is similar to the analysis of CMA in the presence of noise [2]. Hence, write $\mathbf{h} = r\mathbf{h}$ such that $\|\mathbf{h}\|_2 = 1$. Therefore,

$$J = (1 + 2I_h)r^4 + (\kappa \|\mathbf{C}^{-1}\mathbf{h}\|_2^2 - 2)r^2 + 1 \quad (7)$$

where $I_h = 1 - \sum_{i=1}^{N-1} h_i^4$. Minimisation with respect to r yields

$$r^2 = \frac{2 - \kappa \|\mathbf{C}^{-1}\mathbf{h}\|_2^2}{2(1 + 2I_h)} \quad (8)$$

Corollary 1: There are no nontrivial solutions for L-CMA if $\kappa > 2\lambda_{max}(\mathbf{C}\mathbf{C}^T)$.

□

Proof: The solution for r^2 is real only if $\kappa \|\mathbf{C}^{-1}\mathbf{h}\|_2^2 < 2$. However, $\lambda_{min}(\mathbf{C}\mathbf{C}^T)^{-1} \leq \|\mathbf{C}^{-1}\mathbf{h}\|_2^2 \leq \lambda_{max}(\mathbf{C}\mathbf{C}^T)^{-1}$. Therefore, no nontrivial solution exists for r if $\kappa > 2\lambda_{max}(\mathbf{C}\mathbf{C}^T)$.

□

From now onwards, it is assumed that κ is chosen to be less than $\frac{2}{\|\mathbf{C}^{-1}\mathbf{h}\|_2^2}$. Substituting (8) in (7) yields

$$J = 1 - \frac{1}{4} \frac{(2 - \kappa \|\mathbf{C}^{-1}\mathbf{h}\|_2^2)^2}{1 + 2I_h} \quad (9)$$

Hence $\arg \min_{\mathbf{h}} J = \arg \max_{\mathbf{h}} \zeta$, where

$$\zeta = \frac{(2 - \kappa \|\mathbf{C}^{-1}\mathbf{h}\|_2^2)^2}{1 + 2I_h} \quad (10)$$

For a suitable value of κ , maximisation of ζ means minimisation of I_h and $\|\mathbf{C}^{-1}\mathbf{h}\|_2^2$. i.e. *L-CMA aims to*

minimise both the ISI and the noise amplification factor.

Finding the stationary points of ζ means finding the roots of

$$\frac{\partial \zeta}{\partial \mathbf{h}} = \frac{2 - \kappa \|\mathbf{C}^{-1}\mathbf{h}\|_2^2}{1 + 2I_h} (\kappa(1 + 2I_h) \frac{\partial \|\mathbf{C}^{-1}\mathbf{h}\|_2^2}{\partial \mathbf{h}} + (2 - \kappa \|\mathbf{C}^{-1}\mathbf{h}\|_2^2) \frac{\partial I_h}{\partial \mathbf{h}}) \quad (11)$$

Taking the difficulty of finding a closed form solution for (11) into consideration, we aim to establish the behaviour of L-CMA for two special classes of channels, each of them leads to important observations.

Corollary 2: For orthogonal channels, L-CMA has no effect on the cost function other than a radial shrinking of the position of the local minima.

□

Proof: For orthogonal channels $\mathbf{C}\mathbf{C}^T = \lambda^{-1}\mathbf{I}$ where λ is the eigenvalue of $(\mathbf{C}\mathbf{C}^T)^{-1}$. Therefore, $\|\mathbf{C}^{-1}\mathbf{h}\|_2^2 = \lambda$ and $\zeta = \frac{(2 - \lambda\kappa)^2}{1 + 2I_h}$. Hence

$$\arg \max_{\mathbf{h}} \zeta = \arg \min_{\mathbf{h}} I_h$$

and $r_{min} = \sqrt{1 - 0.5\kappa\lambda}$, hence the proof.

□

An interesting consequence of Corollary 2 is that for orthogonal channels all minima perform equally well, hence there is no need to eliminate them, and L-CMA obeys this requirement. Figure 2 depicts the performance surfaces of L-CMA for different values of κ for an orthogonal channel (sub-channel 1 = $[0.6354 \ 0.3103]^T$ and sub-channel 2 = $[-0.3103 \ 0.6354]^T$).

