NORMALIZED SLIDING WINDOW CONSTANT MODULUS ALGORITHMS FOR BLIND EQUALIZATION

Constantinos B. Papadias and Dirk T. M. Slock

Institut EURECOM, 2229 route des Crêtes, B.P. 193, 06904, Sophia Antipolis Cedex, FRANCE

Cette communication présente une nouvelle classe d'algorithmes d'égalisation aveugle pour la transmission des données MAP ou MAQ à travers un canal de communication à phase non-minimale. Cette classe d'algorithmes est derivée en minimisant un critère déterministe qui impose un ensemble de contraintes basées sur la proprieté du module constant de la constellation émise et permet l'accelération de la vitesse de convergence par rapport à celles des algorithmes CMA classiques.

This paper presents a new class of blind equalization algorithms for PAM or QAM data transmission over a possibly non-minimum phase communication channel. This class of algorithms is derived by minimizing a deterministic function that imposes a set of constraints based on the constant modulus property of the emitted symbol constellation and allows for an increased convergence speed as compared to that of classical CMA's.

1 Introduction

Blind equalization of digital communication channels is a domain that has gained increased attention over the last decade. A typical blind-equalization setup is depicted in figure 1. The purpose of the blind equalization algorithm is to make the equalizer match the impulse response of the inverse of the communication channel, thus opening the eye of the communication system and allowing for a correct retrieval of the emitted symbols. Let a_k denote the emitted symbol, x_k the channel's output (possibly corrupted by additive noise), y_k the equalizer's output and W_k the $N \times 1$ equalizer coefficient vector, all at time instant k. Also let $X_k = [x_k \; x_{k-1} \cdots x_{k-N+1}]^H$, where ^{H} denotes complex conjugate transposition. Then the equalizer's output at time instant k may be written as: $y_k = X_k^H W_k$. A

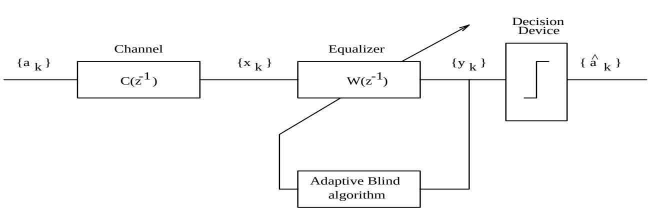

Figure 1: A typical blind equalization scheme

very popular class of blind equalization algorithms based on the constant modulus property of the emitted constellation symbols is the family of Godard blind equalizers [1], [2]. These algorithms minimize by a steepest-descent procedure the following cost function w.r.t. W:

$$J_p(W) = E\left\{\frac{1}{2p}(|y|^p - R_p)^2\right\} \qquad , \; p \in 1, 2, \ldots \; , \qquad (1)$$

where E denotes statistical expectation and R_p is a constant scalar called *dispersion constant* and defined by $R_p = \frac{E|a_k|^{2p}}{E|a_k|^p}$. The corresponding algorithm is given by:

$$W_{k+1} = W_k - \mu X_k y_k |y_k|^{p-2}(|y_k|^p - R_p) \quad , \qquad (2)$$

where μ is the algorithm's stepsize. In the sequel we assume for simplicity that $|a_k| \equiv 1$. The well-known SATO and CMA 2-2 algorithms are special cases of (2) for $p = 1$ and $p = 2$, respectively. A variant of CMA 2-2 called NCMA has been recently introduced in [5] by nulling the algorithm's *a posteriori* error at each iteration. This leads to the algorithm:

$$W_{k+1} = W_k - \frac{1}{\|X_k\|^2} X_k y_k (1 - \frac{1}{|y_k|}) \quad , \qquad (3)$$

where $\|.\|$ denotes the 2-norm of a vector in the Euclidean space. An alternative way of deriving (3) would be to solve at each iteration the following problem:

$$\begin{array}{c} \min\limits_{W_{k+1}} \; (|X_k^H W_{k+1}|^2 - 1)^2 \\ \text{subject to: } \|W_{k+1} - W_k\|^2 = \text{minimal} \, . \end{array} \qquad (4)$$

We will now present a new class of constant-modulus algorithms that minimize a deterministic criterion similar to the one in (4).

