13

BLIND ADAPTIVE FILTERING

13.1 INTRODUCTION

There are a number of applications where the reference signal is either not available or consists of a training signal that in communication systems implies in reduction of useful data transmission. In those cases, we should utilize some alternative objective functions applied to the available data as well as some knowledge related to the nature (properties) of the signals involved.

In this chapter, some adaptive-filtering algorithms are presented which do not utilize reference signal that are collectively known as blind adaptive-filtering algorithms. The algorithms are also called training-less or unsupervised algorithms since their learning do not include any reference or training signal. This chapter makes no attempt to cover this subject in breadth and in depth, but the interested reader can consult some books [1]-[4] for further details.

There are two main types of blind signal processing procedures widely discussed in the literature, namely blind source separation and blind deconvolution. In the former case several signal sources are mixed by an unknown environment and the objective of the blind signal processor is to separate these signal sources [2]. On the other hand, the blind deconvolution aims at removing the effect of a linear time-invariant system on a signal source where the only assumptions are the observation of the signal before the deconvolution process and the probability density of the input signal source.

Blind deconvolution is obviously closely related to blind equalization, and the distinction lies on the fact that in the equalization case it is usually assumed that the input signal belongs to a prescribed finite set (constellation) and the channel is a continuous-time channel. These features of the equalization setup are assets that can be exploited by allowing nonlinear channel equalization solutions, whereas blind deconvolution employs linear solutions because its input signal cannot be considered to belong to a finite set constellation. However, it is fact that several solutions for both problems are closely related and here we emphasize the blind equalization case.

P.S.R. Diniz, *Adaptive Filtering*, DOI: 10.1007/978-0-387-68606-6_13,

In blind equalization the channel model is either identified explicitly or implicitly. The algorithms utilizing as objective function the minimization of the MSE or generating a zero-forcing (ZF) solution[1] in general do not estimate the channel model explicitly. On the other hand, nonlinear solutions for channel equalization such as maximum likelihood sequence detector (MLSD) [8] and the DFE require explicit estimation of the channel model.

As a rule, the blind signal processing algorithms utilize second and higher order statistics indirectly or explicitly. The high-order statistics are directly employed in algorithms based on cummulants, see [9] for details, and they usually have slow convergence and high complexity. There is yet another class of algorithms based on models originated from information theory [3].

This chapter deals with blind algorithms utilizing high-order statistics implicitly for the single-input single-output (SISO) equalization case, e.g. constant modulus algorithm (CMA), and algorithms employing second-order statistics for the single-input multi-output (SIMO) equalization case. Unfortunately the SISO blind solutions have some drawbacks related to the multiple minima solutions, slow convergence, and difficulties in equalizing channels with nonminimum phase[2]. In the SIMO case we are usually dealing with oversampled received signal, that is, the received signal is sampled at rate multiple of the symbol rate (at least twice). Another SIMO situation is whenever we use multiple receive antennas that can be proved to be equivalent to oversampling. Such sampling higher than baud rate results in received signals which are cyclostationary allowing the extraction of phase information of the channel. In the case of baud rate sampling and WSS inputs, the received signal is also WSS and only minimum-phase channels can be identified from second-order statistics since the channel phase information is lost. Under certain assumptions the SIMO configuration allows the identification of the channel model as well as blind channel equalization utilizing only second-order statistics. In particular, this chapter presents the Godard, CM, and Sato algorithms for the SISO case. We also discuss some properties related to the error surface of the CMA. Then we derive the blind CM affine projection algorithm which is then applied to the SISO and SIMO setups.

13.2 CONSTANT-MODULUS ALGORITHM

In this section we present a family of blind adaptive-filtering algorithms that minimizes the distance between the modulus of the equalizer output and some prescribed constant values, without utilizing a reference signal. These constant values are related to the modulus of constellation symbols, denoted by $\mathcal{C}$, of typical modulations utilized in many digital communication systems. The earlier blind equalization proposals addressed the case of Pulse Amplitude Modulation (PAM) for the case the channel model is considered a linear time-invariant Single-Input Single-Output (SISO) system [5]-[6], operating at symbol rate. This approach was latter generalized in [7] by modifying the objective function to consider higher order statistics of the adaptive-filter output signal that accommodates the case of Quadrature Amplitude Modulation (QAM).

[1] In the ZF solution the equalized signal is forced to be equal to the transmitted signal, a solution not recommended whenever the environment noise is not negligible, due to noise enhancement. The ZF equalizer aims at estimating a channel inverse in order to eliminate intersymbol interference.

[2] Channels whose discrete-time models have poles and zeros outside the unit circle.

Let's assume here that symbols denoted by $s(k)$ are transmitted through a communication channel. The channel impulse response described by $h(k)$ convolves with the sequence $s(k)$ generating the received signal given by

$$x(k+J) = s(k)h(J) + \left(\sum_{l=-\infty,\, l\neq k}^{k+J} s(l)h(k+J-l) \right) + n(k+J) \tag{13.1}$$

where J denotes the channel time delay which will be considered zero without loss of generality. The transmitted signals $s(k)$ belong to a set of possible symbols, that is $s(k) \in \mathcal{C}$, with $\mathcal{C}$ representing the constellation set, defined by the chosen constellation such as PAM[3] and the complex QAM. The symbol occurrence is uniformly distributed over the defined elements of the constellation. In the following we present the Godard algorithm which relies on a high-order statistics property of the chosen constellation to define its updating mechanism.

13.2.1 Godard Algorithm

The general objective of the Godard algorithm utilizing the criterion proposed in [7] is to minimize

$$\begin{aligned}\xi_{\text{Godard}} &= E\left[(|\mathbf{w}^H(k)\mathbf{x}(k)|^q - r_q)^p\right] \\ &= E\left[(|y(k)|^q - r_q)^p\right] \\ &= E\left[e^p_{\text{Godard}}(k)\right]\end{aligned} \tag{13.2}$$

with

$$r_q = \frac{E[|s(k)|^{2q}]}{E[|s(k)|^q]} \tag{13.3}$$

where q and p are positive integers. The value of r_q defines the level which $|y(k)|^q$ should approach, with a penalization error powered by p.

The simple stochastic gradient version of this algorithm can be obtained by differentiating the objective function of equation (13.2) with respect to $\mathbf{w}^*(k)$. The resulting updating equation is given by

$$\begin{aligned}\mathbf{w}(k+1) &= \mathbf{w}(k) - \frac{1}{2}\,\mu\, p\, q\, (|y(k)|^q - r_q)^{p-1}\, |y(k)|^{q-2}\, y^*(k)\, \mathbf{x}(k) \\ &= \mathbf{w}(k) - \frac{1}{2}\,\mu\, p\, q\, e^{p-1}_{\text{Godard}}(k)\, |y(k)|^{q-2}\, y^*(k)\, \mathbf{x}(k)\end{aligned} \tag{13.4}$$

The detailed description of the Godard algorithm is provided by Algorithm 13.1.

[3] The M-ary PAM constellation points are represented by $s_i = \tilde{a}_i$, with $\tilde{a}_i = \pm\tilde{d}, \pm 3\tilde{d}, \ldots, \pm(\sqrt{M}-1)\tilde{d}$. The parameter $\tilde{d}$ represents half of the distance between two points in the constellation.

Algorithm 13.1

Godard Algorithm

Initialization
Choose p and q
$\mathbf{x}(0) = \mathbf{w}(0) = \text{random vectors}$
$r_q = \frac{E[|s(k)|^{2q}]}{E[|s(k)|^q]}$
Do for $k > 0$
$y(k) = \mathbf{w}^H(k)\mathbf{x}(k)$
$e_{\text{Godard}}(k) = |y(k)|^q - r_q$
$\mathbf{w}(k+1) = \mathbf{w}(k) - \frac{1}{2}\,\mu\, p\, q\, e_{\text{Godard}}^{p-1}(k)\,|y(k)|^{q-2}\, y^*(k)\,\mathbf{x}(k)$

13.2.2 Constant-Modulus Algorithm

For $q = p = 2$ in the Godard framework, the objective function of equation (13.2) corresponds to the constant-modulus algorithm (CMA) whose objective function is described by

$$\begin{aligned} E\left[e_{\text{CMA}}^2(k)\right] &= E\left[(|\mathbf{w}^H(k)\mathbf{x}(k)|^2 - r_2)^2\right] \\ &= E\left[(|y(k)|^2 - r_2)^2\right] \end{aligned} \tag{13.5}$$

In this case,

$$r_2 = \frac{E[|s(k)|^4]}{E[|s(k)|^2]} \tag{13.6}$$

meaning that whenever the input symbols have constant modulus, the CM error minimization aims at keeping the modulus $|y(k)|^2$ as close as possible to the constant value of r_2. For the CMA, the stochastic gradient update equation is given by

$$\begin{aligned} \mathbf{w}(k+1) &= \mathbf{w}(k) - 2\mu\,(|y(k)|^2 - r_2)\, y^*(k)\,\mathbf{x}(k) \\ &= \mathbf{w}(k) - 2\mu\, e_{\text{CMA}}(k)\, y^*(k)\,\mathbf{x}(k) \end{aligned} \tag{13.7}$$

Algorithm 13.2 describes in detail the CM algorithm.

13.2.3 Sato Algorithm

A historically important objective function somewhat related to the case of the Godard algorithm above is the so-called Sato algorithm whose objective function is defined as

$$e_{\text{Sato}}(k) = y(k) - \text{sgn}[y(k)]r_1 \tag{13.8}$$

Algorithm 13.2

Constant-Modulus Algorithm

Initialization
$\mathbf{x}(0) = \mathbf{w}(0) = \text{random vectors}$
$r_2 = \frac{E[|s(k)|^4]}{E[|s(k)|^2]}$
Do for $k \geq 0$
$y(k) = \mathbf{w}^H(k)\mathbf{x}(k)$
$e_{\text{CMA}}(k) = |y(k)|^2 - r_2$
$\mathbf{w}(k+1) = \mathbf{w}(k) - 2\mu\, e_{\text{CMA}}(k)\, y^*(k)\, \mathbf{x}(k)$

Algorithm 13.3

Sato Algorithm

Initialization
$\mathbf{x}(0) = \mathbf{w}(0) = \text{random vectors}$
$r_1 = \frac{E[|s(k)|^2]}{E[|s(k)|]}$
Do for $k \geq 0$
$y(k) = \mathbf{w}^H(k)\mathbf{x}(k)$
$e_{\text{Sato}}(k) = y(k) - \text{sgn}[y(k)]r_1$
$\mathbf{w}(k+1) = \mathbf{w}(k) - \mu\, e_{\text{Sato}}(k)\, \mathbf{x}(k)$

where $\text{sgn}[y] = \frac{y}{|y|}$ such that for $y = 0$, $\text{sgn}[y] = 1$. Its update equation is described by

$$\begin{aligned}\mathbf{w}(k+1) &= \mathbf{w}(k) - \mu\,(y(k) - \text{sgn}[y(k)]r_1)\,\mathbf{x}(k)\\ &= \mathbf{w}(k) - \mu\, e_{\text{Sato}}(k)\, \mathbf{x}(k)\end{aligned} \tag{13.9}$$

In this case, the target is that the equalized signal $y(k)$ follows the sign of the transmitted symbol, that is, this algorithm follows the decision direction whenever the input signal is a binary PAM signal. The Sato algorithm was the first blind adaptive equalizer taking into consideration PAM transmission signals with multilevel. Algorithm 13.3 describes step by step the Sato algorithm.

