# A FINITE-INTERVAL CONSTANT MODULUS ALGORITHM

*Phillip A. Regalia*

Department of Communications, Image and Information Processing
Institut National des Télécommunications
9, rue Charles Fourier
91011 Evry cedex France
`Phillip.Regalia@int-evry.fr`

**ABSTRACT**

A finite-interval constant modulus algorithm is developed which is vastly simpler than the Analytic Constant Modulus Algorithm and, unlike that algorithm, can claim to minimize a constant modulus criterion. It requires one QR decomposition of a data matrix, followed by a power iteration. Step size bounds which ensure monotonic convergence are obtained in analytic form, and proper tuning leads to an algorithm which converges typically within a few iterations. The algorithm thus gives a computationally feasible method for implementing constant modulus signal restoration in packet-based transmission systems.

## 1. INTRODUCTION

The constant modulus criterion [1]–[3] ranks among the most widely employed methods for blind signal restoration. The most common implementation, in terms of a stochastic gradient algorithm, often requires hundreds or even thousands of iterations to converge, depending on channel conditions, equalizer length, signal to noise ratio, etc., rendering the algorithm ill-suited to block processing or packet data based systems.

In response to this deficiency, van der Veen has developed the so-called Analytic Constant Modulus Algorithm (ACMA) [4], based on a clever technique to factor the received signal in a convolutional form involving a constant modulus signal. The factorization technique is a valid procedure provided perfect signal reconstruction is attainable (e.g., satisfaction of length and disparity conditions in high signal-to-noise ratios). Despite its name, the algorithm does not in general minimize a constant modulus cost function, although it can claim good behavior in practical conditions [5]. The algorithm presents a rather high computational complexity, though, and is complicated by rank conditions at intermediate steps.

Here we present an algorithm which directly minimizes a finite-interval constant modulus criterion, and is computationally much simpler than the ACMA: it requires one QR decomposition of a data matrix, followed by straightforward power method iterations. By proper tuning of the step size parameter, convergence usually occurs within a few iterations (between 5 and 10 iterations is typical). The algorithm thus represents a computationally feasible implementation of the constant modulus criterion for packet or block transmission systems.

## 2. PROBLEM SETTING

Here we develop an alternate formulation of the constant modulus criterion for finite data sets, and show how a simple gradient algorithm may be used to calculate solutions.

Consider a (real-valued) received vector $\mathbf{u}_n$ of the form

$$(P\,\updownarrow) \quad \mathbf{u}_n = \sum_{k=0}^{\infty} \mathbf{H}_k \mathbf{s}_{n-k} + \mathbf{b}_n$$

where $\{\mathbf{H}_k\}$ is the (multi-input/multi-output) channel impulse response, $\mathbf{s}_n$ contains the source signals, and $\mathbf{b}_n$ is the background noise.

The equalizer output is

$$y_n = \sum_{k=0}^{M} \mathbf{g}_k \mathbf{u}_{n-k} = \mathbf{g}\,\mathbf{U}_n$$

where each term $\mathbf{g}_k$ is a row vector containing $P$ components (so that $y_n$ is a scalar-valued sequence), and

$$\mathbf{g} = [\mathbf{g}_0 \;\; \mathbf{g}_1 \;\cdots\; \mathbf{g}_M], \qquad \mathbf{U}_n = \begin{bmatrix} \mathbf{u}_n \\ \mathbf{u}_{n-1} \\ \vdots \\ \mathbf{u}_{n-M} \end{bmatrix}$$