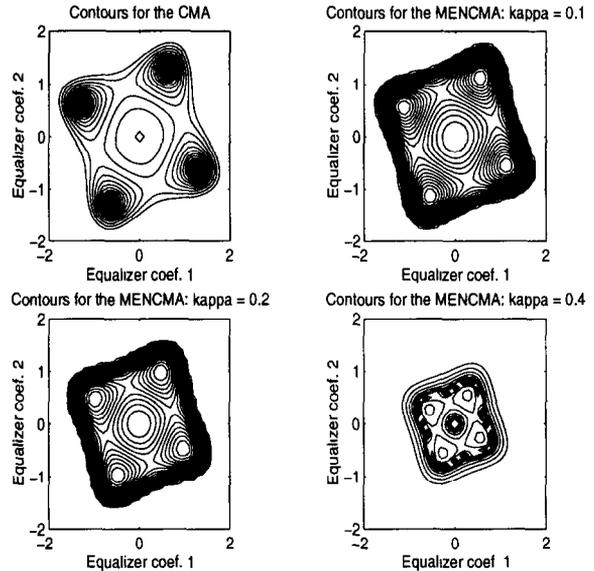


Figure 2: The performance surfaces of CMA and L-CMA

Corollary 3: For channels such that $\mathbf{C}\mathbf{C}^T$ is a diagonal matrix with distinct eigenvalues, a sufficient condition on κ to eliminate all minima except the two global minima is

$$2\lambda_{max2}(\mathbf{C}\mathbf{C}^T) < \kappa < 2\lambda_{max1}(\mathbf{C}\mathbf{C}^T)$$

where $\lambda_{max1}(\mathbf{C}\mathbf{C}^T)$ and $\lambda_{max2}(\mathbf{C}\mathbf{C}^T)$ are respectively the largest and the second largest eigenvalues of $\mathbf{C}\mathbf{C}^T$.

□

Proof: Write $\|\mathbf{C}^{-1}\mathbf{h}\|_2^2 = \lambda_0 + \sum_{i=1}^{N-1} (\lambda_i - \lambda_0)h_i^2$, where λ_i is the i^{th} eigenvalue of $(\mathbf{C}\mathbf{C}^T)^{-1}$. Hence

$$\zeta = \frac{(2 - \kappa \sum_{i=0}^{N-1} \lambda_i h_i^2)^2}{1 + 2I_h} \quad (12)$$

Lemma 1: The cost ζ is maximised only when $\mathbf{h} = \pm \mathbf{e}_v$, i.e. $\mathbf{h} = \{\mathbf{h} : I_h = 0\}$

□

where \mathbf{e}_v is the v^{th} coordinate vector. The proof is given in the Appendix. See also Figure 3 to view ζ for $N = 3$, $\lambda_0 = 0.1$, $\lambda_1 = 0.3$, $\lambda_2 = 0.6$ and $\kappa = 0.05$.

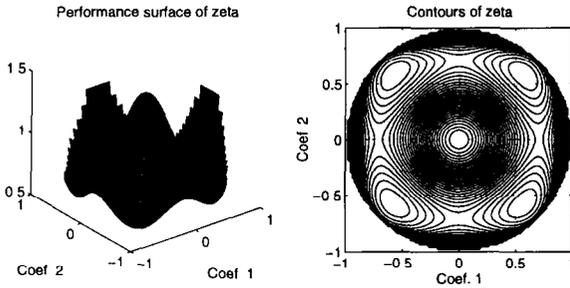


Figure 3: The performance surface and the contours of ζ

Hence h is considered only along the direction of the coordinate vector \mathbf{e}_v . Therefore, (8) yields $r_{min}^2 = 1 - 0.5\kappa\lambda_v$.

When κ is small ($\kappa < 2\lambda_{min}(\mathbf{C}\mathbf{C}^T)$), there are N possible nonnegative r^2 which are the $2N$ candidates for possible maxima of ζ . As κ increases the number of nonzero r^2 decreases.

When $2\lambda_{max2}(\mathbf{C}\mathbf{C}^T) < \kappa < 2\lambda_{max1}(\mathbf{C}\mathbf{C}^T)$, then there is only one non negative r^2 which is the candidate for the two global maxima on ζ , hence the proof.