13.2.4 Error Surface of CMA

In this subsection we derive an expression for the CMA error surface for a simple and yet illustrative case, where both the symbol constellation as well as the adaptive-filter coefficients are real valued. Let's assume the simplest equalization problem where the unknown channel is modeled as

$$H(z) = \frac{\kappa z}{z + a} \tag{13.10}$$

In a noiseless environment this channel has an ideal equalizer (zero forcing) given by

$$\begin{aligned} W(z) &= \pm z^{-i}\left(w_0 + w_1 z^{-1}\right) \\ &= \pm \frac{z^{-i}}{\kappa}[1 + az^{-1}] \end{aligned} \tag{13.11}$$

where i is a nonnegative integer. For $i = 0$ it leads to an equalized signal with zero delay. For the CMA case, the objective function in this particular example can be written as

$$\begin{aligned} \xi_{\text{CMA}} &= E\left\{[|y(k)|^2 - r_2]^2\right\} \\ &= E[|y(k)|^4] - 2E[|y(k)|^2]r_2 + r_2^2 \end{aligned} \tag{13.12}$$

The required expected values for the above equation are given by

$$E[|y(k)|^2] = (w_0^2 + w_1^2)\frac{\kappa^2 E[|s(k)|^2]}{1 - a^2} - 2w_0 w_1 \frac{a\kappa^2 E[|s(k)|^2]}{1 - a^2} \tag{13.13}$$

$$\begin{aligned} E[|y(k)|^4] = {}& (w_0^4 + w_1^4)\left[\frac{\kappa^4 E[|s(k)|^4]}{1 - a^4} + \frac{6a^2\kappa^4\{E[|s(k)|^2]\}^2}{(1 - a^4)(1 - a^2)}\right] \\ &+6w_0^2 w_1^2\left\{a^2\left[\frac{\kappa^4 E[|s(k)|^4]}{1 - a^4} + \frac{6a^2\kappa^4\{E[|s(k)|^2]\}^2}{(1 - a^4)(1 - a^2)}\right] + \frac{\kappa^2\{E[|s(k)|^2]\}^2}{1 - a^2}\right\} \\ &-4w_0 w_1^3 a\left\{\left[\frac{\kappa^4 E[|s(k)|^4]}{1 - a^4}\right] + \frac{6a^2\kappa^4\{E[|s(k)|^2]\}^2}{(1 - a^4)(1 - a^2)}\right\} \\ &-4w_0^3 w_1\left\{a^3\left[\frac{\kappa^4 E[|s(k)|^4]}{1 - a^4} + \frac{6a^2\kappa^4\{E[|s(k)|^2]\}^2}{(1 - a^4)(1 - a^2)}\right] + \frac{3a\kappa^4\{E[|s(k)|^2]\}^2}{1 - a^2}\right\} \end{aligned} \tag{13.14}$$

where the detailed derivations pertaining to the above equations can be found in problem 2.

Example 13.1

Assume a QAM signal with four symbols is transmitted through an AR channel whose transfer function is

$$H(z) = \frac{0.36z}{z + a}$$

for the cases where $a = 0.4$ and $a = 0.8$, respectively. The equalizer is a first-order FIR adaptive filter as described in equation (13.11). For a signal to noise ratio of 10dB, plot the CMA error surface and its corresponding contours.

Solution:

Fig. 13.1 depicts the error surface and its contours for the CM objective function, with $a = 0.4$, where the surface is flattened for certain ranges of w_0 and w_1 in order to allow a better view of valleys and local minima and maxima. As can be verified the surface presents multiple minima, the ones at $w_0 = 0$ do not correspond to global minima. The surface shape indicates that if a good initial point is not given to a CM-based algorithm, the parameters will converge to an undesirable local minima where the equalization performance might be very poor. In addition, if the algorithm traverses a region in the neighborhood of a saddle point the convergence of stochastic gradient algorithms can be particularly slow. Fig. 13.2 shows the error surface and its contours for $a = 0.8$, where in this case the local minima are not so visible but they do exist.

□

Example 13.2

In this example we consider an equalization problem. Perform the equalization of a channel with the following impulse response

$$\mathbf{h} = [1.1 + \jmath 0.5 \;\; 0.1 - \jmath 0.3 \;\; -0.2 - \jmath 0.1]^T$$

The transmitted signals are uniformly distributed four QAM samples with unitary power. An additional Gaussian white noise with variance $10^{-2.5}$ is present at the channel output. Utilize the CMA.

(a) Find the Wiener solution for an equalizer with five coefficients and convolve with the channel impulse response.

(b) Perform a blind equalization also with five coefficients and depict the detected symbols before and after the equalization.

Solution:

(a) In the first step, we compute the Wiener solution and perform the convolution with the channel impulse response in order to verify the effectiveness of the equalizer order in the present example. For a delay of 1, the convolution samples are given by

$$\mathbf{y} = \begin{bmatrix} 0.0052 + \jmath 0.0104 \\ 0.9675 + \jmath 0.0000 \\ 0.0074 + \jmath 0.0028 \\ -0.0548 - \jmath 0.0014 \\ 0.0129 + \jmath 0.0222 \\ -0.0939 - \jmath 0.0075 \\ 0.0328 - \jmath 0.0098 \end{bmatrix}^T$$

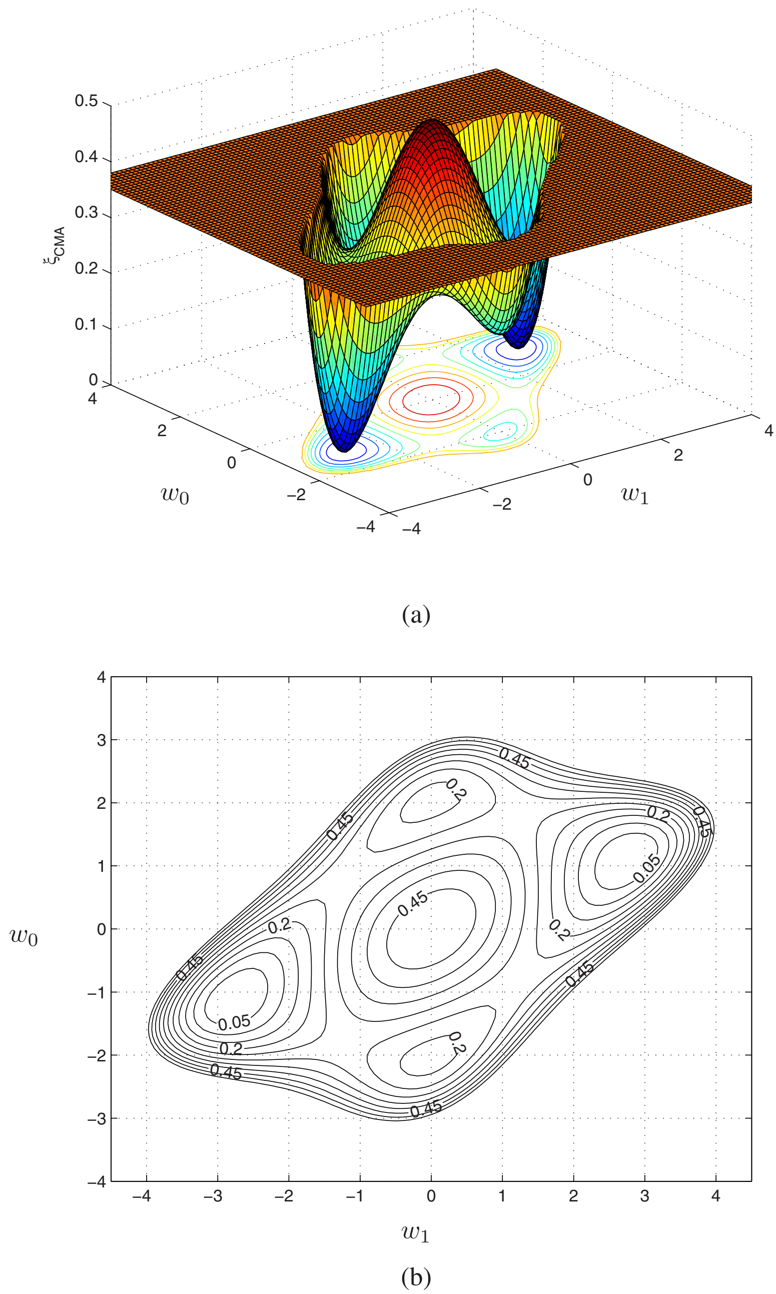
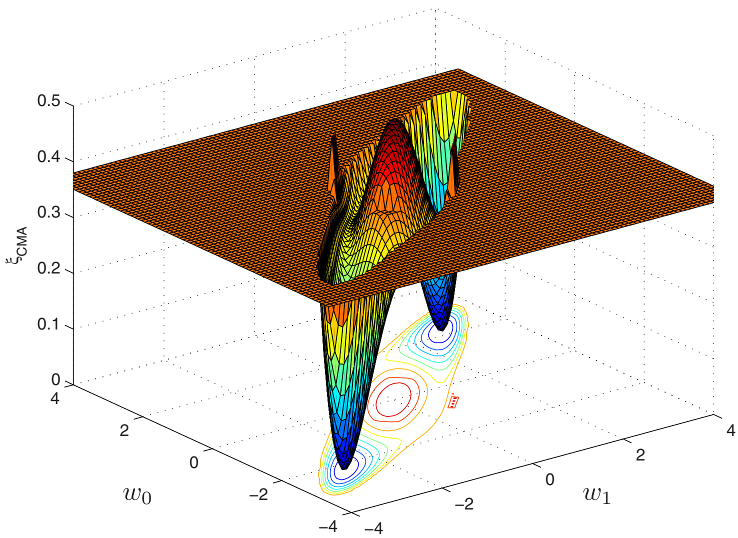
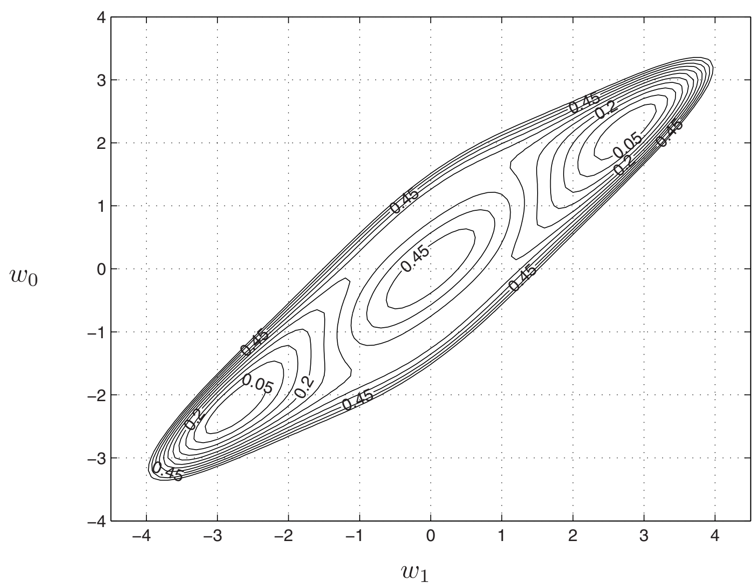


Figure 13.1 (a) CMA error surface, (b) CMA contours; a=0.4.



(a)



(b)

Figure 13.2 (a) CMA error surface, (b) CMA contours; a=0.8.

where as can be observed the real part of the second sample is much higher than the remaining samples, showing that the equalization is successful.

(b) In Fig. 13.3 it is shown how the received signals are distributed in the input signal constellation space, and as can be observed and expected the received signal requires an equalizer for proper detection.