The constant modulus criterion, adapted to a finite time interval $n = 1, 2, \ldots, N$, becomes

$$J(\mathbf{g}) = \sum_{n=1}^{N} (1 - y_n^2)^2$$

0-7803-7402-9/02/$17.00 ©2002 IEEE

Consider now parametrizing the equalizer vector $\mathbf{g}$ in polar form (all norms are the Euclidean norm):

$$\mathbf{g} = \rho \bar{\mathbf{g}}, \qquad \text{with} \quad \begin{cases} \|\bar{\mathbf{g}}\| = 1; \\ \rho = \|\mathbf{g}\| \end{cases}$$

i.e., $\bar{\mathbf{g}}$ sets the angular orientation of the equalizer vector, while $\rho$ (the radial parameter) sets the length. If we let $\bar{y}_n = \bar{\mathbf{g}} \mathbf{U}_n$ denote the output of the unit-norm filter $\bar{\mathbf{g}}$, then $y_n = \rho \bar{y}_n$, and we may rewrite the constant modulus criterion in terms of $\bar{y}_n$ and $\rho$ as

$$J(\mathbf{g}) = \sum_{n=1}^{N} (1 - \rho^2 \bar{y}_n^2)^2 = \sum_{n=1}^{N} (1 - 2\rho^2 \bar{y}_n^2 + \rho^4 \bar{y}_n^4)$$

Consider now optimizing the radial parameter $\rho$, obtained by setting the derivative of $J(\rho \bar{\mathbf{g}})$ with respect to $\rho$ equal to zero:

$$0 = \frac{\partial J(\rho \bar{\mathbf{g}})}{\partial \rho} = 4\rho \sum_{n=1}^{N} (\rho^2 \bar{y}_n^4 - \bar{y}_n^2)$$

The choice $\rho = 0$ can be shown to result in a local maximum of $J$, while the minimum occurs when

$$\rho_{\text{opt}}^2 = \left( \sum_{n=1}^{N} \bar{y}_n^2 \right) \Big/ \left( \sum_{n=1}^{N} \bar{y}_n^4 \right)$$

Observe that with this choice of $\rho$, we have $\sum_n y_n^2 = \sum_n y_n^4$.

The "reduced" cost function (i.e., that obtained upon using the optimal value of $\rho$) becomes

$$\bar{J}(\bar{\mathbf{g}}) \triangleq J(\rho \bar{\mathbf{g}}) \Big|_{\rho = \rho_{\text{opt}}} = N - \frac{\left( \sum_{n=1}^{N} \bar{y}_n^2 \right)^2}{\sum_{n=1}^{N} \bar{y}_n^4}$$

Minimizing $\bar{J}(\bar{\mathbf{g}})$ is therefore accomplished by choosing the angular orientation of $\bar{\mathbf{g}}$ to minimize

$$F(\mathbf{g}) \triangleq \frac{\sum_{n=1}^{N} \bar{y}_n^4}{\left( \sum_{n=1}^{N} \bar{y}_n^2 \right)^2} = \frac{\sum_{n=1}^{N} y_n^4}{\left( \sum_{n=1}^{N} y_n^2 \right)^2} = \frac{m_4}{m_2^2}$$

involving the ratio of the fourth-order moment $m_4$ to the square of the second-order moment $m_2$. Note that this final form is radially invariant: replacing $\bar{y}_n$ by $\beta \bar{y}_n$, with $\beta$ any constant, does not change the value of the ratio; the choice $\beta = \rho$ leads to the indicated equality. We then claim:

**Theorem 1** *For all nonzero equalizer settings* $\mathbf{g}$, *the reduced cost function is bounded as*

$$1 \geq F(\mathbf{g}) \geq \frac{1}{N}.$$

*The lower bound* $1/N$ *is attained if and only if* $\{y_n\}$ *is a constant modulus sequence.*

A proof may be deduced from the identity

$$F = \frac{\sum_{n=1}^{N} y_n^4}{\left( \sum_{n=1}^{N} y_n^2 \right)^2} = \frac{1}{N} \left( 1 + \frac{\sum_{1 \leq i < j \leq N} (y_i^2 - y_j^2)^2}{\left( \sum_{n=1}^{N} y_n^2 \right)^2} \right)$$

