This item was submitted to [Loughborough's Research Repository](#) by the author.
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

Fast convergence algorithms for joint blind equalization and source separation based upon the cross-correlation and constant modulus criterion

PLEASE CITE THE PUBLISHED VERSION

PUBLISHER

© IEEE

VERSION

VoR (Version of Record)

LICENCE

CC BY-NC-ND 4.0

REPOSITORY RECORD

Luo, Yuhui, and Jonathon Chambers. 2019. "Fast Convergence Algorithms for Joint Blind Equalization and Source Separation Based Upon the Cross-correlation and Constant Modulus Criterion". figshare.
<https://hdl.handle.net/2134/5780>.

This item was submitted to Loughborough's Institutional Repository (<https://dspace.lboro.ac.uk/>) by the author and is made available under the following Creative Commons Licence conditions.

For the full text of this licence, please go to:
<http://creativecommons.org/licenses/by-nc-nd/2.5/>

FAST CONVERGENCE ALGORITHMS FOR JOINT BLIND EQUALIZATION AND SOURCE SEPARATION BASED UPON THE CROSS-CORRELATION AND CONSTANT MODULUS CRITERION

Y. Luo and J.A. Chambers

Centre for Digital Signal Processing Research,
School of Physical Sciences and Engineering, King's College London
Strand, London WC2R 2LS, United Kingdom
email: jonathon.chambers@kcl.ac.uk

ABSTRACT

To solve the problem of joint blind equalization and source separation, two new quasi-Newton adaptive algorithms with rapid convergence property are proposed based upon the cross-correlation and constant modulus (CC-CM) criterion, namely the block-Shanno cross-correlation and constant modulus algorithm (BS-CCCMA) and the fast quasi-Newton cross-correlation and constant modulus algorithm (FQN-CCCMA). Simulations studies are used to show that the convergence properties of these algorithms are much improved upon those of the conventional LMS-CCCMA algorithm.

1. INTRODUCTION

We address the problem of joint blind equalization and source separation in a multi-input and multi-output (MIMO) convolutive system. Such a system is representative of spatial division multiple access (SDMA) in wireless communications, where a number U of independent and identically distributed (i.i.d.) source signals is transmitted through a linear channel, which is assumed of order L , and picked up by an array of R antennae to exploit spatial diversity. To retrieve the original transmitted signals, the corrupted received signal is processed by a bank of U parallel space-time equalizers and each of the space-time equalizers consists of R sub-equalizers attached to different antennae. The order of the sub-equalizers is assumed to be N . The composite source signal at time instant k is written as $\mathbf{s}(k) = [\mathbf{s}_1^T(k) \dots \mathbf{s}_U^T(k)]^T$, where $\mathbf{s}_i(k) = [s_i(k) \dots s_i(k - N - L)]^T$ is the i^{th} source vector. By denoting Δ^T as the channel convolution matrix, the space-time equalizer regressor $\mathbf{x}(k) = [\mathbf{x}_1^T(k) \dots \mathbf{x}_R^T(k)]^T$ is written as

$$\mathbf{x}(k) = \Delta^T \mathbf{s}(k) + \mathbf{n}(k) \quad (1)$$

where $\mathbf{x}_j(k) = [x_j(k) \dots x_j(k - N)]^T$ is the j^{th} antenna output and $\mathbf{n}(k)$ represents the additive noise. Notations

$(\cdot)^H$, $(\cdot)^T$ and $(\cdot)^*$ denote Hermitian, transpose and complex conjugate respectively. In determining the equalizer coefficients, blind adaptive algorithms are usually preferred since better channel utilization efficiency is achieved with the elimination of the training sequence. For the CC-CM criterion, which is also known as the multiuser-CMA criterion, the underlying constant modulus property of the communication signals is exploited whilst repeated retrieval of the same source is prevented by penalizing the cross-correlation between multiple output signals [4] and [2]. Proof of its global convergence is shown in [5]. For the l^{th} space-time equalizer, the cost function is written as