2 The proposed algorithms

2.1 Derivation

Consider the following Sliding Window extension of the normalized CMA 2-2 formulation:

$$\min_{W_{k+1}} \left\{ \sum_{i=0}^{L-1} (|X_{k-i}^H W_{k+1}|^p - 1)^2 \right\} \quad (5)$$
$$\text{subject to: } \|W_{k+1} - W_k\|^2 = \text{minimal}\,.$$

It is clear that with $L \leq N$, there exist enough degrees of freedom to choose W_{k+1} such that the sum in (5) takes on its minimal value zero. The first line of (5) leads to the following set of equations:

$$|X_{k-i}^H W_{k+1}| = 1 \quad , \quad i = 0, \cdots, L-1\,. \quad (6)$$

Now, as the quantity $W_{k+1} - W_k$ is a vector of length N, it can in general be decomposed as:

$$W_{k+1} - W_k = \mathbf{X}_k v_k + \Omega_k \ , \quad (7)$$

where $\mathbf{X}_k$ is a $N \times L$ matrix defined as: $\mathbf{X}_k = [X_k \ X_{k-1} \ldots X_{k-L+1}]$. The first term on the right-hand side of (7) represents the component of $W_{k+1} - W_k$ in the L-dimensional subspace of $\mathcal{R}^N$ spanned by $X_{k-i} \quad , i = 0, \cdots, L-1$, and Ω_k the component belonging to the orthogonal complement of this subspace, of dimension $N - L$. Therefore, the second line of (5) can be written as:

$$\|\mathbf{X}_k v_k\|^2 + \|\Omega_k\|^2 = \text{minimal}\,. \quad (8)$$

The term $\|\mathbf{X}_k v_k\|^2$ is completely determined by the system of equations (6), so that (8) is equivalent to minimizing the quantity $\|\Omega_k\|^2$,which results in $\Omega_k = \mathbf{0}$. Reporting this value for Ω_k in (7) and substituting in (6), we obtain:

$$X_{k-i}^H W_k + X_{k-i}^H \mathbf{X}_k v_k = e_i \ , \quad i = 0, \cdots, L-1\ , \quad (9)$$

where $|e_i| = 1, i = 0, \cdots, L-1$. Eq. (9) can be written in matrix form as follows:

$$\mathbf{X}_k^H W_k + \mathbf{X}_k^H \mathbf{X}_k v_k = e = [e_0 \cdots e_{L-1}]^T \ , \quad (10)$$

which gives when solved for v_k:

$$v_k = (\mathbf{X}_k^H \mathbf{X}_k)^{-1}(e - \mathbf{X}_k^H W_k)\,. \quad (11)$$

This equation, combined with (7) and $\Omega_k = \mathbf{0}$ gives the following recursive formula for the equalizer vector W:

$$W_{k+1} = W_k + \mathbf{X}_k (\mathbf{X}_k^H \mathbf{X}_k)^{-1}(e - \mathbf{X}_k^H W_k)\,. \quad (12)$$

In order to choose the *sign* vector e in an optimal way, it must be such that the second condition of (5) holds, i.e. $\|W_{k+1} - W_k\|$ is minimal. This leads [7] to the following choice for e :

$$e = [sign(X_k^H W_k) \ \cdots \ sign(X_{k-L+1}^H W_k)]^T \ , \quad (13)$$

where the *sign* of a complex scalar is defined as: $sign(re^{j\theta}) \equiv e^{j\theta}$ if $r \neq 0$. We also use the convention $sign(0) \equiv 1$. Substituting into (12), we have a recursive algorithm corresponding to the problem (5):

$$W_{k+1} = W_k + \mathbf{X}_k R_k^{-1}(sign(\mathbf{X}_k^H W_k) - \mathbf{X}_k^H W_k)\ , \quad (14)$$

where $R_k = \mathbf{X}_k^H \mathbf{X}_k$ and the *sign* of a vector is defined as the vector whose elements are the *signs* of the vector's elements. One can easily see (by following the same steps) that (14) solves also the following problem:

$$\min_{W_{k+1}} \left\{ \|sign(\mathbf{X}_k^H W_k) - \mathbf{X}_k^H W_{k+1}\|^2 \right\} \quad (15)$$
$$\text{subject to: } \|W_{k+1} - W_k\|^2 = \text{minimal}\,.$$