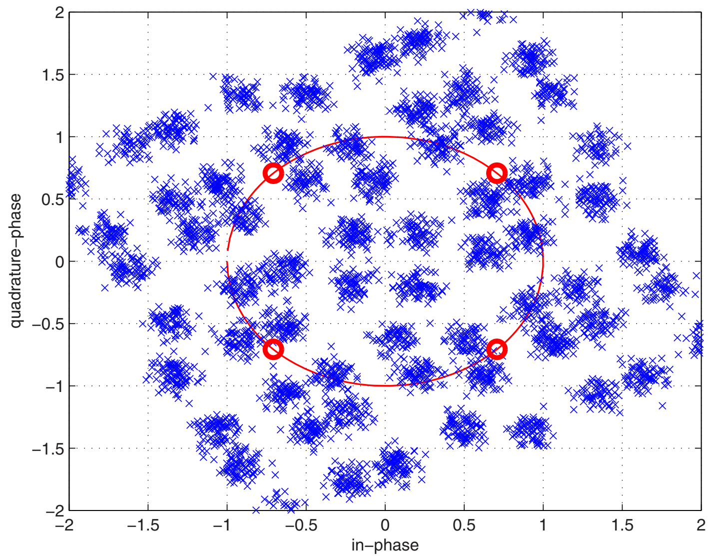


Figure 13.3 Receiver signals before equalization.

By applying the CMA to solve the equalization problem with $\mu = 0.001$, we run the algorithm for 10000 iterations with the results measured by averaging the outcomes of 200 independent runs. By initializing the adaptive-filter coefficients at

$$\mathbf{w}(0) = \begin{bmatrix} -1.627563 - \jmath 0.443856 \\ -0.121194 + \jmath 0.338364 \\ 0.189390 + \jmath 0.063311 \\ 0.575142 - \jmath 0.062878 \\ 0.364852 - \jmath 0.6053977 \end{bmatrix}$$

the last 1000 equalized signals fall in the regions depicted in Fig. 13.4 representing the input signal constellation space. As can be verified, the equalized symbols present four clusters which are not centered at the actual transmitted symbols positions. On the other hand, these clusters are around the same constant modulus position as the transmitted symbols but at different angles, that is, the transmitted constellation is received after equalization rotated by an arbitrary angle. For differentially encoded symbols the mentioned phase shift can be eliminated, allowing proper decoding of the received symbols.

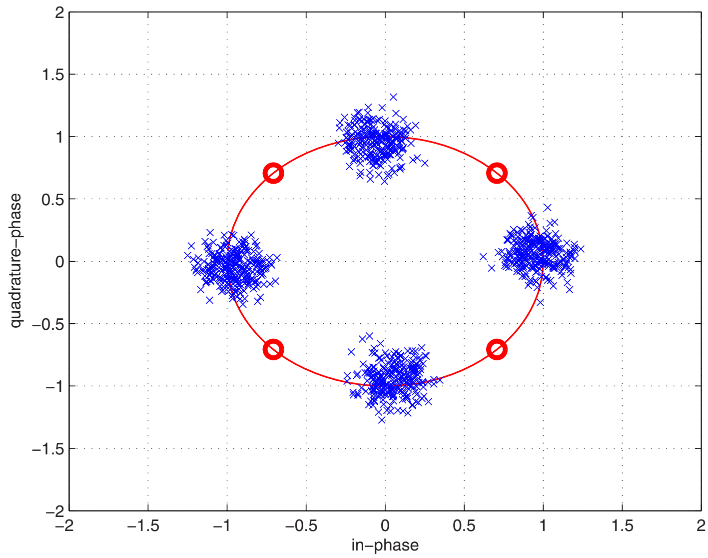


Figure 13.4 Equalized signals for the CM algorithm using the first coefficient initialization.

If the CMA filter coefficients are initialized at

$$\mathbf{w}(0) = \begin{bmatrix} 2.011934 + \jmath 0.157299 \\ 0.281061 + \jmath 0.324327 \\ -0.017917 + \jmath 0.836021 \\ -0.391982 + \jmath 1.144051 \\ -0.185579 - \jmath 0.898060 \end{bmatrix}$$

the resulting clusters are shown in Fig. 13.5, where it is possible to verify that in this case the clusters occur at the right positions with respect to the transmitted symbols.

For illustration, Fig. 13.6 shows the equalization results when using the Wiener solution, where it can be observed by comparing it with Fig. 13.5 that the CMA can lead to Wiener like solutions when properly initialized.

The typical learning curve for the CM algorithm in the present example is illustrated in 13.7 where in this case we utilized random initial coefficients for the adaptive filter.

□

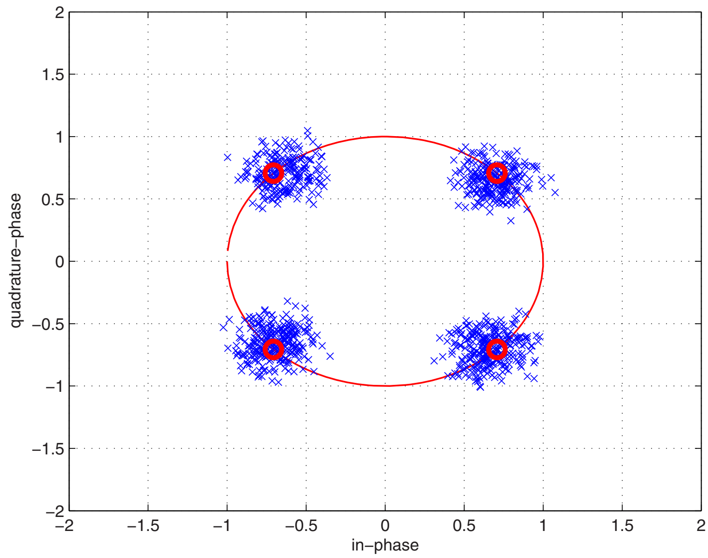


Figure 13.5 Equalized signals for the CM algorithm using the second coefficient initialization.

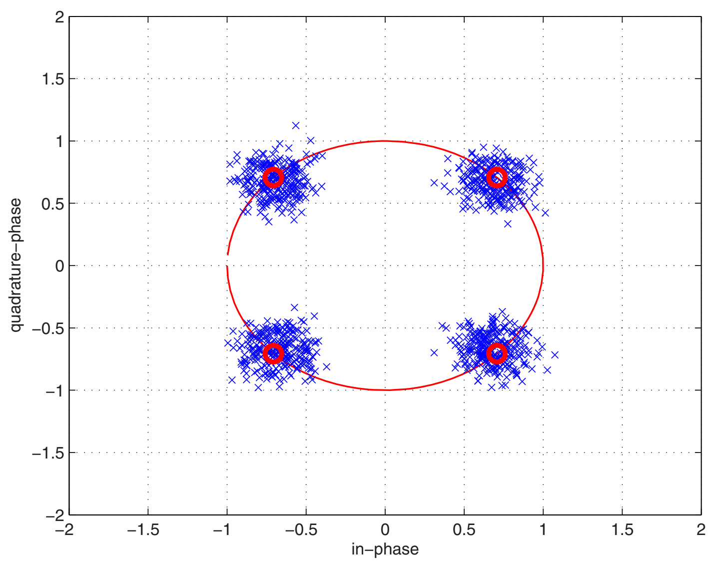


Figure 13.6 Equalized signals for the Wiener filter.

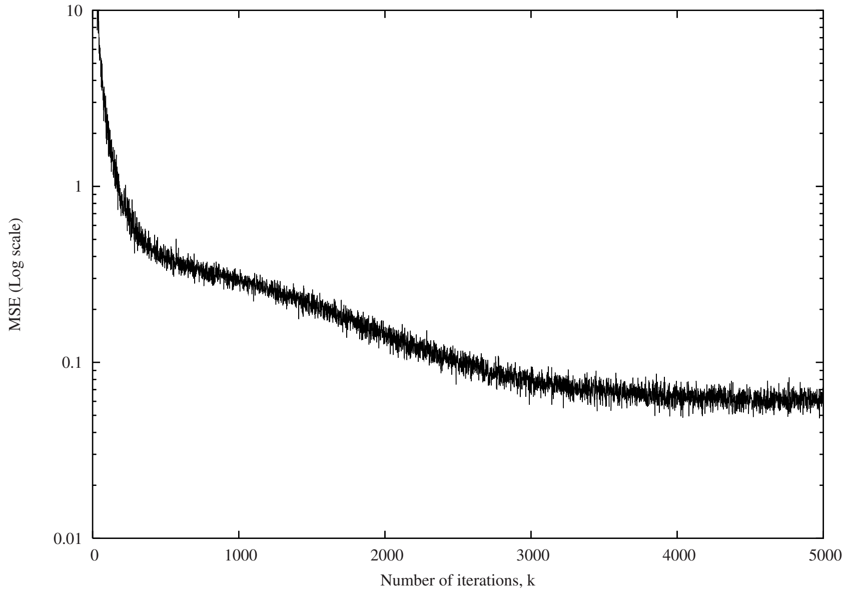


Figure 13.7 Learning curve of the CM algorithm.

13.3 AFFINE PROJECTION CM ALGORITHM

In general the CMA like algorithms present slow convergence when the update equation has a stochastic gradient form. A possible alternative solution when the convergence speed is not acceptable is to utilize the affine projection form. Let's consider the cases where either the desired vector is a CMA like function at each entry of a vector $\mathbf{r}_{\text{ap}}(k)$ or represents a nonlinear function $G_1[\cdot]$ applied to the adaptive-filter output, that is,

$$\mathbf{r}_{\text{ap}}(k) = G_1\left[\mathbf{y}_{\text{ap}}(k)\right] = G_1\left[\mathbf{X}_{\text{ap}}^T(k)\mathbf{w}^*(k)\right] \tag{13.15}$$

where the definitions of the data matrix and vectors of the affine projection algorithm are defined in equations (4.74) and (4.77).

The objective function that the affine projection algorithm minimizes in this case is

$$\begin{aligned}&\|\mathbf{w}(k+1) - \mathbf{w}(k)\|^2\\&\text{subject to :}\\&G_2\left\{\mathbf{r}_{\text{ap}}(k) - \mathbf{X}_{\text{ap}}^T(k)\mathbf{w}^*(k+1)\right\} = \mathbf{0}\end{aligned} \tag{13.16}$$

where $\mathbf{r}_{\text{ap}}(k)$ is a vector replacing $\mathbf{d}_{\text{ap}}(k)$ in the blind formulation whose elements are determined by the type of blind objective function at hand. $G_2[\cdot]$ represents another nonlinear operation applied elementwise on $[\cdot]$, usually given by $(\cdot)^2$ as in the CM algorithm. In any situation, $G_2(\mathbf{0}) = \mathbf{0}$. Also in this case the affine projection algorithm keeps the next coefficient vector $\mathbf{w}(k+1)$ as close as possible to the current one and aims at making the *a posteriori* error to be zero. It is worth mentioning that if the minimization of $\|\mathbf{w}(k+1)-\mathbf{w}(k)\|^2$ is not included in the objective function, the problem of keeping $\mathbf{r}_{\mathrm{ap}}(k)=\mathbf{X}_{\mathrm{ap}}^T(k)\mathbf{w}^*(k+1)$ makes the coefficient vector underdetermined[4] whenever this vector has more than one entry.

