Normalized Sliding Window Constant Modulus and Decision-Directed Algorithms: a Link Between Blind Equalization and Classical Adaptive Filtering

Constantinos B. Papadias[1] *Dirk T.M. Slock*[2]

Abstract

By minimizing a deterministic criterion of the constant modulus (CM) type or of the decision-directed (DD) type, we derive normalized stochastic gradient algorithms for blind linear equalization (BE) of QAM systems. These algorithms allow us to formulate CM and DD separation principles, which help obtain a whole family of CM or DD BE algorithms from classical adaptive filtering algorithms. We focus on the algorithms obtained by using the Affine Projection adaptive filtering Algorithm (APA). Their increased convergence speed and ability to escape from local minima of their cost function, make these algorithms very promising for BE applications.

Permission to publish this abstract separately is granted.

Accepted for publication as a correspondence item in the Special Issue on **Signal Processing for Advanced Communications**, *IEEE Transactions on Signal Processing.*

Present version: September 9, 1996.

Correspondence to:

Constantinos Papadias
Information Systems Laboratory
Durand Bldg.
Stanford University
Stanford, CA 94305-4055
USA

E-mail: papadias@rascals.stanford.edu Tel: +1 415 723 2873 Fax: +1 415 723 8473

[1]Constantinos Papadias was with the Eurécom Institute and is now with the Information Systems Lab, Stanford University, Stanford, CA 94305-4055, USA.

[2]Dirk Slock is with the Eurécom Institute, B.P. 193, 06904 Sophia Antipolis Cedex, France.

1 Derivation of the Normalized CMA and DD Algorithms

We address the problem of blind equalization of linear channels in digital communication systems that employ Quadrature Amplitude Modulation (QAM), as depicted in Figure 1. Assuming a time invariant channel, the channel and equalizer outputs at the Baud rate are given by

$$\begin{aligned} x_k &= \textstyle\sum_{i=0}^{M-1} c_i a_{k-i} + n_k = A_k^H C + n_k \\ y_k &= X_k^H W_k \end{aligned} \tag{1}$$

respectively, where $C = [c(0) \;\cdots\; c(M-1)]^T$ is the (FIR($M-1$)) channel impulse response, $W_k = [w_k(0) \cdots w_k(N-1)]^T$ is the equalizer vector at time instant k, (N is the equalizer length), a_k, n_k the channel input and output additive noise sample, respectively, at time instant k, $A_k = [a_k \;\cdots\; a_{k-M+1}]^H$, $X_k = [x_k \;\cdots\; x_{k-M+1}]^H$ and ^{T},^{H} denote transpose and Hermitian transpose, respectively.

We are interested in deriving appropriately normalized versions of stochastic gradient algorithms for the Godard [1] (or CM [2]), and DD cost functions. These normalized versions will come about by pursuing the deterministic point of view that stochastic gradient algorithms take when faced with the minimization of a statistical average. Consider first the Godard criterion

$$\min_W \; E \left| |X_k^H W|^p - 1 \right|^q \; . \tag{2}$$

A stochastic gradient algorithm drops the expectation operator and minimizes the resulting stochastic cost function by performing one iteration per sample period. So at time k we have a filter estimate W_k and we want to adapt it by considering the following instantaneous optimization problem

$$\min_{W_{k+1}} \left| |X_k^H W_{k+1}|^p - 1 \right|^q \; . \tag{3}$$

It is clear that we can minimize this cost function perfectly (making it zero) leading to

$$|X_k^H W_{k+1}| = 1 \tag{4}$$

$\forall\, p \geq 1,\; q \geq 1$, while leaving W_{k+1} largely undetermined (assuming $N > 1$). In order to fix the remaining degrees of freedom in W_{k+1}, we shall impose that W_{k+1} remains as close as possible to the prior estimate W_k while satisfying the constraint (4) imposed by the new data, leading to

$$\min_{W_{k+1}:\, |X_k^H W_{k+1}|=1} \|W_{k+1} - W_k\|^2 \tag{5}$$

where $\|W\|^2 = W^H W$ for $W \in \mathcal{C}^N$. Let us introduce the unit circle $\mathcal{M} = \{z \in \mathcal{C} : |z| = 1\}$. Then (5) can be written as

$$\min_{W_{k+1},\, d_k \in \mathcal{M}:\, X_k^H W_{k+1} = d_k} \|W_{k+1} - W_k\|^2 \; . \tag{6}$$