Now using a result in [3] we obtain the following result: An exact minimization with respect to W_{k+1} of the deterministic function

$$\|sign(\mathbf{X}_k^H W_k) - \mathbf{X}_k^H W_{k+1}\|^2_{R_k^{-1}} + (\frac{1}{\bar{\mu}} - 1)\|W_{k+1} - W_k\|^2 \ , \quad (16)$$

where $\|x\|_S^2 = x^H S x$, is provided at each iteration by the following algorithm:

$$W_{k+1} = W_k + \bar{\mu} \mathbf{X}_k R_k^{-1}(sign(\mathbf{X}_k^H W_k) - \mathbf{X}_k^H W_k)\,. \quad (17)$$

Note that the problem of minimizing (16) w.r.t. W_{k+1} reduces to the problem in (15) as $\bar{\mu} \to 1$. In (16) the hard constraints of (6) are replaced by a weighted minimization of the terms in (16).

2.2 Discussion

Eq. (17) describes a new parametric class of algorithms for blind equalization. The two adjustable parameters are the normalized stepsize $\bar{\mu}$ and the window length L. According to the deterministic criterion (16) minimized by the algorithm, the stepsize $\bar{\mu}$ controls the deviation of the new equalizer setting W_{k+1} from the previous one W_k in a square-norm sense. It can be shown [3] that the algorithm will have a stable operation for all values of $\bar{\mu} \in (0, 2)$ and moreover the fastest convergence speed will be achieved when $\bar{\mu} = 1$. On the other hand, the choice of L has to do with the degree of constraining of our criterion, i.e. at each iteration we are imposing L soft constraints on the next equalizer vector. These constraints are represented by the first term of the criterion (16). In the case $L = 1$ one immediately recognizes the recently-proposed[5] NCMA algorithm. Actually the derivation above offers an alternative way of deriving the NCMA that covers also the case $\bar{\mu} \neq 1$. On the other hand, when $L = N$ the algorithm becomes Sliding-Window RLS. For all the intermediate choices of $1 < L < N$ one gets other members of this class of algorithms that reside between the two extreme cases of NCMA and Sliding-Window RLS constant modulus algorithms and are expected to compromise for convergence speed and computational complexity in between the two extreme cases. This class of algorithms is in full analogy with the recently proposed BUCFTF[4] or UG/SWC FTF[3] classes of algorithms for adaptive filtering, that reside between the NLMS and Block/Sliding-Window RLS algorithms.

3 Computational organization

The updating of the current equalizer vector W_k described by eq. (17) may be organized in the following way [4]:
1. Compute the $L \times 1$ *a priori* error vector (NL multiplications):

$$E_k = sign(\mathbf{X}_k^H W_k) - \mathbf{X}_k^H W_k \quad (18)$$

2. Update R_k from R_{k-1} (close-to-Toeplitz matrix) ($O(L)$ multiplications).

3. Solve the following linear system by use of the *generalized Levinson* algorithm ($5.5L^2 + O(L^2)$ multiplications):

$$R_k H_k = \bar{\mu} E_k \quad (19)$$

4. Update the equalizer vector as follows (NL multiplications):

$$W_{k+1} = W_k - \mathbf{X}_k H_k \quad (20)$$

A further reduction in complexity may be achieved if the algorithm is implemented in a Block-updating form, i.e:

$$W_{k+M} = W_k - \bar{\mu}\mathbf{X}_k R_k^{-1}(\mathbf{X}_k^H W_k - sign(\mathbf{X}_k^H W_k)) \quad (21)$$