## 3. ALGORITHM DEVELOPMENT

Consider writing successive equalizer outputs in the form

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{U}_1^T \\ \mathbf{U}_2^T \\ \vdots \\ \mathbf{U}_N^T \end{bmatrix}}_{\mathcal{U}} \mathbf{g}^T = \mathbf{Q} \mathbf{R} \mathbf{g}^T = \mathbf{Q} \mathbf{w}$$

where we invoke the QR-decomposition of $\mathcal{U}$ to obtain a $\mathbf{Q}$ matrix having orthonormal columns, and where $\mathbf{R} \mathbf{g}^T = \mathbf{w}$. By partitioning $\mathbf{Q}$ row-wise as

$$\mathbf{Q} = \begin{bmatrix} \mathbf{q}_1^T \\ \mathbf{q}_2^T \\ \vdots \\ \mathbf{q}_N^T \end{bmatrix} \quad \Rightarrow \quad y_n = \mathbf{q}_n^T \mathbf{w},$$

we may calculate the gradient of $F$ with respect to $\mathbf{w}$ as

$$\begin{aligned} \frac{\partial}{\partial \mathbf{w}} \frac{m_4}{m_2^2} &= \frac{\partial}{\partial \mathbf{w}} \frac{\sum_n y_n^4}{\left( \sum_n y_n^2 \right)^2} \\ &= 4 \, \frac{\left( \sum_n y_n^2 \right)^2 \sum_n \mathbf{q}_n y_n^3 - \left( \sum_n y_n^4 \right) \sum_n \mathbf{q}_n y_n}{\left( \sum_n y_n^2 \right)^4} \end{aligned}$$

Since our criterion is radially invariant, we may scale $\mathbf{w}$ to unit norm ($\|\mathbf{w}\| = 1$), which, by orthonormality of the columns of $\mathbf{Q}$, implies that $m_2 = \sum_n y_n^2 = 1$. In this case, the previous gradient calculation simplifies to

$$\frac{\partial}{\partial \mathbf{w}} \frac{m_4}{m_2^2} \Big|_{\|\mathbf{w}\| = 1} = 4 \left( \sum_n \mathbf{q}_n y_n^3 - \left( \sum_n y_n^4 \right) \sum_n \mathbf{q}_n y_n \right)$$

Note that with $m_2 = \sum_n y_n^2 = 1$, we have $F = m_4 = \sum_n y_n^4$, as well as

$$\sum_n \mathbf{q}_n y_n = \underbrace{\left( \sum_n \mathbf{q}_n \mathbf{q}_n^T \right)}_{= \mathbf{Q}^T \mathbf{Q} = \mathbf{I}} \mathbf{w} = \mathbf{w}$$

This gives the following algorithm:

1. Given the data matrix $\mathcal{U}$, run a QR decomposition, and retain the **Q** factor, whose columns provide an orthonormal basis for the column space of $\mathcal{U}$.

2. Using an initial vector $\mathbf{w}_0$, with $\|\mathbf{w}_0\| = 1$, run the gradient descent procedure ($i$ = iteration index)

$$\begin{aligned}\mathbf{v}_{i+1} &= \mathbf{w}_i - \frac{\mu_i}{4}\left.\frac{\partial F(\mathbf{w})}{\partial \mathbf{w}}\right|_{\mathbf{w}=\mathbf{w}_i} \\ &= [1+\mu_i F_i]\,\mathbf{w}_i - \mu_i \sum_{n=1}^{N} y_n^3(i)\,\mathbf{q}_n \\ \mathbf{w}_{i+1} &= \mathbf{v}_{i+1}/\|\mathbf{v}_{i+1}\|\end{aligned}$$

where $y_n(i) = \mathbf{q}_n^T \mathbf{w}_i$.