$$J_l(k) = E\{(|y_l(k)|^2 - R_2)^2\} + \gamma \sum_{m=1}^{l-1} \sum_{\delta=-(L+N)}^{L+N} |E\{\rho_{l,m}(k, \delta)\}|^2 \quad (2)$$

where $E\{(|y_l(k)|^2 - R_2)^2\}$ is the constant modulus cost, $E\{\rho_{l,m}(k, \delta)\} = E\{y_l(k)y_m^*(k-\delta)\}$ is the cross-correlation penalty between equalizer- l and equalizer- m , R_2 is the so-called dispersion constant and $\gamma \in R^+$ is the mixing parameter. This cost function is usually minimized by a steepest descent type method (i.e., the LMS-CCCMA algorithm), in which the update direction $\mathbf{p}_l(k)$ is chosen as the negative gradient with respect to the equalizer tap weight vector. Denoting μ as the step size and $\mathbf{g}_l(k)$ as the instantaneous gradient estimator, the update equation for the l^{th} equalizer of the LMS-CCCMA algorithm is given by

$$\mathbf{w}_l(k+1) = \mathbf{w}_l(k) - \mu \mathbf{g}_l(k) \quad (3)$$

where

$$\mathbf{g}_l(k) = \left[\begin{array}{c} 2(|y_l(k)|^2 - R_2)y_l^*(k) + \\ \gamma \sum_m \sum_{\delta} \hat{E}\{\rho_{l,m}^*(k, \delta)\} y_m^*(k - \delta) \end{array} \right] \mathbf{x}(k) \quad (4)$$

and $\hat{E}\{\rho_{l,m}(k, \delta)\}$ is the estimator of the cross-correlation penalty. With $\lambda \in (0, 1)$ controlling the length of the data

window in the estimation, it can be recursively updated by $\hat{E}\{\rho_{l,m}(k, \delta)\} = \lambda \hat{E}\{\rho_{l,m}(k-1, \delta)\} + (1-\lambda)y_l(k)y_m^*(k-\delta)$. However, the LMS-CCCMA algorithm is observed to have slow convergence rate and therefore in this work, second order Newton methods are proposed to overcome this shortcoming, namely the new BS-CCCMA and the FQN-CCCMA algorithms. As the Hessian information is utilized to obtain better approximations of the error performance surface, faster convergence rate over the LMS-CCCMA algorithm is achieved.

2. THE ALGORITHMS

Denote $H_l(k)$ as the Hessian matrix of the cost function. The update direction of the Newton method is given by $\mathbf{p}_l(k) = -H_l^{-1}(k)\mathbf{g}_l(k)$ and the l^{th} equalizer update equation becomes

$$\mathbf{w}_l(k+1) = \mathbf{w}_l(k) - \mu H^{-1}(k)\mathbf{g}_l(k) \quad (5)$$

As the exact Hessian matrix is difficult to compute, in the development of the following BS-CCCMA and the FQN-CCCMA algorithms two different methods to approximate the inverse Hessian matrix are employed.

2.1. The BS-CCCMA algorithm

In the BS-CCCMA algorithm, the inverse Hessian matrix is approximated by the Shanno method, which is also known as the one-step BFGS algorithm [1]. As the adaptation is performed in a block-by-block fashion, the received data sequence is rearranged such that the k^{th} data block is given by $[\mathbf{x}((k-1)M) \mathbf{x}((k-1)M+1) \dots \mathbf{x}(kM-1)]$, where the block size is selected as $M \gg N+L$ for accurate estimation of the cross-correlation. The extended block CC-CM cost function for the k^{th} data block is defined as

$$J_l(k) = \frac{1}{M} \sum_{i=0}^{M-1} \left[|y_l((k-1)M+i)|^2 - R_2 \right]^2 + \frac{\gamma}{M^2} \sum_m \sum_{\delta} \left| \sum_{i=0}^{M-1} [\rho_{l,m}((k-1)M+i, \delta)] \right|^2 \quad (6)$$