The algorithm (21) will have a complexity M times smaller than (17) but also a lower convergence speed. It is possible to use FFT techniques to perform the computations in (18) and (20) in $O(Nlog_2 L)$ operations. Also the computations in (19) can be reduced to $O(Llog_2 L)$ operations by using the displacement representation of R_k^{-1} (this necessitates the use of the FTF algorithm for propagating the generators of R_k^{-1} in $O(L)$ operations). These alternative computations are interesting when L is large. Finally, if we introduce an approximation in NSWCMA corresponding to taking

$$e = [sign(X_k^H W_k) \ \cdots \ sign(X_{k-L+1}^H W_{k-L+1})]^T \quad (22)$$

(compare to (13)), the USWC FTF algorithm of [3] may be used to run NSWCMA in $12N + O(L)$ operations/sample.

4 Algorithm's performance

We will separately examine the algorithm's performance in the noiseless and the noise-present cases. In contrast to the classical adaptive filtering context where the noise influence is modeled as an additive noise contaminating the desired signal, the modelling of noise in an equalization setup is somewhat different, i.e. the received *data* at the equalizer's input are assumed to be corrupted by additive (usually white Gaussian) noise. Moreover, as the communication channel is usually modeled as a FIR filter and the equalizer is a FIR filter itself too, there is a second kind of "noise" arising from the fact that the equalizer has only a finite number of taps and thus can never exactly match the infinite-length impulse response of the channel's inverse. Therefore, in order to examine a "noiseless" case we will model the channel as an AR(N-1) filter and the equalizer as an MA(N-1) filter thus avoiding order-mismatch noise.

4.1 Noiseless case

Based on the above assumption, we consider the algorithm when $\bar{\mu} = 1$ and $L = N$. In a classical adaptive filtering context where the desired signal is available, the algorithm would converge after one single iteration [4]. In the blind-equalization context the desired signal is replaced by the vector $[sign(\mathbf{X}_k^H W_k) \ \ sign(\mathbf{X}_{k-1}^H W_k) \cdots sign(\mathbf{X}_{k-L+1}^H W_k)]^T$. When the elements of this vector coincide with the transmitted symbols at time instants $k, k-1, \ldots, k-L+1$, then the algorithm will converge in one more iteration. This rather deterministic than stochastic approach makes sense in this case, since the algorithm forces exactly at each iteration a deterministic criterion rather than a stochastic one. Moreover, as L grows from 1 to N the algorithm becomes less and less sensitive to the colouring of the input signal in contrast to NCMA (NSWCMA with $L = 1$) which will exhibit a very slow convergence rate when the input signal is strongly coloured. Moreover, as the absence of noise allows for the use of a big stepsize $\bar{\mu}$, these algorithms may avoid converging to a local minimum of their cost function as opposed to classical CMA's that may get more easily trapped by local-minima (see [6] for a more elaborate discussion).

4.2 Noise present

We now consider the realistic case of a FIR channel and an additive white Gaussian noise corrupted received signal. We can again draw a parallel with the performance of BUCFTF. As explained in [4], the conditioning of the sample covariance matrix R_k has a significant impact on the convergence of the BUCFTF algorithm when noise is present. In fact the noise contribution to the MMSE is directly proportional to the following term:

$$\alpha = \bar{\mu}\mathbf{X}_k R_k^{-2}\mathbf{X}_k^H \ . \quad (23)$$

Now if one uses the SVD of the data matrix $\mathbf{X}_k = U\Sigma V^H$, (23) becomes:

$$\alpha = \bar{\mu}U\Sigma^{-2}U^H \ . \quad (24)$$

This last equation shows that when the matrix $R_k = \mathbf{X}_k^H\mathbf{X}_k$ has a big eigenvalue spread, the noise will get amplified in a large and very disproportional manner along the different eigen-directions of this matrix, thus resulting in a big steady-state MSE. This will become more and more severe as L moves from 1 to N, which will be the worst case from this point of view. Different remedies to this problem might be either regularizing in some way the sample covariance matrix R_k or reducing the stepsize $\bar{\mu}$. In the first case one deviates from satisfying exactly the deterministic criterion (16), in the second case one reduces the convergence speed. In the NSWCMA class of algorithms, a similar behaviour is expected, deteriorated also by the fact that the vector $[sign(\mathbf{X}_k^H W_k) \ \ sign(\mathbf{X}_{k-1}^H W_k) \cdots sign(\mathbf{X}_{k-L+1}^H W_k)]^T$ might deviate very much at some iterations from the actually transmitted symbols at time instants $k, k-1, \ldots, k-L+1$. However, a faster convergence than NCMA's may still be achieved.