For the decision-directed case, the cost function is

$$\min_{W,d_k\in\mathcal{A}} E\,|d_k - X_k^H W|^2 \tag{7}$$

where $\mathcal{A}$ is the symbol alphabet (constellation). For a stochastic gradient algorithm, this becomes the instantaneous cost function

$$\min_{W_{k+1},d_k\in\mathcal{A}} |d_k - X_k^H W_{k+1}|^2 \tag{8}$$

which can again be minimized perfectly by making W_{k+1} and d_k satisfy the following constraint

$$X_k^H W_{k+1} \;=\; d_k \;\in\; \mathcal{A}\,. \tag{9}$$

Keeping again W_{k+1} as close as possible to W_k while satisfying the constraint (9) results in

$$\min_{W_{k+1},\,d_k\in\mathcal{A}:\,X_k^H W_{k+1}=d_k} \|W_{k+1} - W_k\|^2\;. \tag{10}$$

Hence both the Godard and the decision-directed cost functions lead to the following optimization problem

$$\min_{W_{k+1},\,d_k\in\mathcal{D}:\,X_k^H W_{k+1}=d_k} \|W_{k+1} - W_k\|^2 \quad,\quad \mathcal{D} = \begin{cases} \mathcal{M}\,, & \text{Godard (CMA)} \\ \mathcal{A}\,, & \text{decision-directed (DD)}\end{cases}\;. \tag{11}$$

This constrained problem with the hard constraints (4) or (9) can be seen to be the limiting case of the following soft-constrained problem

$$\min_{W_{k+1},\,d_k\in\mathcal{D}} \left\{ |d_k - X_k^H W_{k+1}|^2 + (\frac{1}{\overline{\mu}}-1)\,\|X_k\|^2\,\|W_{k+1}-W_k\|^2 \right\} \tag{12}$$

where $\overline{\mu}\in(0,1)$. Indeed, as $\overline{\mu}\to 1$, problem (12) becomes problem (11) (this can be verified also by checking that as $\overline{\mu}\to 1$, the solution of the soft-constrained problem, which we shall develop below, becomes the solution to the hard-constrained problem). The parameter $\overline{\mu}$ regulates the compromise between satisfying the new data at time k (the first term) and sticking to the a priori knowledge (the second term). To solve (12), remark that the problem is separable. Indeed, if at first we assume d_k to be known, then we can solve the remaining quadratic problem for W_{k+1}. The solution can be found to be

$$\begin{aligned} e_k &= d_k - X_k^H W_k \\ W_{k+1} &= W_k + \overline{\mu} X_k (X_k^H X_k)^{-1}\, e_k \end{aligned} \tag{13}$$

where e_k is the a priori error signal. Among all the possible values of d_k we must choose the one that satisfies the criterion (12). It turns out that the solution for the desired response d_k is

$$d_k \;=\; f(y_k) \quad \text{where} \quad f(\cdot) = \begin{cases} sign(\cdot) & ,\ \mathcal{D}=\mathcal{M} \quad \text{(CMA)} \\ dec(\cdot) & ,\ \mathcal{D}=\mathcal{A} \quad \text{(DD)}\end{cases}\;. \tag{14}$$

$y_k = X_k^H W_k$ is the prior equalizer output, $sign(z) = \frac{z}{|z|}$ for $z \in \mathcal{C}$ (with the convention $sign(0) = 1$), and the decision function $dec(z)$ returns the element in the alphabet $\mathcal{A}$ that is closest to z. In summary, the solution to both problems (11) and (12) is

$$W_{k+1} \;=\; W_k + \overline{\mu} X_k (X_k^H X_k)^{-1} \left(f(X_k^H W_k) - X_k^H W_k \right) \;. \tag{15}$$

The resulting algorithm minimizes the deterministic criterion (12) at each time step. The algorithm is normalized in that the step-size parameter $\overline{\mu}$ can be chosen anywhere in the signal-independent interval $\overline{\mu} \in (0,1]$ for stable operation with fastest convergence for $\overline{\mu} = 1$ (for $\overline{\mu} \in (1,2)$, we can replace $\overline{\mu}$ by $2-\overline{\mu}$ and get the same convergence dynamics but more estimation noise, so there is no point in choosing $\overline{\mu} \in (1,2)$). The posterior and prior error signals are related as

$$\epsilon_k \;=\; d_k - X_k^H W_{k+1} \;=\; (1-\overline{\mu})\, e_k \tag{16}$$

which shows that an update reduces the error for $\overline{\mu} \in (0,1]$. A related signal is the posterior output

$$\mathrm{y}_k \;=\; X_k^H W_{k+1} \;=\; d_k + (1-\overline{\mu}) e_k \;=\; \overline{\mu}\, d_k + (1-\overline{\mu})\, y_k \;. \tag{17}$$