As described in Chapter 4 by utilizing the method of Lagrange multipliers the constrained minimization problem of equation (13.16) becomes

$$F[\mathbf{w}(k+1)] = \|\mathbf{w}(k+1)-\mathbf{w}(k)\|^2 + \boldsymbol{\lambda}_{\mathrm{ap}}^H(k) G_2\left\{\mathbf{r}_{\mathrm{ap}}(k)-\mathbf{X}_{\mathrm{ap}}^T(k)\mathbf{w}^*(k+1)\right\} \tag{13.17}$$

where $\boldsymbol{\lambda}_{\mathrm{ap}}(k)$ is the $(L+1)\times 1$ vector of Lagrange multipliers. In order to facilitate the gradient computation let's rewrite the above expression as

$$\begin{aligned} F[\mathbf{w}(k+1)] &= [\mathbf{w}(k+1)-\mathbf{w}(k)]^H[\mathbf{w}(k+1)-\mathbf{w}(k)] \\ &\quad + G_2\left\{\mathbf{r}_{\mathrm{ap}}^T(k)-\mathbf{w}^H(k+1)\mathbf{X}_{\mathrm{ap}}(k)\right\}\boldsymbol{\lambda}_{\mathrm{ap}}^*(k) \end{aligned} \tag{13.18}$$

The gradient of $F[\mathbf{w}(k+1)]$ with respect to $\mathbf{w}^*(k+1)$[5] is given by

$$\begin{aligned} \mathbf{g}_{\mathbf{w}^*}\{F[\mathbf{w}(k+1)]\} &= [\mathbf{w}(k+1)-\mathbf{w}(k)] \\ &\quad + \mathbf{X}_{\mathrm{ap}}(k)\mathbf{g}_{\bar{\mathbf{y}}_{\mathrm{ap}}}\left\{G_2\left[\mathbf{r}_{\mathrm{ap}}^T(k)-\bar{\mathbf{y}}_{\mathrm{ap}}^T(k)\right]\right\}\boldsymbol{\lambda}_{\mathrm{ap}}^*(k) \end{aligned} \tag{13.19}$$

where $\bar{\mathbf{y}}_{\mathrm{ap}}(k)$ represents the *a posteriori* adaptive-filter output signal. After setting the gradient of $F[\mathbf{w}(k+1)]$ with respect to $\mathbf{w}^*(k+1)$ equal to zero, we get

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \mathbf{X}_{\mathrm{ap}}(k)\mathbf{g}_{\bar{\mathbf{y}}_{\mathrm{ap}}}\left\{G_2\left[\mathbf{r}_{\mathrm{ap}}^T(k)-\bar{\mathbf{y}}_{\mathrm{ap}}^T(k)\right]\right\}\boldsymbol{\lambda}_{\mathrm{ap}}^*(k) \tag{13.20}$$

By premultiplying equation (13.20) by $\mathbf{X}_{\mathrm{ap}}^H(k)$, using the constraint relation of equation (13.16), and considering the fact that $G_2(\mathbf{0})=\mathbf{0}$ so that $\mathbf{X}_{\mathrm{ap}}^H(k)\mathbf{w}(k+1)=\mathbf{r}_{\mathrm{ap}}^*(k)$, we obtain

$$-\mathbf{X}_{\mathrm{ap}}^H(k)\mathbf{X}_{\mathrm{ap}}(k)\mathbf{g}_{\bar{\mathbf{y}}_{\mathrm{ap}}}\left\{G_2\left[\mathbf{r}_{\mathrm{ap}}^T(k)-\bar{\mathbf{y}}_{\mathrm{ap}}^T(k)\right]\right\}\boldsymbol{\lambda}_{\mathrm{ap}}^*(k) + \mathbf{X}_{\mathrm{ap}}^H(k)\mathbf{w}(k) = \mathbf{r}_{\mathrm{ap}}^*(k) \tag{13.21}$$

This expression leads to

$$\mathbf{g}_{\bar{\mathbf{y}}_{\mathrm{ap}}}\left\{G_2\left[\mathbf{r}_{\mathrm{ap}}^T(k)-\bar{\mathbf{y}}_{\mathrm{ap}}^T(k)\right]\right\}\boldsymbol{\lambda}_{\mathrm{ap}}^*(k) = \left[\mathbf{X}_{\mathrm{ap}}^H(k)\mathbf{X}_{\mathrm{ap}}(k)\right]^{-1}\left\{-\mathbf{r}_{\mathrm{ap}}^*(k)+\mathbf{X}_{\mathrm{ap}}^H(k)\mathbf{w}(k)\right\} \tag{13.22}$$

By substituting equation (13.22) in equation (13.20), the update equation can be rewritten as

$$\begin{aligned} \mathbf{w}(k+1) &= \mathbf{w}(k) + \mathbf{X}_{\mathrm{ap}}(k)\left(\mathbf{X}_{\mathrm{ap}}^H(k)\mathbf{X}_{\mathrm{ap}}(k)\right)^{-1}\left\{\mathbf{r}_{\mathrm{ap}}^*(k)-\mathbf{X}_{\mathrm{ap}}^H(k)\mathbf{w}(k)\right\} \\ &= \mathbf{w}(k) + \mathbf{X}_{\mathrm{ap}}(k)\left(\mathbf{X}_{\mathrm{ap}}^H(k)\mathbf{X}_{\mathrm{ap}}(k)\right)^{-1}\mathbf{e}_{\mathrm{ap}}^*(k) \end{aligned} \tag{13.23}$$

From the above equation it follows that

$$\|\mathbf{w}(k+1)-\mathbf{w}(k)\|^2 = \mathbf{e}_{\mathrm{ap}}^T(k)\left(\mathbf{X}_{\mathrm{ap}}^H(k)\mathbf{X}_{\mathrm{ap}}(k)\right)^{-1}\mathbf{e}_{\mathrm{ap}}^*(k) \tag{13.24}$$

[4] A solution exists but it is not unique.

[5] We could also formulate this solution employing the gradient with respect to $\mathbf{w}(k+1)$, leading to the same results.

such that the minimization of the terms on the left- and right-hand sides are equivalent. However, the minimization of the right-hand side term does not mean minimizing $\|\mathbf{e}^*_{\rm ap}(k)\|$ unless the matrix $\left(\mathbf{X}^H_{\rm ap}(k)\mathbf{X}_{\rm ap}(k)\right)^{-1}$ is a diagonal matrix with equal nonzero values in the main diagonal. Despite of that, in order to generate a tractable solution we minimize $\|\mathbf{e}^*_{\rm ap}(k)\|$ and interpret the objective function that is actually minimized.

If we assume $\mathbf{r}^*_{\rm ap}(k)$ has constant modulus elementwise, the minimization of

$$\|\mathbf{e}^*_{\rm ap}(k)\|^2 = \|\mathbf{r}^*_{\rm ap}(k) - \mathbf{X}^H_{\rm ap}(k)\mathbf{w}(k)\|^2$$

occurs when $\mathbf{r}^*_{\rm ap}(k)$ is in the same direction as (is colinear with) $\mathbf{X}^H_{\rm ap}(k)\mathbf{w}(k)$. In this case the following choice should be made

$$\mathbf{r}^*_{\rm ap}(k) = \text{sgn}[\mathbf{X}^H_{\rm ap}(k)\mathbf{w}(k)] \tag{13.25}$$

where for a complex number y, $\text{sgn}[y] = \frac{y}{|y|}$, and whenever $y = 0$, $\text{sgn}[y] = 1$.

In the update equation (13.24) the convergence factor is unity, and as previously discussed a trade-off between final misadjustment and convergence speed is achieved by including convergence factor as follows

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu\mathbf{X}_{\rm ap}(k)\left(\mathbf{X}^H_{\rm ap}(k)\mathbf{X}_{\rm ap}(k)\right)^{-1}\left\{\mathbf{r}^*_{\rm ap}(k) - \mathbf{X}^H_{\rm ap}(k)\mathbf{w}(k)\right\} \tag{13.26}$$

As before, with a convergence factor different from one (smaller than one) *a posteriori* error is no longer zero. The reader might question why $G_2[\cdot]$ did not appear in the final update expression of equation (13.22), the reason is the assumption that the constraint in equation (13.16) is satisfied exactly leading to a zero *a posteriori* error.

The objective function that equation (13.26) actually minimizes is given by

$$\begin{aligned}&\left(\frac{1}{\mu} - 1\right)\|\mathbf{w}(k+1) - \mathbf{w}(k)\|^2 + \|\mathbf{r}_{\rm ap}(k) - \mathbf{X}^T_{\rm ap}(k)\mathbf{w}^*(k+1)\|^2_{\mathbf{P}} = \\ &\left(\frac{1}{\mu} - 1\right)\|\mathbf{w}(k+1) - \mathbf{w}(k)\|^2 + \|\text{sgn}[\mathbf{X}^H_{\rm ap}(k)\mathbf{w}(k)] - \mathbf{X}^T_{\rm ap}(k)\mathbf{w}^*(k+1)\|^2_{\mathbf{P}}\end{aligned} \tag{13.27}$$

where $\mathbf{P} = \left(\mathbf{X}^H_{\rm ap}(k)\mathbf{X}_{\rm ap}(k)\right)^{-1}$ and $\|\mathbf{a}\|^2_{\mathbf{P}} = \mathbf{a}^H\mathbf{P}\mathbf{a}$.

Proof:

In order to simplify the derivations let's define

$$\alpha = \left(\frac{1}{\mu} - 1\right)$$

The objective function to be minimized with respect to the coefficients $\mathbf{w}^*(k+1)$ is given by

$$\xi(k) = \alpha\|\mathbf{w}(k+1) - \mathbf{w}(k)\|^2 + \|\mathbf{r}_{\rm ap}(k) - \mathbf{X}^T_{\rm ap}(k)\mathbf{w}^*(k+1)\|^2_{\mathbf{P}}$$

The derivative of the objective function is then given by

$$\frac{\partial \xi(k)}{\partial \mathbf{w}^*(k+1)} = \alpha[\mathbf{w}(k+1) - \mathbf{w}(k)] - \mathbf{X}_{\text{ap}}(k)\mathbf{P}\left[\mathbf{r}^*_{\text{ap}}(k) - \mathbf{X}^H_{\text{ap}}(k)\mathbf{w}(k+1)\right]$$

By setting this result to zero it follows that

$$\left[\alpha\mathbf{I} + \mathbf{X}_{\text{ap}}(k)\mathbf{P}\mathbf{X}^H_{\text{ap}}(k)\right]\mathbf{w}(k+1) = \alpha\mathbf{w}(k) + \mathbf{X}_{\text{ap}}(k)\mathbf{P}\mathbf{r}^*_{\text{ap}}(k) \tag{13.28}$$

By applying the matrix inversion lemma we obtain

$$\begin{aligned}
\left[\alpha\mathbf{I} + \mathbf{X}_{\text{ap}}(k)\mathbf{P}\mathbf{X}^H_{\text{ap}}(k)\right]^{-1} &= \frac{1}{\alpha}\mathbf{I} - \frac{1}{\alpha}\mathbf{I}\mathbf{X}_{\text{ap}}(k)\left[\mathbf{X}^H_{\text{ap}}(k)\frac{1}{\alpha}\mathbf{I}\mathbf{X}_{\text{ap}}(k) + \mathbf{P}^{-1}\right]^{-1}\mathbf{X}^H_{\text{ap}}(k)\frac{1}{\alpha}\mathbf{I} \\
&= \frac{1}{\alpha}\mathbf{I} - \frac{1}{\alpha}\mathbf{I}\mathbf{X}_{\text{ap}}(k)\left[\frac{\mathbf{P}^{-1}}{\alpha} + \mathbf{P}^{-1}\right]^{-1}\mathbf{X}^H_{\text{ap}}(k)\frac{1}{\alpha}\mathbf{I} \\
&= \frac{1}{\alpha}\left[\mathbf{I} - \mathbf{X}_{\text{ap}}(k)\frac{\alpha}{1+\alpha}\mathbf{P}\mathbf{X}^H_{\text{ap}}(k)\frac{1}{\alpha}\mathbf{I}\right] \\
&= \frac{1}{\alpha}\left[\mathbf{I} - \frac{\mathbf{X}_{\text{ap}}(k)\mathbf{P}\mathbf{X}^H_{\text{ap}}(k)}{1+\alpha}\right]
\end{aligned}$$