If we let $\mathbf{y}^{\odot 3}$ denote the Hadamard cube of the vector $\mathbf{y}$ (i.e., the vector whose components are $y_n^3$), the above algorithm can be rewritten as

$$\begin{aligned}\mathbf{v}_{i+1} &= \mathbf{w}_i - \mu_i\left(\mathbf{Q}^T\mathbf{y}^{\odot 3}(i) - F_i\,\mathbf{w}_i\right) \\ \mathbf{w}_{i+1} &= \mathbf{v}_{i+1}/\|\mathbf{v}_{i+1}\| \end{aligned} \qquad (1)$$

## 4. CONVERGENCE OF ITERATIVE ALGORITHM

We obtain in this section a bound on the step-size $\mu_i$ which ensures monotonic convergence of the algorithm.

Relax the unit-norm constraint on $\mathbf{w}$ and let $\mathcal{B}$ denote the unit ball in $\mathbb{R}^{(M+1)P}$:

$$\mathcal{B} = \{\mathbf{w} \in \mathbb{R}^{(M+1)P} : \|\mathbf{w}\| \leq 1\}.$$

This is a convex subset of $\mathbb{R}^{(M+1)P}$.

The Hessian matrix of the numerator $m_4$ of $F$ is readily calculated as

$$\frac{\partial^2 m_4}{\partial \mathbf{w}\,\partial \mathbf{w}^T} = 12\,\mathbf{Q}^T\,\mathrm{diag}(\mathbf{y}^{\odot 2})\,\mathbf{Q},$$

where $\mathrm{diag}(\mathbf{y}^{\odot 2})$ denotes the diagonal matrix

$$\mathrm{diag}(\mathbf{y}^{\odot 2}) = \begin{bmatrix} y_1^2 & 0 & \cdots & 0 \\ 0 & y_2^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & y_N^2 \end{bmatrix}.$$

We let $\gamma$ denote one-fourth the largest eigenvalue $\lambda_+$ of this Hessian matrix over $\mathcal{B}$:

$$\gamma \triangleq \frac{1}{4}\max_{\|\mathbf{w}\|\leq 1} \lambda_+\left(\frac{\partial^2 m_4}{\partial \mathbf{w}\,\partial \mathbf{w}^T}\right). \qquad (2)$$

Letting $F_i = F(\mathbf{w}_i)$, we then have:

**Theorem 2** *If $\mathbf{w}_i$ is not a stationary point of the procedure (1), then $F_{i+1} < F_i$ if $\mu_i$ obeys the bound*

$$0 < \mu_i < \frac{2(\gamma - F_i)}{F_i^2 + (\gamma - F_i)^2 - \|\mathbf{Q}^T\mathbf{y}^{\odot 3}(i)\|^2}.$$

*Proof:* Consider the functional

$$f(\mathbf{w}) = 2\gamma(\mathbf{w}^T\mathbf{w} - 1) - m_4(\mathbf{w}), \qquad \|\mathbf{w}\| \leq 1,$$

which on the unit sphere $\|\mathbf{w}\| = 1$ fulfills $f(\mathbf{w}) = -F(\mathbf{w})$. Minimizing $F(\mathbf{w})$ on the unit sphere $\|\mathbf{w}\| = 1$ is thus equivalent to maximizing $f(\mathbf{w})$ there.

Now, in view of our choice of $\gamma$ [cf. (2)], the Hessian matrix of $f(\mathbf{w})$ is positive semi-definite over $\mathcal{B}$:

$$\frac{\partial^2 f(\mathbf{w})}{\partial \mathbf{w}\,\partial \mathbf{w}^T} = 4\gamma\mathbf{I} - \frac{\partial^2 m_4(\mathbf{w})}{\partial \mathbf{w}\,\partial \mathbf{w}^T} \geq 0, \qquad \text{for all } \|\mathbf{w}\| \leq 1.$$