For notational simplicity, we write $\mathbf{x}((k-1)M+i) \equiv \mathbf{x}(i)$. Thus, if not mentioned, the k^{th} block is being processed. Since the published Shanno algorithm literature requires real input data, an equivalent form of eqn (6) is derived to accommodate complex signals

$$J_l(k) = \frac{1}{M} \sum_{i=0}^{M-1} (\mathbf{w}_l^T X(i) \mathbf{w}_l' - 1)^2 + \frac{\gamma}{M^2} \sum_m \sum_{\delta} \left(\frac{(\mathbf{w}_l^T \sum_i X'(i, \delta) \mathbf{w}_m')^2 + (\mathbf{w}_l^T \sum_i X''(i, \delta) \mathbf{w}_m')^2}{(\mathbf{w}_l^T \sum_i X''(i, \delta) \mathbf{w}_m')^2} \right) \quad (7)$$

where $\mathbf{w}_l' = [\text{Re}(\mathbf{w}_l^T) \text{Im}(\mathbf{w}_l^T)]^T$ represents the l^{th} space-time equalizer tap weight vector in the real field. The matrices $X(i)$, $X'(i, \delta)$ and $X''(i, \delta)$ are defined as followed.

$$X(i) = \mathbf{x}_f(i) \mathbf{x}_f^T(i) + \mathbf{x}_b(i) \mathbf{x}_b^T(i) \\ X'(i, \delta) = \begin{bmatrix} A(i, \delta) & B(i, \delta) \\ -B(i, \delta) & A(i, \delta) \end{bmatrix} \\ X''(i, \delta) = \begin{bmatrix} B(i, \delta) & -A(i, \delta) \\ A(i, \delta) & B(i, \delta) \end{bmatrix}$$

$$\text{where } \mathbf{x}_f(i) = \begin{bmatrix} \text{Re}(\mathbf{x}(i)) \\ -\text{Im}(\mathbf{x}(i)) \end{bmatrix}, \mathbf{x}_b(i) = \begin{bmatrix} \text{Im}(\mathbf{x}(i)) \\ \text{Re}(\mathbf{x}(i)) \end{bmatrix} \\ A(i, \delta) = \text{Re}(\mathbf{x}(i)) \text{Re}(\mathbf{x}^T(i-\delta)) + \text{Im}(\mathbf{x}(i)) \text{Im}(\mathbf{x}^T(i-\delta)) \\ B(i, \delta) = \text{Im}(\mathbf{x}(i)) \text{Re}(\mathbf{x}^T(i-\delta)) - \text{Re}(\mathbf{x}(i)) \text{Im}(\mathbf{x}^T(i-\delta))$$

This block cost function is iteratively minimized by updating the real equalizer tap weight vector $\mathbf{w}_l'$. That is, for the k^{th} block, the update equation of the BS-CCCMA algorithm at the n^{th} iteration is given by $\mathbf{w}_l'^{n+1} = \mathbf{w}_l'^n + \mu^n \mathbf{p}_l^n$, where $\mathbf{p}_l^n = -(H_l^n)^{-1} \mathbf{g}_l^n$ is the update direction vector and $\mathbf{g}_l^n$ denotes the block gradient of the cost function with respect to the real equalizer vector $\mathbf{w}_l'^n$.

$$\mathbf{g}_l^n = \frac{4}{M} \sum_{i=0}^{M-1} (\mathbf{w}_l'^T X(i) \mathbf{w}_l' - 1) X(i) \mathbf{w}_l' + \quad (8)$$

$$\frac{2\gamma}{M^2} \sum_m \sum_{\delta} \left(\frac{\left(\mathbf{w}_l'^T \sum_i X^0(i, \delta) \mathbf{w}_m^0 \right) \sum_i X^0(i, \delta) \mathbf{w}_m^0 + \left(\mathbf{w}_l'^T \sum_i X^0(i, \delta) \mathbf{w}_m^0 \right) \sum_i X^0(i, \delta) \mathbf{w}_m^0}{\left(\mathbf{w}_l'^T \sum_i X^0(i, \delta) \mathbf{w}_m^0 \right) \sum_i X^0(i, \delta) \mathbf{w}_m^0} \right)$$