This shows that the posterior output is a convex combination of the prior output and the desired response. Hence, the posterior output will be closer to the desired response than the prior output. The stepsize $\overline{\mu}$ controls the extent to which the posterior output approaches the desired response. This issue is illustrated explicitly in Figure 5 for the normalized CMA and DD algorithms. For the CMA algorithm, the desired response consists of the projection of the prior output onto the unit circle. The result of doing an update is to bring the posterior equalizer output closer to the unit circle without changing the phase. This is a deterministic version of the CM philosophy. For the DD algorithm, the desired response is the closest constellation point and the posterior output approaches it also to an extent that is controlled by $\overline{\mu}$.

The term Normalized CMA (NCMA) algorithm was coined in [3] where it was derived for $\overline{\mu} = 1$ by setting the posterior Godard criterion equal to zero.

2 A Separation Principle for Blind Equalization

We can generalize to some extent the criterion that led to the normalized CMA/DD algorithms. The generalized criterion is

$$\min_{W_{k+1},\, d_k \in \mathcal{D}} \left\{ |d_k - X_k^H W_{k+1}|^2 + \|W_{k+1} - W_k\|^2_{T_k} \right\} \tag{18}$$

where $\|W\|_A^2 = W^H A W$ and $T_k = T_k^H > 0$. The criterion (18) is again separable. Assuming at first d_k to be known, we can optimize w.r.t. W_{k+1}. The solution can again be found to be

$$W_{k+1} = W_k + T_k^{-1} X_k (1 + X_k^H T_k^{-1} X_k)^{-1} \left(f(X_k^H W_k) - X_k^H W_k \right) . \tag{19}$$

We can interpret this solution to the blind equalization criterion (18) as corresponding to the solution of the adaptive filtering problem obtained from (18) by considering the d_k to be a given desired-response signal, augmented with the construction of the desired response d_k from the prior output y_k as $d_k = f(y_k)$. This leads to the following principle.

A separation principle for BE: *an adaptive blind equalization algorithm of the CMA/DD-type can be obtained by taking a classical adaptive filtering algorithm and replacing its desired response signal d_k by $f(X_k^H W_k)$.*

The separation principle applies exactly for all adaptive filtering algorithms of the form (19) meaning that for the corresponding blind equalization criteria (18), the solution is obtained exactly as the corresponding adaptive filtering algorithm augmented with the desired-response construction $d_k = f(y_k)$. However, we can apply the principle in a loose sense to adaptive filtering algorithms of another form. A number of applications of the separation principle are sketched in Table 1. For the first application (LMS), we take $T_k = (\frac{1}{\mu} - \|X_k\|^2)\, I_N$ in (18). For the previously derived normalized algorithms, we take $T_k = (\frac{1}{\bar{\mu}} - 1)\|X_k\|^2\, I_N$ in (18). For the third application (RLS), we take $T_k = \lambda R_{k-1}$ where $R_k = \sum_{i=0}^{k} \lambda^{k-i} X_i X_i^H + R_{-1}$ and λ is the exponential weighting factor. For the RLS algorithm (see also [4]), any fast version could be used such as the Fast Transversal Filter algorithm or the Fast Lattice/QR algorithms [5]. Also the Fast Newton Transversal Filter algorithm [6] could be used. The last application involves the APA algorithm which we discuss now.

3 The NSWCMA and NSWDD algorithms

The Affine Projection Algorithm (APA) is a generalization of the NLMS algorithm in which the data at the L latest time instants are explicitly taken into account

$$\|D_k - \mathbf{X}_k^H W_{k+1}\|^2_{P_k^{-1}} + (\frac{1}{\bar{\mu}} - 1)\|W_{k+1} - W_k\|^2 \tag{20}$$

where $P_k = \mathbf{X}_k^H \mathbf{X}_k$ and $\mathbf{X}_k$, D_k are defined as

$$\mathbf{X}_k = [X_k\; X_{k-1} \ldots X_{k-L+1}] \;\; (N \times L)\,, \quad D_k = [d_k\; d_{k-1} \cdots d_{L-1}]^T . \tag{21}$$