5 Simulations

The behaviour discussed above of the proposed class of algorithms has been verified through computer simulations. We will again discriminate between a noiseless and a noise-present case. The emitted symbols are assumed to belong to a binary alphabet (2-PAM constellation) and the transmitting channel to be a linear filter. As a measure of performance we will use the evolution in time of the so-called *closed-eye measure* of the communication system defined as follows for a constant modulus constellation:

$$\rho = \frac{\sum_i |h_i| - max_i|h_i|}{max_i|h_i|} \ , \quad (25)$$

where $\{h_i\}$ represents the convolution of the channel and the equalizer impulse responses.

5.1 Noiseless AR channel

In a first experiment we consider an AR(7) linear noiseless channel and an MA(7) equalizer ($N = 8$). The channel's poles are located at $\rho_1 = 0.1, \rho_{2,3} = 0.2e^{\pm j\pi/4}, \rho_{4,5} = 5e^{\pm j\pi/4}, \rho_{6,7} = 6e^{\pm j\pi/6}$, i.e. far from the unit circle. This corresponds to a well conditioned sample covariance matrix R_k. The SATO and CMA algorithms with step-sizes that have been found by trial and error to guarantee stability as well as the NSWCMA with unit stepsize ($\bar{\mu} = 1$) and three different choices for L (1,3,8) have been implemented. Figure 2(a) shows the evolution of the closed-eye measure of the communication system for the different algorithms during 1000 iterations. One can see the faster convergence provided by NCMA w.r.t. SATO and CMA and the further increase of convergence speed of NSWCMA as L grows up towards N. In a second experiment another AR(7) channel is chosen with poles $\rho_1 = 0.3, \rho_{2,3} = 0.5e^{\pm j\pi/4}, \rho_{4,5} = 1.5e^{\pm j\pi/4}, \rho_{6,7} = 2e^{\pm j\pi/6}$. This channel has its poles closer to the unit circle than the previous one, thus resulting in a more strongly coloured received signal and a more ill-conditioned matrix R_k. Figure 2(b) shows again the evolution of the closed-eye measure by different algorithms in this case. We can see that all algorithms except NSWCMA $L = N$ have now a lower convergence speed because of the colouring of the received signal. However, as expected, NSWCMA $L = N$ seems to be insensitive to this colouring and to converge at the same speed as for the previous channel.

5.2 Noisy FIR channel

We now examine a more realistic case of a FIR channel, corrupted by 20 dB of additive noise. The channel's impulse response is $[1\ 2\ 0.6]^T$ (relatively ill-conditioned). The equalizer is implemented as an MA(5) filter (corresponding to a 6 coefficients tapped-delay line). The emitted sequence is again 2-PAM and the NSWCMA algorithm is implemented. In order to avoid a very big steady-state error, we have used the following regularization for the calculation of the matrix R_k^{-1}: as the generalized-Levinson algorithm provides us automatically a LDU decomposition of the matrix R_k ($R_k = LDL^T$ with L lower triangular), we calculate R_k^{-1} as $R_k^{-1} = L^{-T}D^{-1}L^{-1}$. In order to avoid the effect of severe ill-conditioning of the matrix R_k at certain time instants when R_k has a very big eigenvalue spread, we consider all diagonal elements of D that are less than 5000 times smaller than its maximum-valued element to be zero, and thus their inverse is also set to zero (pseudo-inverse of D). Figure 3 shows the closed-eye measure evolution of NSWCMA for three different values of L ($L = 1$ (NCMA), 3 and 5) averaged over 100 different experiments (realizations) for each algorithm. As L increases, we reduce the stepsize in order to combat the growing influence of noise. We notice that both algorithms ($L = 3, \bar{\mu} = 0.2$ and $L = 5, \bar{\mu} = 0.1$) provide a faster convergence as compared to the case $L = 0, \bar{\mu} = 1$ (NCMA of maximum convergence speed). We have also run simulations for the modified NSWCMA corresponding to (22) and have found a negligible degradation in convergence speed for moderate values of L.