□

When $\kappa = 2\lambda_{max2}(\mathbf{C}\mathbf{C}^T)$, $r_{min}^2 = 1 - \frac{\lambda_{max2}(\mathbf{C}\mathbf{C}^T)}{\lambda_{max1}(\mathbf{C}\mathbf{C}^T)}$ and $J = 1 - r_{min}^4$. The bias in the location of the minimum $B(r)$ and in the cost $B(J)$ can be respectively written as $B(r) = \frac{\lambda_{max2}(\mathbf{C}\mathbf{C}^T)}{\lambda_{max1}(\mathbf{C}\mathbf{C}^T)}$ and $B(J) = \frac{\lambda_{max2}(\mathbf{C}\mathbf{C}^T)}{\lambda_{max1}(\mathbf{C}\mathbf{C}^T)}(2 - \frac{\lambda_{max2}(\mathbf{C}\mathbf{C}^T)}{\lambda_{max1}(\mathbf{C}\mathbf{C}^T)})$. For a two dimensional channel + equalizer response case, $B(r) = \frac{1}{\chi(\mathbf{R})}$ and $B(J) = \frac{1}{\chi(\mathbf{R})}(2 - \frac{1}{\chi(\mathbf{R})})$ and the bias decreases as the condition number $\chi(\mathbf{R}) (= \frac{\lambda_{max}(\mathbf{C}\mathbf{C}^T)}{\lambda_{min}(\mathbf{C}\mathbf{C}^T)})$ increases. Suppose, $\lambda_0(\mathbf{R}) < \lambda_1(\mathbf{R}) \cdots \lambda_{N-2}(\mathbf{R}) < \lambda_{N-1}(\mathbf{R})$ and $\lambda_{j+1}(\mathbf{R}) \approx \lambda_{j+2}(\mathbf{R}) \approx \cdots \approx \lambda_{N-1}(\mathbf{R})$, it is suggested to choose $\kappa = 2\lambda_j(\mathbf{R})$ as in this case the $2j$ worst minima will disappear and the remaining $2(N-j)$ minima may perform approximately equally.

Even though the results provided in Corollary 3 are valid only when $(\mathbf{C}\mathbf{C}^T)^{-1} = \mathbf{Q}\mathbf{A}\mathbf{Q}^T$ and all columns of \mathbf{Q} are the coordinate vectors, for the two dimensional case, simulation results suggest that the above results are valid for arbitrary channels with columns of \mathbf{Q} which are only approximately the coordinate vectors. This is under current study.

3. CONCLUSION

We proposed and analysed a leaky constant modulus algorithm which jointly minimises ISI and the amount of noise amplification of the equalizer. L-CMA has potential to eliminate undesirable local minima, which have large noise amplification, associated with the fractionally spaced CMA. Moreover, our analytical result provides a range of leakage factor, κ , to avoid ill convergence of the equalizer to local minima.

Appendix: Proof of Lemma 1

Write $\mathbf{h} = [\sqrt{1 - \sum_{i=1}^{N-1} h_i^2} \ h_1 \ \cdots \ h_{N-1}]$, i.e., $h_0 = \sqrt{1 - \sum_{i=1}^{N-1} h_i^2}$, hence (12) can be written as

$$\zeta = \frac{(2 - \kappa \sum_{i=0}^{N-1} \lambda_i h_i^2)^2}{1 + 2(1 - \sum_{i=0}^{N-1} h_i^4)} \quad (13)$$

$$= \frac{((2 - \kappa\lambda_0) - \kappa \sum_{i=1}^{N-1} (\lambda_i - \lambda_0)h_i^2)^2}{1 + 4 \sum_{i=1}^{N-1} h_i^2 - 2(\sum_{i=1}^{N-1} h_i^2)^2 - 2 \sum_{i=1}^{N-1} h_i^4}$$

and substituting $h_i = r_1 t_i$ and $\mathbf{t} = \{\mathbf{t} : \sum_{i=1}^{N-1} t_i^2 = 1\}$,

$$\zeta = \frac{((2 - \kappa\lambda_0) + \kappa r_1^2(\lambda_0 - \sum_{i=1}^{N-1} \lambda_i t_i^2))^2}{1 + 4r_1^2 - 2r_1^4(1 + \sum_{i=1}^{N-1} t_i^4)} = \frac{n^2(t)}{d(t)} \quad (14)$$

and $r_1 \leq 1$ because $\sum_{i=1}^{N-1} h_i^2 \leq 1$.