The criterion (20) is minimized exactly at each iteration by the following algorithm:

$$W_{k+1} \;=\; W_k + \overline{\mu}\mathbf{X}_k P_k^{-1}(D_k - \mathbf{X}_k^H W_k)\ . \tag{22}$$

The APA algorithm ([7]) was first proposed by Ozeki and Umeda in 1984. Its derivation based on a deterministic criterion was given in [8], where a fast version, the USWCFTF algorithm, was derived. A fast block version was derived in [9] (and independently in [10]) and given the name of BUCFTF algorithm. The fastest version was derived by Gay in [11] (see also [12]) and termed the Fast AP (FAP) algorithm.

A loose application of the separation principle to the APA algorithm (instead of taking $d_k = f(X_k^H W_k)$ we take $D_k = f(\mathbf{X}_k^H W_k)$ where $f(\cdot)$ operating on a vector gives the vector of the elementwise operations) yields the following class of algorithms for blind equalization:

$$W_{k+1} = W_k + \bar{\mu}\mathbf{X}_k(\mathbf{X}_k^H\mathbf{X}_k)^{-1}(D_k - \mathbf{X}_k^H W_k) \tag{23}$$

where

$$D_k \;=\; \begin{cases} sign(\mathbf{X}_k^H W_k) \\ dec(\mathbf{X}_k^H W_k) \end{cases} \tag{24}$$

In the case $D_k = sign(\mathbf{X}_k^H W_k)$, (23) is the update equation of a CM-type class of algorithms called *Normalized Sliding Window Constant Modulus Algorithms (NSWCMA)*, whereas in the case $D_k = dec(\mathbf{X}_k^H W_k)$ (23) describes the class of *Normalized Sliding Window Decision Directed (NSWDD) Algorithms*. In the rest of the paper we will mainly focus on the NSWCMA:

$$W_{k+1} = W_k + \bar{\mu}\mathbf{X}_k(\mathbf{X}_k^H\mathbf{X}_k)^{-1}(sign(\mathbf{X}_k^H W_k) - \mathbf{X}_k^H W_k) \tag{25}$$

which corresponds to an exact minimization at each time step of the criterion

$$\min_{W_{k+1}} \left\{ \|sign(\mathbf{X}_k^H W_k) - \mathbf{X}_k^H W_{k+1}\|^2_{P_k^{-1}} + (\frac{1}{\bar{\mu}} - 1)\|W_{k+1} - W_k\|^2 \right\}\ . \tag{26}$$

Eq. (25) describes a new parametric class of algorithms for BE of CM signals, parameterized by the data window length L and the stepsize $\bar{\mu}$. L is an integer that expresses the number of "CM constraints" imposed on the next equalizer setting at each iteration and varies from 1 to N. The imposition of L constraints results in a prewhitening of the input signal by linear prediction of order $L-1$ (see [8]). The prediction filter involved corresponds to the $L \times L$ sample covariance matrix (P_k), which is constructed from the $N \times L$ data matrix ($\mathbf{X}_k$) and hence corresponds to passing the data through a *rectangular sliding window*. $\bar{\mu}$ is a real scalar that controls the deviation of the new equalizer setting w.r.t. the previous one. Strictly speaking, only when $\bar{\mu} = 1$ do we impose a set of L constraints on W_{k+1}. When $\bar{\mu} \neq 1$, we compromise between these (soft) constraints and the change in the filter through the minimization of two weighted additive terms in the deterministic criterion (26). Compared to the classical Godard or CMA BE algorithms, the new algorithms replace a stochastic gradient technique with the exact minimization at each iteration of a deterministic criterion, which involves $L-1$ previous regressors and hence adds memory. This memory aspect has a beneficial impact on the convergence speed, namely the convergence speed will increase as L increases. The particular deterministic criterion also provides the important feature of *normalization*. As a result, it is possible to choose the stepsize $\bar{\mu}$ for stability ($\bar{\mu} \in (0, 2)$) or fastest convergence ($\bar{\mu} = 1$) independently of the input signal. This fact furthermore has a positive impact in avoiding the problem of ill-convergence often observed in BE algorithms of the Godard type (see [13],[17]). The influence of noise on the algorithm has been discussed in [14],[15]: when L approaches N, noise amplification can become severe, however the problem is less pronounced when fractionally-spaced equalization is used. A method to further overcome the noise amplification problem has been also presented in [15].