This gradient information is used in the approximation of $(H_l^n)^{-1}$. Denoting $\mathbf{u}_l^n = \mathbf{g}_l^n - \mathbf{g}_l^{n-1}$ as the difference gradient vector between the two most recent iterations, the Shanno method approximates the inverse Hessian matrix by

$$(H_l^n)^{-1} = I - \frac{\mathbf{u}_l^n (\mathbf{p}_l^{n-1})^T + \mathbf{p}_l^{n-1} (\mathbf{u}_l^n)^T + a \mathbf{p}_l^{n-1} (\mathbf{p}_l^{n-1})^T}{(\mathbf{p}_l^{n-1})^T \mathbf{u}_l^n} \quad (9)$$

where $a = -\mu^{n-1} + \frac{|\mathbf{u}_l^n|^2}{\mathbf{p}_l^{n-1} \mathbf{u}_l^n}$ and I is the identity matrix.

In fact, the computation of $(H_l^n)^{-1}$ can be avoided as the direct computation of the descent direction $\mathbf{p}_l^n$ is possible, i.e.,

$$\mathbf{p}_l^n = -\mathbf{g}_l^n + b \mathbf{u}_l^n + (c - ab) \mathbf{p}_l^{n-1} \quad (10)$$

where the scalars b and c are given by $b = \frac{(\mathbf{p}_l^{n-1})^T \mathbf{g}_l^n}{(\mathbf{p}_l^{n-1})^T \mathbf{u}_l^n}$ and $c = \frac{(\mathbf{u}_l^n)^T \mathbf{g}_l^n}{(\mathbf{p}_l^{n-1})^T \mathbf{u}_l^n}$. In the step size selection, the parameter μ^n should be chosen such that the following conditions

$$J_l^n - J_l^{n+1} \geq -\alpha \mu^n (\mathbf{g}_l^n)^T \mathbf{p}_l^n \quad (11)$$

$$(\mathbf{g}_l^{n+1})^T \mathbf{p}_l^n \geq \beta (\mathbf{g}_l^n)^T \mathbf{p}_l^n \quad (12)$$

are satisfied, where $0 < \alpha < \beta < 1$. Furthermore, to avoid instability, a multiple checking procedure regarding the magnitude of the equalizer tap weights and the validity of the searching direction (i.e., whether the search direction is orthogonal to the gradient vector) is suggested [1]. For each data block, such adaptation continues until either a sufficiently small value of $\|\mathbf{g}_t^n\|_2^2$ is reached or a pre-determined number d of iterations has been performed. Then adaptation on the next data block can start. Here the counter d can be chosen as the dimension of the equalizer regressor since experience shows that a fairly good approximation to the inverse of the Hessian can be constructed after this number of iterations. In term of the BS-CCCMA convergence property, since the Shanno method ensures that every step of the algorithm is always a descent direction by retaining the positive definite Hessian matrix, ill-convergence is avoided. Also notice that the BS-CCCMA algorithm is closely related to the block conjugate gradient type algorithm in the sense that by setting $b = 0$, a block conjugate gradient cross-correlation and constant modulus algorithm is obtained.

2.2. FQN-CCCMA algorithm

Compared with the BS-CCCMA algorithm, the FQN-CCCMA algorithm uses the cost function in eqn (2) and approximates the inverse Hessian matrix with repeated applications of the matrix inversion lemma. To avoid the statistical expectation involved in the exact expression of the Hessian matrix, the following approximation is used

$$H_l(k) = 2 \frac{1-\lambda}{1-\lambda^k} \Psi(k) + \gamma \left(\frac{1-\lambda}{1-\lambda^k} \right)^2 \Phi(k) \quad (13)$$

where $\Psi(k)$ and $\Phi(k)$ are Hermitian matrices given by

$$\Psi(k) = \sum_{i=1}^k \lambda^{k-i} \left(2 |y_l(i)|^2 - R_2 \right) \mathbf{x}(i) \mathbf{x}^H(i) \quad (14)$$

$$\Phi(k) = \sum_{m=1}^{l-1} \sum_{\delta=-(N+L)}^{N+L} \mathbf{v}_m(k, \delta) \mathbf{v}_m^H(k, \delta) \quad (15)$$