By replacing the last expression in the updating equation (13.28), we obtain

$$\begin{aligned}
\mathbf{w}(k+1) &= \left[\mathbf{I} - \frac{\mathbf{X}_{\text{ap}}(k)\mathbf{P}\mathbf{X}^H_{\text{ap}}(k)}{1+\alpha}\right]\mathbf{w}(k) + \frac{1}{\alpha}\left[\mathbf{I} - \frac{\mathbf{X}_{\text{ap}}(k)\mathbf{P}\mathbf{X}^H_{\text{ap}}(k)}{1+\alpha}\right]\mathbf{X}_{\text{ap}}(k)\mathbf{P}\mathbf{r}^*_{\text{ap}}(k) \\
&= \mathbf{w}(k) - \frac{\mathbf{X}_{\text{ap}}(k)\mathbf{P}\mathbf{y}^*_{\text{ap}}(k)}{1+\alpha} + \frac{1}{\alpha}\mathbf{X}_{\text{ap}}(k)\mathbf{P}\mathbf{r}^*_{\text{ap}}(k) - \frac{1}{\alpha}\frac{\mathbf{X}_{\text{ap}}(k)\mathbf{P}\mathbf{r}^*_{\text{ap}}(k)}{1+\alpha} \\
&= \mathbf{w}(k) - \mu\mathbf{X}_{\text{ap}}(k)\mathbf{P}\mathbf{y}^*_{\text{ap}}(k) + \mu\mathbf{X}_{\text{ap}}(k)\mathbf{P}\mathbf{r}^*_{\text{ap}}(k) \\
&= \mathbf{w}(k) + \mu\mathbf{X}_{\text{ap}}(k)\left(\mathbf{X}^H_{\text{ap}}(k)\mathbf{X}_{\text{ap}}(k)\right)^{-1}\mathbf{e}^*_{\text{ap}}(k)
\end{aligned}$$

□

The description of the affine projection CM algorithm is provided in Algorithm 13.4, where as standard an identity matrix multiplied by a small constant was added to the matrix $\mathbf{X}^H_{\text{ap}}(k)\mathbf{X}_{\text{ap}}(k)$ in order to avoid numerical problems in the matrix inversion.

It is worth mentioning that the update equation (13.22) represents other important application such as the case where $\mathbf{r}^*_{\text{ap}}(k) = \text{dec}[\mathbf{X}^H_{\text{ap}}(k)\mathbf{w}(k)]$, which corresponds to a decision directed blind algorithm, where $\text{dec}[\cdot]$ represents a hard limiter where each entry of its argument is mapped into the closest symbol of the constellation used in the transmission [10].

Now let's consider the special scalar case where the nonlinear operations to be applied to the output error of the normalized LMS algorithm are as following described. The objective function to be

Algorithm 13.4

The Affine Projection CM Algorithm

Initialization
$\mathbf{x}(0) = \mathbf{w}(0) = \text{random vectors}$
choose μ in the range $0 < \mu \leq 2$
$\gamma =$ small constant
Do for $k > 0$
$\mathbf{y}^*_{\text{ap}}(k) = \mathbf{X}^H_{\text{ap}}(k)\mathbf{w}(k)$
$\mathbf{r}^*_{\text{ap}}(k) = \text{sgn}[\mathbf{X}^H_{\text{ap}}(k)\mathbf{w}(k)]$
$\mathbf{e}^*_{\text{ap}}(k) = \mathbf{r}^*_{\text{ap}}(k) - \mathbf{y}^*_{\text{ap}}(k)$
$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu\mathbf{X}_{\text{ap}}(k)\left(\mathbf{X}^H_{\text{ap}}(k)\mathbf{X}_{\text{ap}}(k) + \gamma\mathbf{I}\right)^{-1}\mathbf{e}^*_{\text{ap}}(k)$

minimized is

$$\begin{aligned}
&\|\mathbf{w}(k+1) - \mathbf{w}(k)\|^2 \\
&\text{subject to :} \\
&|1 - |\mathbf{x}^H(k)\mathbf{w}(k+1)|^q|^p = 0
\end{aligned} \tag{13.29}$$

The resulting update equation is

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu\mathbf{x}(k)\left(\mathbf{x}^H(k)\mathbf{x}(k)\right)^{-1}\left\{\text{sgn}\left[\mathbf{x}^H(k)\mathbf{w}(k)\right] - \mathbf{x}^H(k)\mathbf{w}(k)\right\} \tag{13.30}$$

corresponding to a scalar normalized LMS CM algorithm.

Example 13.3

Repeat Example 13.2 for the case of the affine projection CM algorithm, for $L = 1$ and $L = 3$ and compare the result with the CM algorithm with $q = 2$.

Solution:

Using $\mu = 0.001$ and the CM algorithm, the equalizer took well over 1000 iterations to converge as depicted in Fig. 13.8. The same figure shows that the affine projection CM algorithm with $L = 3$ has the fastest convergence, around 100 iterations, while leading to higher MSE after convergence when compared with the cases of $L = 1$ and the CMA. For the affine projection cases the convergence factor is $\mu = 0.1$. Fig. 13.9 depicts the equalized signals after convergence for the case where $L = 3$. All these figures were generated by averaging the outcomes of 50 independent runs.

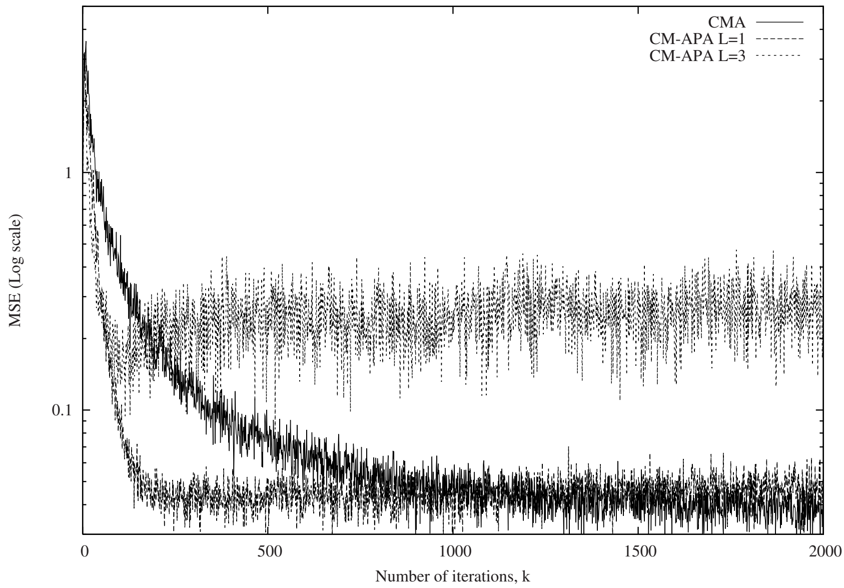


Figure 13.8 Learning curves for the CM and affine projection CM algorithms, with $L = 1$ and $L = 3$.

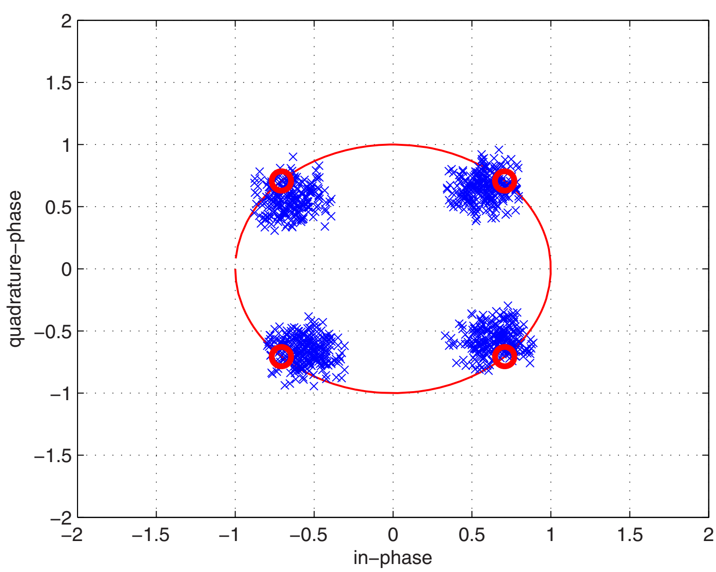


Figure 13.9 Equalized signals for the affine projection CM algorithm, with $L = 3$.

□

13.4 BLIND SIMO EQUALIZERS

The symbol spaced blind CMA equalizer methods described in previous section may converge to unacceptable local minima induced by the finite-length of the FIR equalizers, despite these minima being correct whenever the equalizer is a double sided filter with infinite order [1]. This situation changes favorably in the case a fractionally spaced equalizer is employed as following discussed. Many of the early blind equalizer methods utilized SISO channel model and relied on high-order (greater than second-order) statistics which lead to multiple minima and slow convergence. These equalizers are more sensitive to noise than those using second-order statistics. On the other hand, the availability of multiple measures of the received signal gives rise to SIMO configuration that in turn allows for blind channel equalization using second-order statistics. For example, oversampling the channel output signal by an integer factor l leads to a cyclostationary process with period l, such that the received discrete signal has cyclic correlation function allowing, under certain conditions, the identification of the channel modulus and phase [1] blindly. The SIMO configuration can be obtained by exploring diversity of antennas or by oversampling (also known as fractionally sampling) the received signal.

It is worth mentioning that the SIMO methods are not only useful to estimate a SIMO channel inverse filter but can be also used to perform channel identification. Many identification and equalization approaches can be constructed from the observed data such as subspace methods [11] and prediction methods [12]-[14] among others. The subspace methods are in general computationally complex. Furthermore they are sensitive to the channel order uncertainty causing dimension errors in the constructed signal and noise subspaces. Prediction error methods (PEM) are robust to overmodeling [15] and lend themselves to adaptive implementations.

These SIMO approaches can be extended in a rather straightforward way to device CDMA receivers [21] where blind multiuser detections are required [22]-[28], and in the cases semi-blind solutions are possible [29]. In addition, in multiple transmitter and receiver antennas systems several types of blind MIMO receivers can be derived [30]-[33]. In this section we briefly introduce the formulation for SIMO blind equalization [1], [16], and point out how this formulation brings useful solutions to blind equalization problems.

Let's consider the single-input I-output linear system model depicted in Fig. 13.10, representing an oversampling and/or the presence of multiple antenna at the receiver. In this case, the received signal can be described by

$$\mathbf{r}(k) = \sum_{i=0}^{M} x(k-i)\mathbf{h}(i) + \mathbf{n}(k) \tag{13.31}$$

where

$$\begin{aligned}
\mathbf{r}(k) &= [r_1(k)\ \ r_2(k) \cdots r_I(k)]^T \\
\mathbf{n}(k) &= [n_1(k)\ \ n_2(k) \cdots n_I(k)]^T \\
\mathbf{h}(m) &= [h_1(m)\ \ h_2(m) \cdots h_I(m)]^T
\end{aligned}$$

The elements of vector $\mathbf{r}(k)$ represent the I received signals at instant k, $\mathbf{n}(k)$ collects the noise samples from each subchannel at the same instant. The elements of vector $\mathbf{h}(m)$, that is $h_i(m)$, represent the mth sample of the ith subchannel model, for $m = 0, 1, \ldots, M$ and $i = 1, 2, \ldots, I$.