If we apply the (FAP) algorithm [11] the algorithm's complexity is $2N+20L$ operations/iteration, the additive linear term in L representing the price paid for the extra constraining as compared to the LMS algorithm.

When using the NSWCMA on non-CM constellations, the function $sign(\cdot)$ should be replaced by $f(\cdot) = r_1\, sign(\cdot)$ (r_p is the *dispersion constant* defined as $r_p = \frac{E|a_k|^{2p}}{E|a_k|^p}$).

4 Computer simulation results

In a first experiment, we consider the typical mobile multipath channel given in [16], through which we transmit a white 4-QAM sequence. The received signal has an SNR of 30 dB, it is sampled at the rate $2/T$ and we use an equalizer of 33 taps. Figure 3 shows the evolution of the closed-eye measure for the CMA and two members of the NSWCMA algorithms, used with the following parameters: in (a) $\bar{\mu} = 0.3,\ L = 4$, in (b) $\bar{\mu} = 1,\ L = 1$, and in (c) $\mu = 0.05$ (the closed eye measure is defined as $\rho = \frac{\sum_i |h_i| - \max_i(|h_i|)}{\max_i(|h_i|)}$, where h is the channel-equalizer cascade impulse response). Note how the convergence speed increases for the normalized algorithms, and the more so as the number L of CM constraints increases.

In a second experiment, we consider an all-pole AP(1) noiseless channel defined by $C(z) = \frac{1}{1+0.25z^{-1}}$, through which we transmit a white 2-PAM sequence. Figure 4 shows the convergence trajectories for 40 different initializations on a circle of radius 2 in the equalizer space (the FIR equalizer length is $N = 2$) of three different algorithmic implementations: in (4a), we have tested the CMA (with a stepsize found by trial and error to guarantee stability), in (4b) the NSWCMA with $\bar{\mu} = 0.05$, $L = 2$, and in (4c) the NSWCMA with $\bar{\mu} = 1$, $L = 2$. As can be seen from the figure, the CMA may end up, depending on its initialization, either at one of its two global minima, or at one of the two local minima on the axis $w(0) = 0$. On the other hand, the NSWCMA with a small stepsize ($\bar{\mu} = 0.05$) manages to escape in some cases from its local towards the global minima, whereas it ends up only at its global minima when used with the large stepsize $\bar{\mu} = 1$. This verifies the theoretical analysis in [13], [17], which predicts the ability of the normalized algorithms to escape more easily from the local minima of their cost function. A similar behavior has been observed for noisy FIR channels.

5 Conclusions

Based on the derivation of normalized algorithms of the CM or the DD type for blind equalization, we have proposed a general methodology for the design of blind adaptive algorithms of these types. Our focus on the NSWCMA algorithm has shown its improved performance in terms of convergence speed and ability to avoid the problem of ill convergence.

References

[1] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. on Communications*, vol. COM-28, No 11, 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. on Acoustics, Speech, and Signal Processing*, vol. ASSP-31, pp. 459-472, April 1983.

[3] 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*, pp. 135-138, Brussels, Belgium, Aug. 24-27, 1992.

[4] B. G. Agee, "The least-squares CMA: a new technique for rapid correction of constant modulus signals," *Proc. ICASSP-86*, pp. 953-956, Tokyo, 1986.

[5] R. Gooch, M. Ready, and J. Svoboda, "A lattice-based constant modulus adaptive filter," *Proc.*

Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, 1987.

[6] G. V. Moustakides, S. Theodoridis, "Fast Newton transversal filters - a new class of adaptive estimation algorithms, " *IEEE Trans. on Signal Processing*, vol. 39, No. 10, pp. 2184-2193, Oct. 91.

[7] K. Ozeki, T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties," *Electronics and Communications in Japan*, vol. 67-A, No. 5, 1984.

[8] D. T. M. Slock, "Underdetermined growing and sliding window covariance fast transversal filter RLS algorithms," *Proc. EUSIPCO 92, VIth European Signal Processing Conference*, pp. 1169-1172, Brussels, Belgium, Aug. 24-27, 1992.

[9] 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.

[10] M. Montazeri and P. Duhamel, "A set of algorithms linking NLMS and block RLS algorithms," *IEEE Trans. on Signal Processing*, vol. 43, No. 2, pp. 444-453, Feb. 1995.