The vector $\mathbf{v}_m(k, \delta)$ in eqn (15) is written as $\mathbf{v}_m(k, \delta) = \sum_{i=1}^k \lambda^{k-i} y_m^*(i - \delta) \mathbf{x}(i)$. In order to reduce the computational complexity of the FQN-CCCMA algorithm, we utilize only the diagonal terms of $\Phi(k)$. Representing the square root of the diagonal of $\Phi(k)$ as $\mathbf{b}(k)$. The matrix $\Phi(k)$ is approximated by $\Phi(k) \approx \left(\mathbf{b}(k) \mathbf{b}^H(k) \right) * I$, where $*$ is the operation of point-by-point multiplication. The effect of this approximation is slightly slower convergence speed at the very beginning of the adaptation compared with BS-CCCMA. But better curvature information of the CC-CM cost function can still be built up within a number of iterations than with the conventional LMS-CCCMA algorithm,

as will be confirmed by the simulations results. Notice that there exists two recursive relationship

$$\begin{aligned} \Psi(k) &= \lambda \Psi(k-1) + (2 |y_l(k)|^2 - R_2) \mathbf{x}(k) \mathbf{x}^H(k) \\ \mathbf{v}_m(k) &= \lambda \mathbf{v}_m(k-1) + y_m^*(k - \delta) \mathbf{x}(k) \end{aligned}$$

By applying the matrix inversion lemma [6] to eqn (13), the inverse Hessian matrix is given by

$$H_l^{-1}(k) = \frac{1-\lambda^k}{2(1-\lambda)} \Psi^{-1}(k) - \frac{\frac{1-\lambda^k}{2(1-\lambda)} \Psi^{-1}(k) \text{diag}(\mathbf{b}(k) \mathbf{b}^H(k)) \frac{1-\lambda^k}{2(1-\lambda)} \Psi^{-1}(k)}{\gamma^{-1} \left(\frac{1-\lambda^k}{1-\lambda} \right)^2 + \mathbf{b}^H(k) \frac{1-\lambda^k}{2(1-\lambda)} \Psi^{-1}(k) \mathbf{b}(k)} \quad (16)$$

where the computation of the matrix $\Psi^{-1}(k)$ is solved by applying the matrix inversion lemma again, i.e.,

$$\Psi^{-1}(k) = \frac{1}{\lambda} \Psi^{-1}(k-1) - \frac{\lambda^{-2} \Psi^{-1}(k-1) \mathbf{x}(k) \mathbf{x}^H(k) \Psi^{-1}(k-1)}{(2 |y_l(k)|^2 - R_2)^{-1} + \lambda^{-1} \mathbf{x}^H(k) \Psi^{-1}(k-1) \mathbf{x}(k)} \quad (17)$$

In the initialization, $\Psi^{-1}(0) = \varsigma I$, where ς is a small constant with typical value of 0.01. It is feasible for the matrix in eqn (17) to lose its positive definite nature and monitoring may therefore be necessary in practice. However, no problem was observed in the simulation studies, as in [6]. Thus with the approximation of $H_l^{-1}(k)$ in eqn (16) and the gradient in eqn (4), the FQN-CCCMA update equation for the l^{th} equalizer is shown in eqn (5). Moreover, the step size parameter μ can be adaptively chosen as $\mu = \frac{\alpha}{\epsilon + \mathbf{x}^H(k) H_l^{-1}(k) \mathbf{x}(k)}$ to reduce the fluctuation in the estimation of the inverse Hessian matrix, where the parameters ϵ and α are small constants within the range of (0, 1]. In term of convergence, by inspection of eqn (13), provided the positive definite property of the approximated Hessian matrix is retained, the update direction is ensured to be a descent direction in the adaptation. For the computational complexity, similar to the BS-CCCMA algorithm, the price for fast convergence is in the increase of computational complexity, which is $O(d^2)$. But if the inputs of the equalizer are statistically stationary or slowly varying compared with the convergence of the filter tap weights, the update of the inverse Hessian matrices can be performed every d samples once steady state is reached.