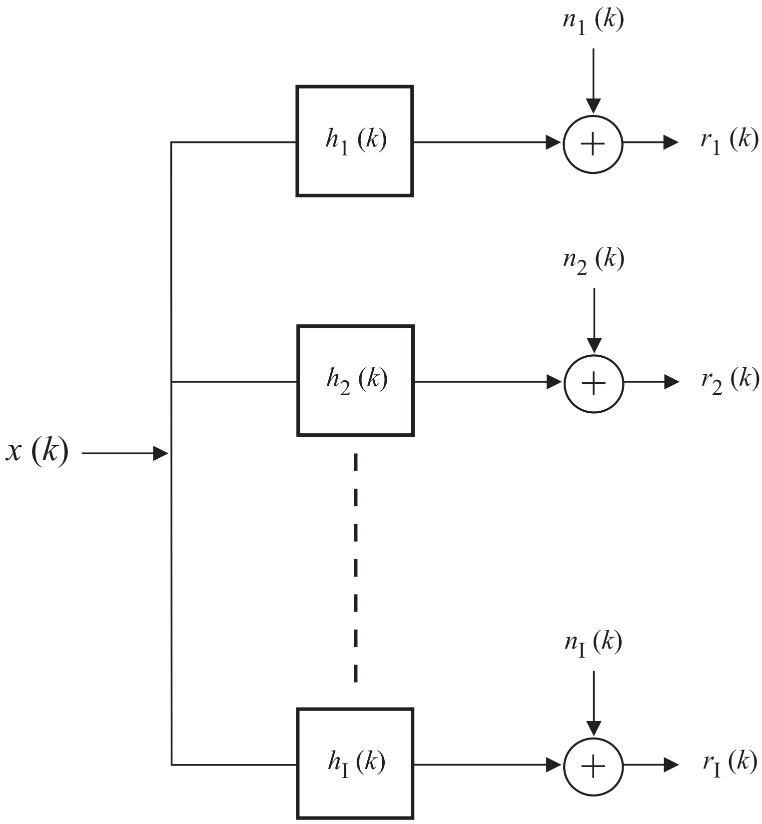


Figure 13.10 Single-input multiple-output model.

Now let's collect N samples of information vectors and pile them up in long vectors such that the received signal vector is function of the input signal block as follows

$$\bar{\mathbf{r}}(k) = \mathbf{H}\mathbf{x}(k) + \bar{\mathbf{n}}(k) \tag{13.32}$$

where

$$\bar{\mathbf{r}}(k) = \left[\mathbf{r}^T(k)\ \mathbf{r}^T(k-1) \cdots \mathbf{r}^T(k-N+1)\right]^T$$
$$\bar{\mathbf{n}}(k) = \left[\mathbf{n}^T(k)\ \mathbf{n}^T(k-1) \cdots \mathbf{n}^T(k-N+1)\right]^T$$
$$\bar{\mathbf{x}}(k) = [x(k)\ x(k-1) \cdots x(k-M-N+1)]^T$$
$$\mathbf{H} = \begin{bmatrix} \mathbf{h}(0) & \cdots & \mathbf{h}(M) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{h}(0) & \cdots & \mathbf{h}(M) & \mathbf{0} & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{h}(0) & \cdots & \mathbf{h}(M) \end{bmatrix}$$

Vectors $\bar{\mathbf{r}}(k)$ and $\bar{\mathbf{n}}(k)$ have dimension NI, the input signal vector $\bar{\mathbf{x}}(k)$ has dimension $N + M$ whereas the channel model matrix $\mathbf{H}$ has dimension $NI \times M + N$ and is a block Toeplitz matrix.

Applying a linear combiner equalizer to the system of equation (13.32) the following relation results

$$y(k) = \bar{\mathbf{w}}^H(k)\bar{\mathbf{r}}(k) = \bar{\mathbf{w}}^H(k)\mathbf{H}\bar{\mathbf{x}}(k) + \bar{\mathbf{w}}^H(k)\bar{\mathbf{n}}(k) \tag{13.33}$$

The coefficient vector $\bar{\mathbf{w}}(k)$ is the equalizer vector of length NI described as

$$\bar{\mathbf{w}}(k) = \left[\tilde{\mathbf{w}}_0^T(k) \;\; \tilde{\mathbf{w}}_1^T(k) \cdots \tilde{\mathbf{w}}_{N-1}^T(k)\right]^T \tag{13.34}$$

where the vector $\tilde{\mathbf{w}}_n(k)$ represents the weights applied to $\mathbf{r}(k-n)$, for $n = 0, 1, \ldots, N-1$. The ith element of $\tilde{\mathbf{w}}_n(k)$, for $i = 1, 2, \ldots, I$, represents the ith weight applied to the corresponding element of $\mathbf{r}(k-n)$.

In a noiseless environment the zero-forcing equalizer is the desired solution such that

$$\bar{\mathbf{w}}^H(k)\mathbf{H} = [0 \ldots 0 \;\; 1 \;\; 0 \ldots 0]^T \tag{13.35}$$

However, the possible noise enhancement originated by $\bar{\mathbf{w}}^T(k)\bar{\mathbf{n}}(k)$ makes the zero-forcing solution not practical in many situations.

13.4.1 Identification Conditions

An FIR channel is identifiable utilizing second-order statistics whenever the block Toeplitz matrix $\mathbf{H}$ in equation (13.32) has full column rank, such that there is a left inverse. Alternatively, we can say that the system of equation (13.32) can be equalized according to some objective function, if for a set of subchannels each with order M the following conditions are met

1. $\text{rank}[\mathbf{H}] = M + N$.

   This means that matrix $\mathbf{H}$ has full column rank.

2. $NI \geq N + M$, i.e., $\mathbf{H}$ is a tall matrix in the case $NI > N + M$.

   In the latter case, this means that matrix $\mathbf{H}$ has more rows than columns.

For the case $N \geq M$, condition 1 is equivalent to say that the transfer functions

$$H_i(z) = \sum_{m=0}^{M} h_i(m) z^{-m} \tag{13.36}$$

for $i = 1, 2, \ldots, I$, have no common zeros [1], that is, the polynomials $H_i(z)$ are coprime. In the case $\frac{M}{I-1} \leq N < M$, we cannot infer that whenever $H_i(z)$, for $i = 1, 2, \ldots, I$, have no common zeros, the matrix $\mathbf{H}$ will have full column rank. In case the $H_i(z)$ have common zeros there is no left inverse matrix for $\mathbf{H}$. In addition, it can also be shown that even if the subchannels are coprime, the matrix $\mathbf{H}$ has its rank reduced if $N < M$. Condition 2 is equivalent to say that the channel matrix $\mathbf{H}$ has full column rank, making possible the channel equalization as well as identification using second-order statistics. Several alternative proofs related to the identifiability of a SIMO system are available in the literature such as in [17]-[19], no proof is included here.

Once satisfied the conditions for identifiability in the SIMO system, the finite-length input signal included in $\bar{\mathbf{x}}(k)$ should contain a large number of modes meaning it should have rich spectral content. This way, in a noiseless environment the SIMO channel can be perfectly identified, except for a gain ambiguity[6], through several methods available in the literature [1], [11]-[14]. The requirements on the channel input signal statistics vary from method to method, with some requiring that it is uncorrelated while others not.

The same type of results applies for the SIMO blind equalizers, that is, a single-input I-output channel can be equalized whenever:

- At least one of the subchannels has length $M+1$, i.e., $h_i(0) \neq 0$ and $h_i(M) \neq 0$, for any $i = 1, 2, \ldots, I$.
- $H_i(z)$ for $i = 1, 2, \ldots, I$, have no common zeros.
- $N \geq M$.

These conditions are necessary and sufficient for the SIMO channel identifiability or equalization utilizing second-order statistics of the I outputs.

Many of the available solutions for blind channel identification and equalization based on second-order statistics are very sensitive to channel order or rank estimation. Some of them rely on singular value decomposition(s) (SVD) which are very computationally complex and are usually meant for batch form of implementation. The emphasis here is to present a recursive solution which is more robust to order estimation errors and is computationally attractive such that it can be applied to track time-varying channels. An online blind SIMO equalizer is introduced in the following section.

13.5 SIMO-CMA EQUALIZER

This section discusses an important result that suggests that by combining the techniques implicitly utilizing high-order statistics such as the CMA, with SIMO systems using second-order statistics can be very beneficial. Let's start by stating the following result whose proof can be found in [1], [20]:

In a noiseless channel, if the Multiple-Input Single-Output (MISO) FIR equalizer has length $N \geq M$, then the SIMO CMA equalizer is globally convergent if the subchannels $H_i(z)$ for $i = 1, 2, \ldots, I$, have no common zeros.

[6] A constant value multiplying the channel model.

The reader should notice that a SIMO setup utilizing a CM objective function can be interpreted as fractionally spaced constant-modulus equalizer.

The expression for the SIMO equalizer output signal as described in equation (13.33) can rewritten as

$$y(k) = \sum_{i=1}^{I} \mathbf{w}_i^H(k)\mathbf{r}_i(k) \tag{13.37}$$

where the nth element of vector $\mathbf{w}_i(k)$ corresponds to the $(i+n-1)$th element of $\bar{\mathbf{w}}(k)$, and the nth element of vector $\mathbf{r}_i(k)$ corresponds to $r_i(k-n)$, for $i = 1, 2, \ldots, I$, and $n = 0, 1, \ldots, N-1$. The equivalent SIMO system is depicted in Fig. 13.11, where it can be observed that the overall equalization consists of using a separated sub-equalizer for each sub-channel with a global output signal used in the blind adaptation algorithm.