[11] S. L. Gay, "A fast affine projection algorithm," *Proc. ICASSP-95*, pp. 3023-3026, Detroit, Michigan, May 8-12, 1995.

[12] M. Tanaka, Y. Kaneda, S. Makino, J. Kojima, "Fast Projection Algorithm and Its Step Size Control," *Proc. ICASSP-95*, pp. 945-948, Detroit, Michigan, May 8-12, 1995.

[13] 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*, pp. 245-250, Nicosia, Cyprus, July 14-16, 1993.

[14] C. B. Papadias and D. T. M. Slock, "Normalized sliding window constant modulus algorithms for blind equalization," 14^{th} *GRETSI Symposium on Signal and Image Processing*, pp. 507-510, Juan les Pins, France, Sept. 13-16, 1993.

[15] C. B. Papadias and D. T. M. Slock, "New adaptive blind equalization algorithms for constant modulus constellations," *Proc. ICASSP-94*, pp. III-321-324, Adelaide, Australia, April 19-22, 1994.

[16] J. J. Shynk, R. P. Gooch, G. Krishnamurthy and C. K. Chan, "A comparative performance study of several blind equalization algorithms," *SPIE*, vol. 1565, pp. 102-117, 1991.

[17] C. B. Papadias, "Methods for blind equalization and identification of linear channels," PhD thesis, Ecole Nationale Supérieure des Télécommunications, ENST-95 E 006, Paris, March 1995.

LMS	$W_{k+1} = W_k + \mu X_k (d_k - y_k)$
CMA 1-2/DD	$W_{k+1} = W_k + \mu X_k (f(y_k) - y_k)$
NLMS	$W_{k+1} = W_k + \frac{\overline{\mu}}{\|\|X_k\|\|^2} X_k (d_k - y_k)$
NCMA/NDD	$W_{k+1} = W_k + \frac{\overline{\mu}}{\|\|X_k\|\|^2} X_k (f(y_k) - y_k)$
RLS	$W_{k+1} = W_k + R_k^{-1} X_k (d_k - y_k)$
RLSCMA/RLSDD	$W_{k+1} = W_k + R_k^{-1} X_k (f(y_k) - y_k)$
APA	$W_{k+1} = W_k + \overline{\mu}\mathbf{X}_k P_k^{-1} (D_k - Y_k)$
NSWCMA/NSWDD	$W_{k+1} = W_k + \overline{\mu}\mathbf{X}_k P_k^{-1} (f(Y_k) - Y_k)$

Table 1: Applications of the BE separation principle.

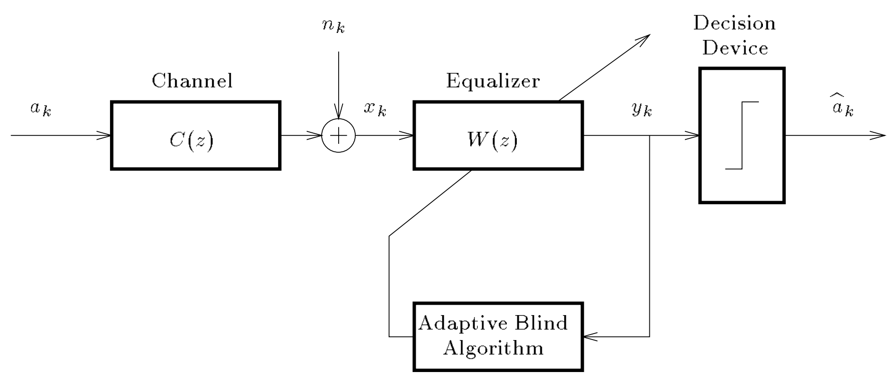

Figure 1: A typical blind linear equalization (BE) scheme in baseband and after baud rate sampling.