3. SIMULATIONS

A three users and four antennae QPSK system with source alphabet $\left\{ \pm \frac{1}{\sqrt{2}} \pm \frac{1}{\sqrt{2}} j \right\}$ is assumed. The order of the channel is $L = 1$ and the order of the sub-equalizer is $N = 2$. Hence the length of the space-time equalizers is $R(N +$

Fig. 1. Algorithms comparison in term of residual interference (a) EQ-1 (b) EQ-2 (c) EQ-3

1) = 12. The channel coefficients are generated as random numbers with full rank convolution matrix. Additive white Gaussian noise of $SNR = 30dB$ is assumed. The performance of the LMS-CCCMA, BS-CCCMA and FQN-CCCMA algorithms are compared in term of residual interference at the three equalizer outputs. Denoting the vector $h_l(k)$ as the combined channel + equalizer- l impulse response, residual interference at the equalizer- l is measured as $\frac{\|h_l(k)\|^2 - \max(\|h_l(k)\|)^2}{\max(\|h_l(k)\|)^2}$. In Figure 1, the fast convergence property of the FQN-CCCMA and BS-CCCMA algorithms over the conventional LMS-CCCMA algorithm is confirmed. For equalizer-1, only 200 data samples are required for the open eye constellation diagram to be achieved, compared with 800 and 1500 samples required for BS-CCCMA and LMS-CCCMA respectively. For equalizer-2, 450 samples are required by FQN-CCCMA, compared with approximately 1000 and 2000 samples by the BS-CCCMA and the LMS-CCCMA algorithms respectively. For equalizer-3, 800 data samples are required for FQN-CCCMA while 2000 samples for BS-CCCMA and at least 5000 samples for LMS-CCCMA. This is depicted in Figure 2. Successful separation of different sources is supported by the plots of the combined channel + equalizer impulse response, which can be found in [3].

4. CONCLUSION

Based upon the CC-CM criterion for solving the problem of joint blind equalization and source separation, the BS-CCCMA and the FQN-CCCMA algorithms are proposed to overcome the slow convergence of the conventional LMS-CCCMA algorithm by utilizing the Hessian information. The proposed methods belong to the quasi-Newton family and their rapid convergence is confirmed by simulations.

Fig. 2. Algorithms comparison (a) FQN-CCCMA, 200 samples (b) EQ-1: BS-CCCMA, 200 samples (c) EQ-1: LMS-CCCMA, 200 samples (d) EQ-2: FQN-CCCMA, 450 samples (e) EQ-2: BS-CCCMA, 450 samples (f) EQ-2: LMS-CCCMA, 450 samples (g) EQ-3: FQN-CCCMA, 800 samples (h) EQ-3: BS-CCCMA, 800 samples (i) EQ-3: LMS-CCCMA, 800 samples

5. REFERENCES

- [1] U.G. Jani, E.M. Dowling, R.M. Golden, and Z. Wang. Multiuser interference suppression using block Shanno constant modulus algorithm. *IEEE Trans. on Signal Processing*, 48(5):1503–1506, 2000.
- [2] S. Lambotaran and J.A. Chambers. On the surface characteristics of a mixed constant modulus and cross-correlation criterion for the blind equalization of an MIMO channel. *Signal Processing*, 74:209–216, 1998.
- [3] Y. Luo. *A Mixed Cross-Correlation and Constant Modulus Adaptive Algorithm for Joint Blind Equalization and Source Separation*. PhD thesis, Imperial College of Science, Technology and Medicine, University of London, UK, 2001.
- [4] C.B. Papadias and A.J. Paulraj. A space-time constant modulus algorithm for SDMA systems. In *Proc. of IEEE Vehicular Technology Conf.*, pages 86–90, Atlanta, 1996.
- [5] A. Touzni, I. Fijalkow, M.G. Larimore, and J.R. Treichler. A globally convergent approach for blind MIMO adaptive deconvolution. *IEEE Trans. on Signal Processing*, 49(6):1166–1178, 2001.
- [6] G. Yan and H. Fan. A Newton-like algorithm for complex variables with applications in blind equalization. *IEEE Trans. on Signal Processing*, 48:553–556, 2000.