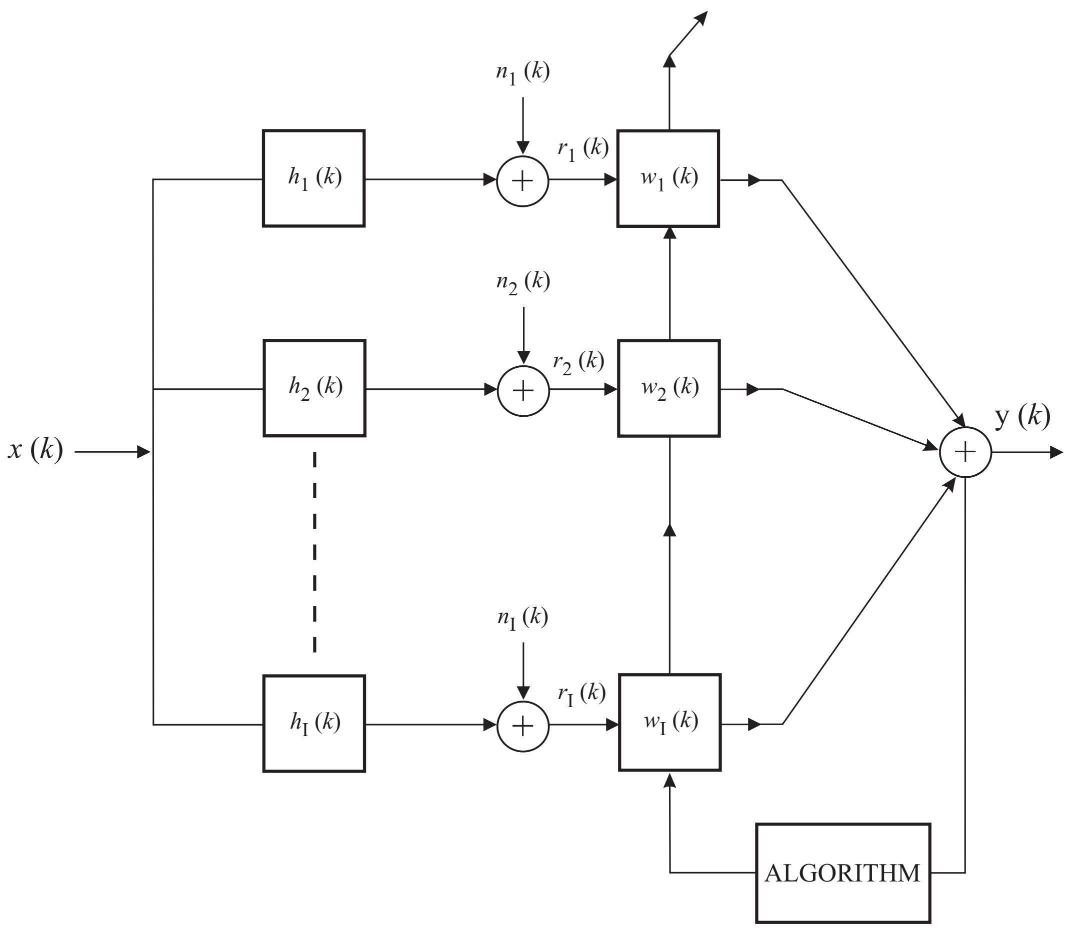


Figure 13.11 SIMO equalizer.

Algorithm 13.5

SIMO Affine Projection CM Algorithm

Initialization
$\bar{\mathbf{r}}(0) = \bar{\mathbf{w}}(0) = \text{random vectors}$
choose μ in the range $0 < \mu \leq 2$
$\gamma = $ small constant
Do for $k > 0$
$\mathbf{y}^*_{\text{ap}}(k) = \mathbf{X}^H_{\text{ap}}(k)\bar{\mathbf{w}}(k)$
$\mathbf{r}^*_{\text{ap}}(k) = \text{sgn}[\mathbf{X}^H_{\text{ap}}(k)\bar{\mathbf{w}}(k)]$
$\mathbf{e}^*_{\text{ap}}(k) = \mathbf{r}^*_{\text{ap}}(k) - \mathbf{y}^*_{\text{ap}}(k)$
$\bar{\mathbf{w}}(k+1) = \bar{\mathbf{w}}(k) + \mu \mathbf{X}_{\text{ap}}(k)\left(\mathbf{X}^H_{\text{ap}}(k)\mathbf{X}_{\text{ap}}(k) + \gamma \mathbf{I}\right)^{-1} \mathbf{e}^*_{\text{ap}}(k)$

In the case we adopt a CMA objective function along with the affine projection algorithm to derive a SIMO equalizer, the $\mathbf{X}_{\text{ap}}(k)$ matrix, assuming we keep the last $L+1$ input signal vectors, has the following form

$$\mathbf{X}_{\text{ap}}(k) = [\bar{\mathbf{r}}(k)\, \bar{\mathbf{r}}(k-1) \ldots \bar{\mathbf{r}}(k-L)] \tag{13.38}$$

The adaptive-filter output vector is described by

$$\begin{aligned}
\mathbf{y}^*_{\text{ap}}(k) &= \mathbf{X}^H_{\text{ap}}(k)\bar{\mathbf{w}}(k) \\
&= \begin{bmatrix} \bar{\mathbf{r}}^H(k) \\ \bar{\mathbf{r}}^H(k-1) \\ \vdots \\ \bar{\mathbf{r}}^H(k-L) \end{bmatrix} \bar{\mathbf{w}}(k) \\
&= \begin{bmatrix} \bar{\mathbf{r}}^H(k) \\ \bar{\mathbf{r}}^H(k-1) \\ \vdots \\ \bar{\mathbf{r}}^H(k-L) \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{w}}_0(k) \\ \tilde{\mathbf{w}}_1(k) \\ \vdots \\ \tilde{\mathbf{w}}_{N-1}(k) \end{bmatrix}
\end{aligned} \tag{13.39}$$

where in the last equality we adopted the description of $\bar{\mathbf{w}}(k)$ as given by equation (13.34). By following the same derivations of section 13.3 it is possible to generate the SIMO affine projection CM algorithm as described in Algorithm 13.5. The affine projection algorithm is expected to converge to the global optimum using normalized steps originated by the minimal distance principle utilized in its derivations, as discussed in Chapter 4.

Example 13.4

Given the one-input two-output channel whose model is described below. Assume a QAM signal with four symbols is transmitted through these channels and simulate a blind equalization using the SIMO affine projection CM algorithm of order 12, for a signal to noise ratio of 20dB measured at the receiver input.

$$\begin{bmatrix} \mathbf{h}_1^T \\ \mathbf{h}_2^T \end{bmatrix} = \begin{bmatrix} 0.1823 & -0.7494 & -0.4479 & 0.2423 & 0.0047 & -0.41 \\ 0.3761 & -0.1612 & -0.1466 & 0.6437 & 0.5952 & -0.2060 \end{bmatrix}$$

Solution:

We utilize the affine projection CM algorithm to solve the SIMO equalization problem with $\mu = 0.1$, $L = 2$ and $\gamma = 10^{-6}$. The symbol error rate is measured by averaging the outcoming results of 50 independent runs, and the initial conditions utilized correspond to the Wiener solution randomly disturbed. Fig. 13.12 shows the evolution of the errors in the symbols, and as can be observed minimum symbol error rate occurs after 500 iterations. This result is expected since the conditions for the correct channel equalization is met in this case, see subsection 13.4.1, and there is some channel noise. Fig. 13.13 depicts the MSE between the equalized signal and the transmitted symbols where the convergence of the affine projection CM algorithm takes places in around 1000 iterations. Fig. 13.14 illustrates the effectiveness of the equalizer through the appropriate combination of signals measured in each antenna.

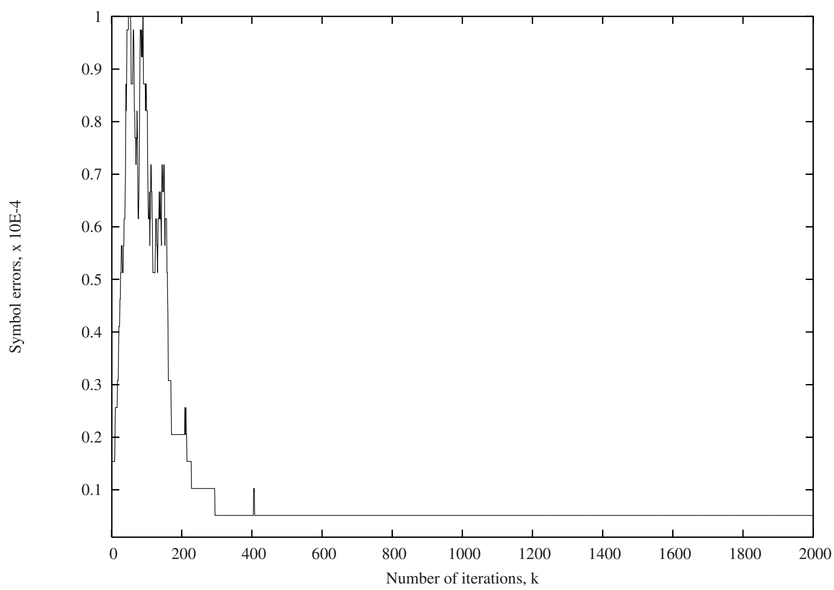


Figure 13.12 Symbol errors; affine projection CM algorithm.

□

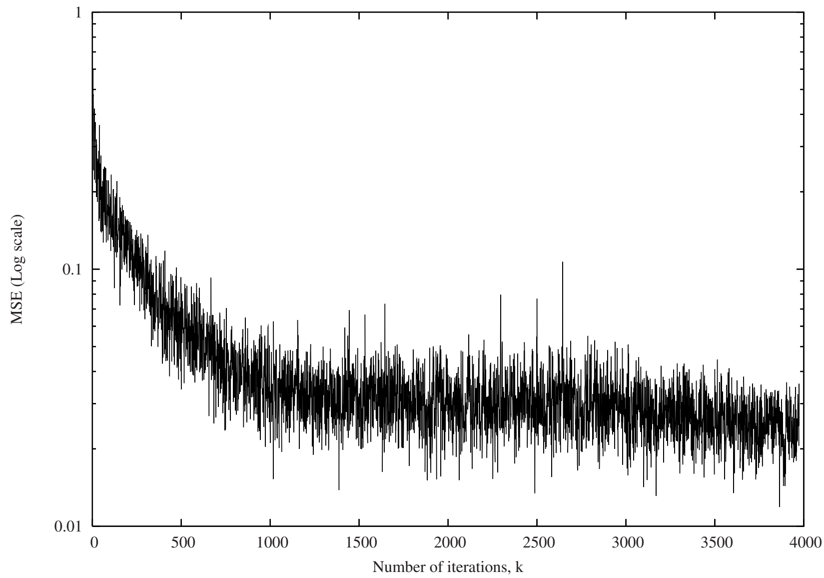


Figure 13.13 Learning curve (MSE), $\mu = 0.1$, SNR = 20 dB, order=12.

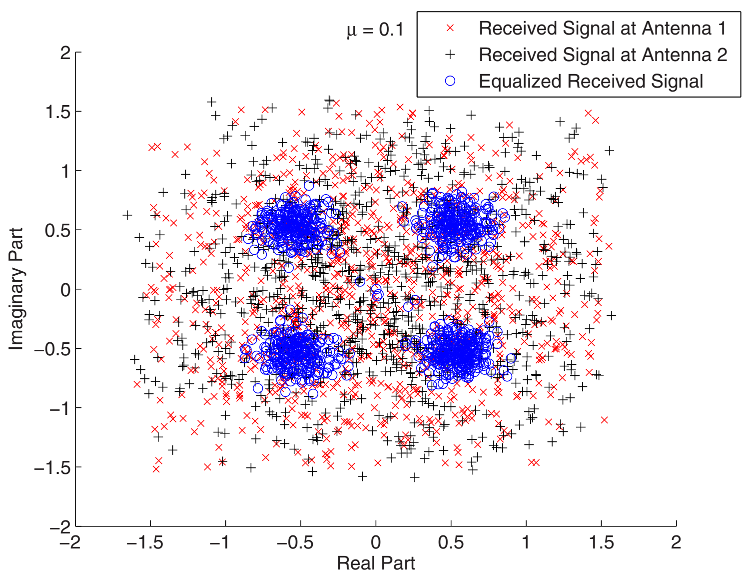


Figure 13.14 Equalized signals for the SIMO affine projection CM algorithm, with $L = 2$.

Example 13.5

Repeat the Example 6.3 by measuring through simulations the MSE performance of an equalizer implemented with the SIMO affine projection CM algorithm, when it is available two received signals obtained through different antennas. Choose the appropriate parameters and comment on the results.