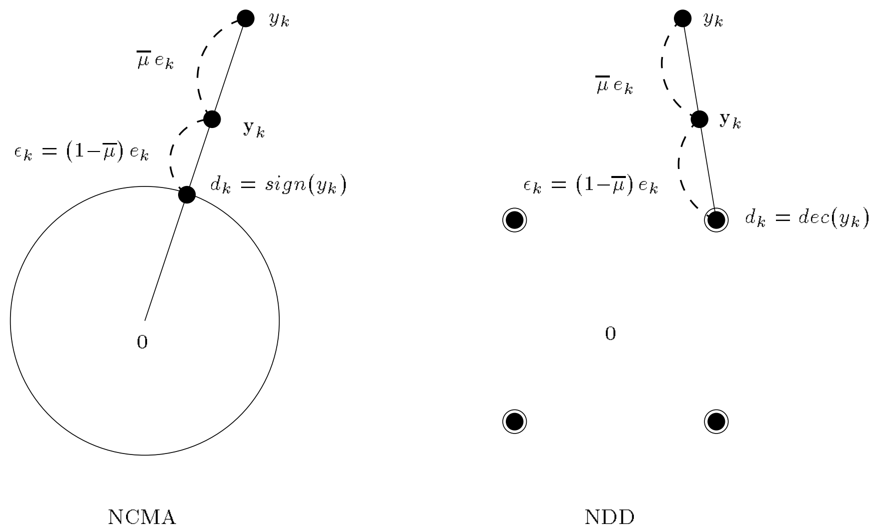

Figure 2: The configuration of the a priori output y_k, the a posteriori output y_k, the desired response d_k, the a priori error signal e_k and the a posteriori error signal ϵ_k for the normalized CMA and DD algorithms (assuming a 4-QAM constellation).

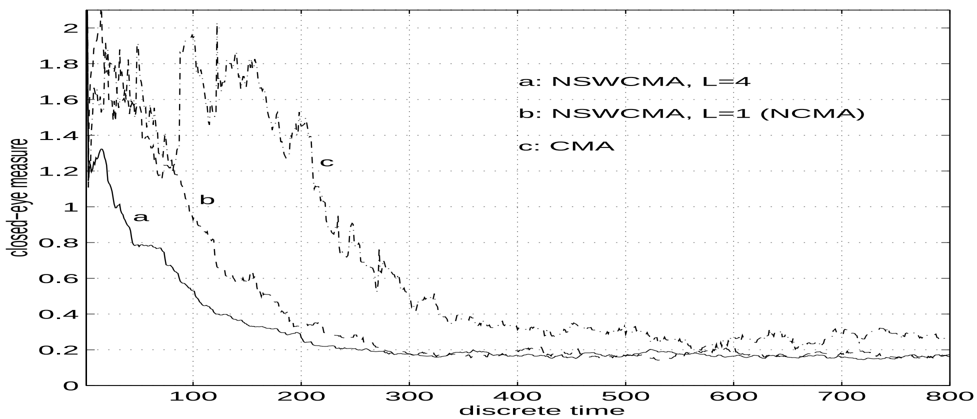

Figure 3: Noisy FIR channel, comparison in terms of convergence speed.

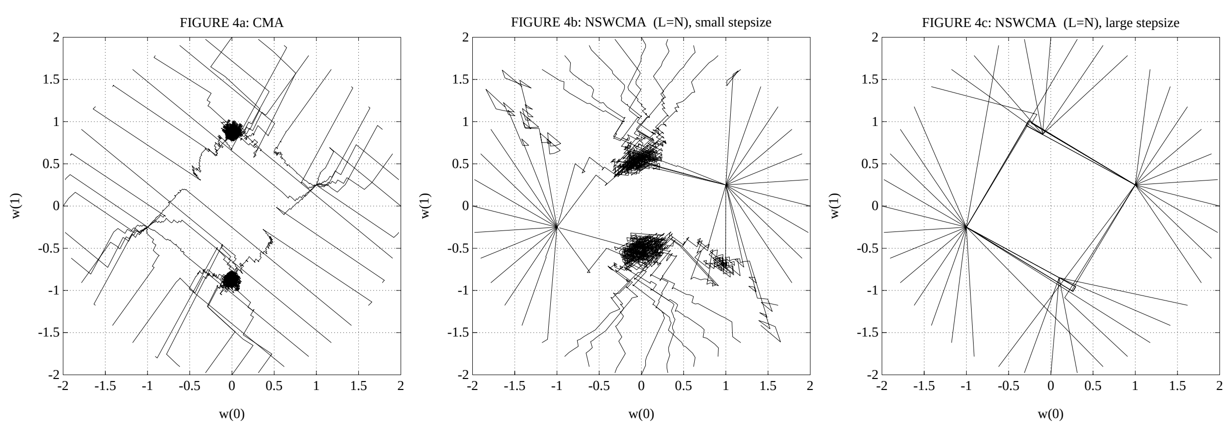

Figure 4: A comparison of CMA and NSWCMA (L=2) for an AP(1) channel.