This is necessary and sufficient [6] for $f(\mathbf{w})$ to be a convex function of $\mathbf{w}$ over $\mathcal{B}$. The gradient inequality applicable to convex functions [6] then implies that, for all vectors $\mathbf{w}_i$ and $\mathbf{w}_{i+1}$ in $\mathcal{B}$,

$$f(\mathbf{w}_{i+1}) - f(\mathbf{w}_i) \geq \left(\nabla f(\mathbf{w}_i)\right)^T\left(\mathbf{w}_{i+1} - \mathbf{w}_i\right) \qquad (3)$$

where $\nabla f(\mathbf{w}) = 4(\gamma\,\mathbf{w} - \mathbf{Q}^T\mathbf{y}^{\odot 3})$ is the gradient vector of $f(\mathbf{w})$. Positivity of the right-hand side will imply $f(\mathbf{w}_{i+1}) > f(\mathbf{w}_i)$, or $F_{i+1} < F_i$. It suffices thus to choose $\mu_i$ such that the right-hand side is positive. This constraint takes the form

$$\Big(\underbrace{\gamma\,\mathbf{w}_i - \mathbf{Q}^T\mathbf{y}^{\odot 3}}_{\nabla f(\mathbf{w}_i)}\Big)^T\left(\underbrace{\frac{\mathbf{w}_i - \mu_i(\mathbf{Q}^T\mathbf{y}^{\odot 3}(i) - F_i\mathbf{w}_i)}{\|\mathbf{w}_i - \mu_i(\mathbf{Q}^T\mathbf{y}^{\odot 3}(i) - F_i\mathbf{w}_i)\|}}_{\mathbf{w}_{i+1}} - \mathbf{w}_i\right) > 0.$$

Solving for $\mu_i$ compatible with this constraint gives the bound of the theorem statement. ◇

*Remark:* An "optimal" choice for $\mu_i$ would maximize the right-hand side of (3), which in turn amounts to maximizing $\left(\nabla f(\mathbf{w}_i)\right)^T\mathbf{w}_{i+1}$. By the Cauchy-Schwarz inequality,

$$\left(\nabla f(\mathbf{w}_i)\right)^T\mathbf{w}_{i+1} \leq \|\nabla f(\mathbf{w}_i)\| \cdot \underbrace{\|\mathbf{w}_{i+1}\|}_{=1}$$

with equality iff $\mathbf{w}_{i+1}$ is colinear with $\nabla f(\mathbf{w}_i)$. This colinearity is satisfied with

$$\mu_i^{\mathrm{opt}} = \frac{1}{\gamma - F_i} \qquad (4)$$

which is thus an "optimal" step size choice. ◇

The dependence on $\gamma$ from (2) complicates the direct application of these results. One may show, similar to [7], the chain of inequalities

$$F_i \leq \|\mathbf{Q}^T\mathbf{y}^{\odot 3}(i)\| \leq \gamma/3 = \max_{\|\mathbf{w}\|=1} F(\mathbf{w}),$$

which can assist the selection of $\mu_i$ from (4). In simulations, the stepsize choice

$$\mu_i = \frac{\alpha}{F_i}$$

with $\frac{1}{2} \le \alpha \le \frac{2}{3}$, has been found to work very satisfactorily.

## 5. SIMULATION EXAMPLE

Here we consider a single-input–two-output channel whose impulse response terms are

$$\begin{bmatrix} \mathbf{H}_0^T \\ \mathbf{H}_1^T \\ \mathbf{H}_2^T \\ \mathbf{H}_3^T \\ \mathbf{H}_4^T \\ \mathbf{H}_5^T \\ \mathbf{H}_6^T \end{bmatrix} = \begin{bmatrix} -0.335923 & 0.087186 \\ 0.228832 & -0.011060 \\ 0.465334 & -0.474379 \\ 0.623545 & -0.273567 \\ 0.387568 & -0.535722 \\ 0.267495 & -0.081896 \\ 0.087663 & 0.631425 \end{bmatrix}$$

The input is a real binary sequence, and white Gaussian noise is added to the channel output to obtain a signal to noise ratio of 14dB. Figure 1 shows the evolution of $F_i$ using fourteen coefficients per equalizer branch, a step size of $\mu_i = 2/(3F_i)$, and a block length of $N = 262$. Center tap initialization is employed, and convergence is seen to occur in a few iterations. Figure 2 shows the resulting output sequence $\{y_k\}$, which is nearly binary except for leading and trailing terms which are nearly zero, induced by an increased sequence length from the channel-equalizer convolutional cascade.