Proposition 1: There are no maxima for ζ in $r_1 \in [0, 1]$.

□

Proof:

$$\frac{\partial \zeta}{\partial r_1} = 4 \frac{r_1 n(t) g(t)}{d^2(t)} \quad (15)$$

where $g(t) = 2((2 - \kappa \sum_{i=1}^{N-1} \lambda_i t_i^2) + (2 - \kappa\lambda_0) \sum_{i=1}^{N-1} t_i^4) r_1^2 - \kappa(\sum_{i=1}^{N-1} \lambda_i t_i^2 - \lambda_0) - 2(2 - \kappa\lambda_0)$. The roots of $n(t) = 0$, $g(t) = 0$ and $r_1 = 0$ are the stationary points of ζ on every radial directions. The second derivative at these stationary points can be written as

$$H = \frac{4}{d^2(t)} (r n(t) \frac{\partial g(t)}{\partial r_1} + r g(t) \frac{\partial n(t)}{\partial r_1} + g(t) n(t)) \quad (16)$$

The stationary point that corresponds to $n(t) = 0$ is

$$r_1^2 = \frac{-(2 - \kappa\lambda_0)}{\kappa(\lambda_0 - \sum_{i=1}^{N-1} \lambda_i t_i^2)} = \Gamma_1 \quad (17)$$

The stationary point that corresponds to $g(t) = 0$ is

$$r_1^2 = \frac{2(2 - \kappa\lambda_0) - \kappa(\lambda_0 - \sum_{i=1}^{N-1} \lambda_i t_i^2)}{(2 - \kappa\lambda_0)(1 + \sum_{i=1}^{N-1} t_i^2) + \kappa(\lambda_0 - \sum_{i=1}^{N-1} \lambda_i t_i^2)} = \Gamma_2 \quad (18)$$

These stationary points are examined for two different cases which corresponds to the sign of $\lambda_0 - \sum_{i=1}^{N-1} \lambda_i t_i^2$.

Case 1: $\lambda_0 - \sum_{i=1}^{N-1} \lambda_i t_i^2 > 0$

The value Γ_1 of (17) is negative in this case, and hence this solution is discarded. For the second stationary point $r_1^2 = \Gamma_2$ of (18), the second derivative is $H = \frac{16r_1^2}{d^2(r)}(2 - \kappa \sum_{i=1}^{N-1} \lambda_i t_i^2 + (2 - \kappa\lambda_0) \sum_{i=1}^{N-1} t_i^4)$ which is positive, hence this stationary point is a minimum. Therefore, there are no maxima for ζ in $r_1 \in [0, 1]$.

Case 2: $\lambda_0 - \sum_{i=1}^{N-1} \lambda_i t_i^2 < 0$

First consider the stationary point of $r_1 = 0$. The second derivative at this stationary point is $H = -4(2 - \kappa\lambda_0)(2(2 - \kappa\lambda_0) + \kappa(\sum_{i=1}^{N-1} \lambda_i t_i^2 - \lambda_0))$ which is negative and this stationary point is a maximum.

Substituting (17) into (18), the following coupled equation is obtained.

$$\Gamma_1 = \frac{1}{2} \frac{2\Gamma_2 + 1}{(1 + \sum_{i=1}^{N-1} t_i^2)\Gamma_2 - 1} \quad \Gamma_2 = \frac{1}{2} \frac{2\Gamma_1 + 1}{(1 + \sum_{i=1}^{N-1} t_i^2)\Gamma_1 - 1} \quad (19)$$

Therefore, since $0 < \sum_{i=1}^{N-1} t_i^2 \leq 1$, $\Gamma_1 \notin [0, 1]$ for $\Gamma_2 \in [0, 1]$ and vice versa. i.e., there is utmost one stationary point in $r_1 \in [0, 1]$. Suppose this stationary point is a maximum, since $r_1 = 0$ is also a maximum, there should be a minimum between these two maxima which contradicts with the proof that there is utmost one stationary point in $r_1 \in [0, 1]$. Hence, there are no maxima in $r_1 \in [0, 1]$

Therefore, considering case 1 and case 2, it can be concluded that there are no maxima for ζ in $r_1 \in [0, 1]$. \square