6 Conclusions

A new class of algorithms for blind equalization called Normalized Sliding Window Constant Modulus Algorithms has been introduced. These algorithms correspond to an exact minimization of the deterministic criterion (16) and may be viewed as "blind" counterparts of the BUCFTF[4] and UG/SWC FTF[3] classes of adaptive filtering algorithms. The algorithms of this class have been shown to have a faster convergence speed w.r.t. classical CMA's.

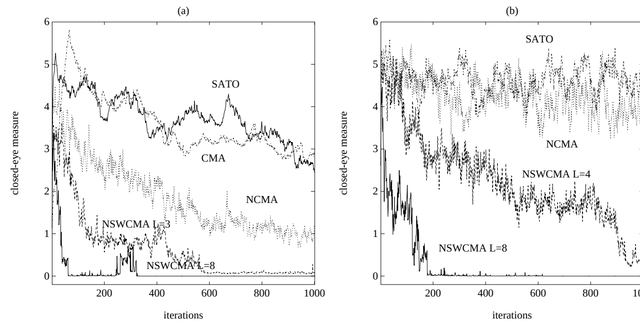

Figure 2: Noiseless case : two examples

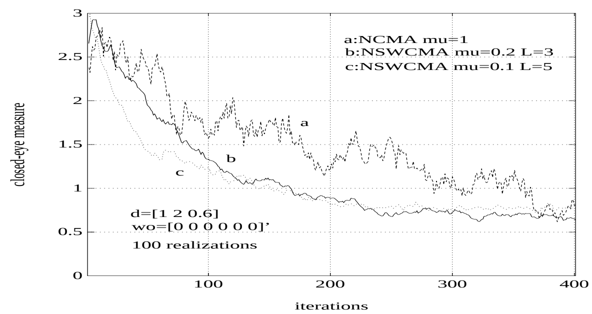

Figure 3: Noisy FIR channel

REFERENCES

[1] D. N. Godard, "Self-Recovering Equalization and Carrier Tracking in Two-Dimensional Data Communications Systems," *IEEE Trans. Commun.*, vol. COM-28, pp. 1867-1875, Nov. 1980.

[2] J. R. Treichler and B. G. Agee, "A new Approach to Multipath Correction of Constant Modulus Signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 459-472, Apr. 1983.

[3] D. T. M. Slock, "Underdetermined Growing and Sliding Window Covariance Fast Transversal Filter RLS algorithms," *Proc. EUSIPCO 92, VIth European Signal Processing Conference*, pages 1169-1172, Brussels, Belgium, Aug. 24-27, 1992.

[4] D. T. M. Slock, "The Block Underdetermined Covariance (BUC) Fast Transversal Filter (FTF) Algorithm for Adaptive Filtering," *Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Oct. 26-28, 1992.

[5] K. Hilal and P. Duhamel, "A Convergence study of the Constant Modulus Algorithm Leading to a Normalized-CMA and a Block-Normalized-CMA," *Proc. EUSIPCO 92, VIth European Signal Processing Conference*, pages 135-138, Brussels, Belgium, Aug. 24-27, 1992.

[6] C. B. Papadias and D. T. M. Slock, "On the Convergence of Normalized Constant Modulus Algorithms for Blind Equalization," *Proc. DSP International Conference on Digital Signal Processing*, Nicosia, Cyprus, July 14-16, 1993.

[7] C. B. Papadias and D. T. M. Slock, "A class of Normalized Sliding Window Constant Modulus Algorithms for Blind Equalization," *research report 1993*, Institut Eurecom.