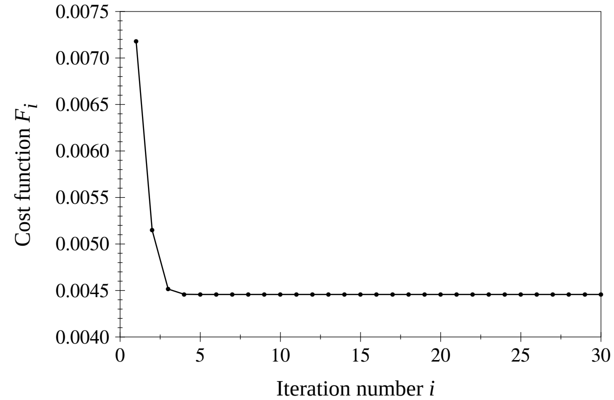


**Figure 1:** Evolution of cost function $F_i$.

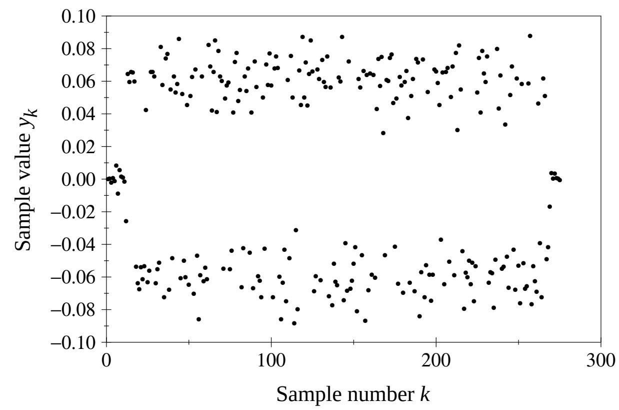


**Figure 2:** Output sequence $\{y_k\}$ after convergence.

## 6. CONCLUDING REMARKS

The iterative algorithm developed here is well suited to signal restoration over finite intervals. The computational complexity is dominated by the initial QR decomposition, with subsequent power iterations contributing but a modest computational load. By contrast, the ACMA [4] (itself an iterative algorithm) requires a singular value decomposition, QZ iterations, matrix inversions, and so forth, presenting a significantly higher computational burden.

For simplicity, real signals and channels have been considered, as the extension to complex signals and channels is straightforward.

## 7. REFERENCES

[1] D. N. Godard, “Self-recovering equalization and carrier tracking in a two-dimensional data communication system” *IEEE Trans. Communications*, vol. 28, pp. 1867–1875, November 1980.

[2] 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. 459–471, April 1983.

[3] C. R. Johnson *et al.*, “Blind equalization using the constant modulus criterion: A review,” *Proc. IEEE*, vol. 86, pp. 1927–1950, October 1998.

[4] A.-J. van der Veen and A. Paulraj, “An analytic constant modulus algorithm,” *IEEE Trans. Signal Processing*. vol. 44, pp. 1136–1155, May 1996.

[5] A.-J. van der Veen, “Asymptotic properties of the algebraic constant modulus algorithm,” *IEEE Trans. Signal Processing*, vol. 49, pp. 1796–1807, August 2001.

[6] T. R. Rockafellar, *Convex Analysis,* Princeton Univ. Press, 1970.

[7] M. Mboup and P. A. Regalia, “A gradient search interpretation of the super-exponential algorithm,” *IEEE Trans. Information Theory*, vol. 46, pp. 2731–2734, November 2000.