Proposition 1 implies that candidates for possible maxima of ζ are either when $r_1 = 0$ (i.e. $\mathbf{h} = \pm \mathbf{e}_1$) or when $r_1 = 1$. When $r_1 = 0$, the cost ζ can be written as

$$\zeta_1 = \frac{(2 - \kappa \sum_{i=1}^{N-1} \lambda_i t_i^2)^2}{1 + 2(1 - \sum_{i=1}^{N-1} t_i^4)} \quad \text{and} \quad \sum_{i=1}^{N-1} t_i^2 = 1 \quad (20)$$

which is in the same form as (14), but it is in $(N - 1)$ dimensions. Applying Proposition 1 again, the candidates for possible maxima of ζ_1 are either when all t_i are zero (i.e. $\mathbf{h} = \pm \mathbf{e}_2$) or when the associated radius is unity, i.e the maxima of ζ_2

$$\zeta_2 = \frac{(2 - \kappa \sum_{i=2}^{N-1} \lambda_i t_i^2)^2}{1 + 2(1 - \sum_{i=2}^{N-1} t_i^4)} \quad \text{and} \quad \sum_{i=2}^{N-1} t_i^2 = 1 \quad (21)$$

which is again in the same form as (14), but it is in $(N - 2)$ dimensions. We could apply Proposition 1 repeatedly until the dimension become 2, i.e.

$$\zeta_{N-2} = \frac{(2 - \kappa \sum_{i=N-2}^{N-1} \lambda_i t_i^2)^2}{1 + 2(1 - \sum_{i=N-2}^{N-1} t_i^4)} \quad \text{and} \quad \sum_{i=N-2}^{N-1} t_i^2 = 1 \quad (22)$$

Writing $t_{N-1} = t$ and $t_{N-2} = \sqrt{1 - t^2}$

$$\zeta_{N-2} = \frac{((2 - \kappa\lambda_{N-2}) + \kappa(\lambda_{N-2} - \lambda_{N-1})t^2)^2}{1 + 4t^2 - 4t^4} \quad (23)$$

and applying Proposition 1 yields candidates for possible maxima of (23) which are either $t = 0$, i.e. $\mathbf{h} = \pm \mathbf{e}_{N-1}$ or $t = 1$, i.e. $\mathbf{h} = \pm \mathbf{e}_N$.

Therefore, possible candidates for the maxima of ζ are $\mathbf{h} = \pm \mathbf{e}_v$, i.e. $\mathbf{h} = \{\mathbf{h} : I_h = 0\}$. \square

4. REFERENCES

- [1] D.J. Brooks, S. Lambbotharan and J.A. Chambers, "Optimum Delay and Mean Square Error Using CMA," ICASSP, Seattle, vol. 6, pp. 3361-3364, May 1998.
- [2] W. Chung and J.P. LeBlanc, "The Local Minima of Fractionally Spaced CMA Blind Equalizer Cost Function in the Presence of Channel Noise," ICASSP, Seattle, vol. 6, pp. 3345-3348, May 1998.
- [3] S. Lambbotharan, and J.A. Chambers, "On the Surface Characteristics of a Mixed Constant Modulus and Cross-Correlation for the Blind Equalisation of an MIMO Channel", accepted for Signal Processing, Sept. 1997.
- [4] S. Lambbotharan, J.A. Chambers and A.G. Constantinides, "A Novel Re-Initialization Technique for CMA in the Presence of Channel Noise," to appear in Asilomar Conf. on Sig., Sys. and Comp., Pacific Grove, CA, USA, Nov. 1998.
- [5] S. Lambbotharan, A.G. Constantinides and J.A. Chambers, "Channel Equalization Under Constant Modulus and Reduced Noise Amplification Objectives", IEEE DSP Workshop, Utah, August 1998.
- [6] Y. Li and Z. Ding, "Global Convergence of Fractionally Spaced Godard CMA Adaptive Equalizers," IEEE Trans. Sig. Proc., vol. 44, pp. 818-826, Apr. 1996.
- [7] 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 Letters, vol. 4(6), pp. 178-181, June 1997.
- [8] L. Tong and H.H. Zeng, "Channel-Surfing Re-initialization of CMA", IEEE workshop in Signal Processing Advances in Wireless Communications SPAWC, Paris, pp. 45-48, 1997.