Solution:

The channels available for the detection of the transmitted symbols correspond to the transfer function from the transmitter to each antenna. The blind affine projection CM algorithm is employed to update the sub equalizers of the SIMO system. The parameters chosen after some simulation trials are $\mu = 0.3$, $L = 1$, and $\gamma = 10^{-6}$. The measures of MSE reflect an average taken from the outcomes of 50 independent runs, where in the initialization one of the receiver filters is set to the Wiener solution during the first 350 iterations. Each sub equalizer has order 30. Fig. 13.15 illustrates the MSE evolution and as can be observed only after a few thousand iterations the curve shows a non decreasing behavior. In comparison with the results from Example 6.3, the learning process takes a lot more iterations when compared to the algorithms employing some sort of training. However, despite of slower convergence the equalization is feasible since the conditions for the correct channel equalization are met.

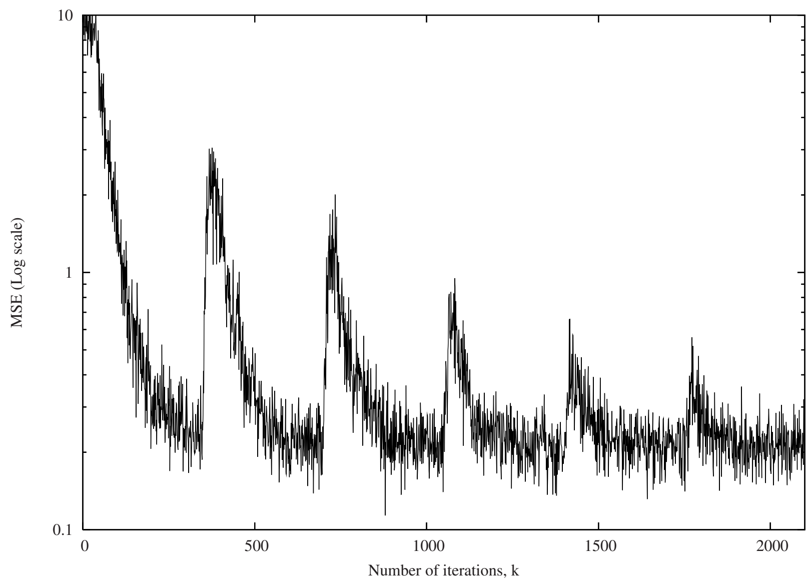


Figure 13.15 Learning curve of the SIMO affine projection CM algorithm, $L = 1$.

□

The SIMO formulation presented in this chapter can be extended to the Multi-Input Multi-Output (MIMO) case in rather straightforward way, under some assumptions such as independence of the sources. There are several communication system setups that can be modeled as MIMO systems by properly stacking the transmitted and received information. In some applications the setup is originally MIMO such as in multiuser communication systems [21]-[28], and in case we use antenna array at transmitter and receiver in order to increase the communication capacity [30]-[33]. In many MIMO applications adaptive-filtering algorithms are often utilized with training or in a blind form.

The affine projection CM algorithm presented in this chapter can be extended to include selective updating using the set-membership approach presented in Chapter 6. In addition, for multiuser environments such as CDMA systems, it is possible to incorporate some blind measurements related to the multi-access and additional noise interferences in order to improve the overall performance of blind receivers based on the set-membership affine projection CM algorithm, as discussed in [34]. The set-membership affine projection algorithm can be very efficient in SIMO as well as in MIMO setups.

13.6 CONCLUDING REMARKS

This chapter presented some blind adaptive-filtering algorithms mostly aimed at direct blind channel equalization. The subject of blind signal processing is quite extensive, as a result our emphasis was to present the related issues and to introduce some useful algorithms. In particular it was introduced some algorithms utilizing high-order statistics in an implicitly way, since the resulting algorithms have low computational complexity[7] while presenting slow convergence and possible convergence to local minima. The cases introduced in this class were the constant-modulus, Godard, and Sato algorithms, respectively. Some issues related to the error surface of the CM algorithm were also illustrated through a simple example.

In order to improve the convergence speed of the CMA family of algorithms its affine projection version was presented. This algorithm certainly alleviates the speed limitations of the CM algorithms at the expense of increased computational complexity. In addition, this chapter discussed the single-input multi-output methods which allow under certain conditions the correct identification and equalization of unknown channels using only second-order statistics and do not have local minima. In fact, the combination of the algorithms with implicit high-order statistics, with the affine projection update equation and the single-input multi-output setup leads to very interesting solutions for blind channel equalization. The resulting algorithm has rather fast convergence and has only global solutions under certain conditions.

In specific cases, we can conclude that fractionally spaced equalizers using indirect high-order statistics such as the CM algorithms are not suitable to equalize channels with zeros in common. In case this happens an additional equalizer after the SIMO equalizer might help in combating the remaining intersymbol interference. On the other hand, the SIMO equalizers are suitable to equalize channels with zeros on the unit circle, a rough situation for symbol spaced equalizers. In this case, the SIMO equalizer can be used with an implicit high-order statistics objective function or with training signal, as long as the subchannels do not have common zeros. For situations with common zeros on the unit circle, or close to it, the standard way out is to employ DFE.

[7] In comparison with the algorithms using high-order statistics explicitly.

13.7 REFERENCES

1. Z. Ding and Y. Li, *Blind Equalization and Identification*, Marcel Dekker, New York, N.Y., 2001.

2. S. Haykin, Editor, *Unsupervised Adaptive Filtering, Vol. I: Blind Source Separation*, John Wiley & Sons, New York, NY, 2000.

3. S. Haykin, Editor, *Unsupervised Adaptive Filtering, Vol. II: Blind Deconvolution*, John Wiley & Sons, New York, NY, 2000.

4. C.-Y. Chi, C.-C. Feng, C.-H. Chen, and C.-Y. Chen, *Blind Equalization and System Identification*, Springer, London, UK, 2006.

5. Y. Sato, "A method of self-recovering equalization for multi-level amplitude modulation," *IEEE Trans. on Communications*, vol. COM-23, pp. 679-682, June 1975.

6. A. Benveniste, M. Gousat, and R. Ruget, "Robust identification nonminimum phase system: Blind adjustment of a linear equalizer in data communications," *IEEE Trans. on Automatic Control*, vol. AC-25, pp. 385-399, June 1980.

7. D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication system," *IEEE Trans. on Communications*, vol. COM-28, pp. 1867-1875, Nov. 1980.

8. J. R. Barry, E. A. Lee, and D. G. Messerschmitt, *Digital Communication*, Kluwer Academic Publishers, Boston, MA, 3rd edition, 2004.

9. C. L. Nikias and A. P. Petropulu, *Higher-order spectra analysis: A nonlinear signal processing framework*, Prentice Hall, Englewood Cliffs, NJ, 1993.

10. C. B. Papadias and D. T. M. Slock, "Normalized sliding window constant-modulus and decision-directed algorithms: A link between blind equalization and classical adaptive filtering," *IEEE Trans. on Signal Processing*, vol. 45, pp. 231-235, Jan. 1997.

11. E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayarargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. on Signal Processing*, vol. 43, pp. 516-525, Feb. 1995.

12. D. Gesbert and P. Duhamel, "Unbiased blind adaptive channel identification," *IEEE Trans. on Signal Processing*, vol. 48, pp. 148-158, Jan. 2000.

13. X. Li and H. Fan, "Direct estimation of blind zero-forcing equalizers based on second-order statistics," *IEEE Trans. on Signal Processing*, vol. 48, pp. 2211-2218, Aug. 2000.

14. L. Tong, and Q. Zhao, "Jointly order detection and blind channel estimation by least squares smoothing," *IEEE Trans. on Signal Processing*, vol. 47, pp. 2345-2355, Sept. 1999.

15. A. P. Liavas, P. A. Regalia, and J.-P. Delmas, "Blind channel approximation: effective channel order determination," *IEEE Trans. on Signal Processing*, vol. 47, pp. 3336-3344, Dec. 1999.

16. L. Tong and Z. Ding, "Single-user channel estimation and equalization," *IEEE Signal Processing Magazine*, vol. 17, pp. 17-28, May 2000.

17. L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. on Information Theory*, vol. 40, pp. 340-349, March 1994.

18. L. Tong, G. Xu, B. Hassibi, and T. Kailath, "Blind identification and equalization based on second-order statistics: A frequency-domain approach," *IEEE Trans. on Information Theory*, vol. 41, pp. 329-334, March 1994.

19. E. Serpedin and G. B. Giannakis, "A simple proof of a known blind channel identifiability result," *IEEE Trans. on Signal Processing*, vol. 47, pp. 591-593, Feb. 1999.

20. Y. Li and Z. Ding, "Global convergence of fractionally spaced Godard (CMA) adaptive equalizers," *IEEE Trans. on Signal Processing*, vol. 44, pp. 818-826, April 1996.

21. A. J. Viterbi, *Principles of Spread Spectrum Communication*, Addison Wesley, Reading, MA, 1995.

22. M. Honig and M. K. Tsatsanis, "Adaptive techniques for multiuser CDMA receivers," *IEEE Signal Processing Magazine*, vol. 17, pp. 49-61, May 2000.

23. S. Verdu, *Multiuser Detection*, Cambridge University Press, Cambridge, UK, 1998.

24. M. L. Honig, U. Madhow, and S. Verdú, "Blind adaptive multiuser detection," *IEEE Trans. on Information Theory*, vol. 41, pp. 944-960, July 1995.

25. M. K. Tsatsanis, "Inverse filtering criteria for CDMA systems," *IEEE Trans. on Signal Processing*, vol. 45, pp. 102-112, Jan. 1997.

26. Z. Xu and M. K. Tsatsanis, "Blind adaptive algorithms for minimum variance CDMA receivers," *IEEE Trans. on Signal Processing*, vol. 49, pp. 180-194, Jan. 2001.

27. Z. Xu, P. Liu, and X. Wang, "Blind multiuser detection: From MOE to subspace methods," *IEEE Trans. on Signal Processing*, vol. 52, pp. 510-524, Feb. 2004.

28. X. Wang and H. V. Poor, *Wireless Communication Systems: Advanced Techniques for Signal Reception*, Prentice Hall, Upper Saddle River, NJ, 2003.

29. E. de Carvalho and D. T. M. Slock, "Blind and semi-blind FIR multichannel estimation: (Global) Identifiability conditions," *IEEE Trans. on Signal Processing*, vol. 52, pp. 1053-1064, April 2004.

30. A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*, Cambridge University Press, Cambridge, UK, 2003.

31. E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*, Cambridge University Press, Cambridge, UK, 2003.

32. A. Hottinen, O. Tirkkonen, and R. Wichman, *Multi-Antenna Transceiver Techniques for 3G and Beyond*, John Wiley, New York, NY, 2003.

33. H. L. Van Trees, *Optimum Array Processing: Detection, Estimation, and Modulation Theory, Part IV*, John Wiley Interscience, New York, NY, 2002.

34. R. C. de Lamare and P. S. R. Diniz, “Blind constrained set-membership algorithms with time-varying bounds for CDMA interference suppression,” *Proc. IEEE Intern. Conf. on Acoust. Speech and Signal Processing*, Toulouse, France, pp. IV-617 - IV-620, May 2006.

13.8 PROBLEMS

1. Derive the Godard algorithm for real input signal constellations.

2. Derive equations (13.13) and (13.14).

   Hint: Utilize the difference equation that describes $x(k)$.

3. Perform the equalization of a channel with the following impulse response

$$h(k) = ku(k) - (2k-9)u(k-5) + (k-9)u(k-10)$$

   using as transmitted signal a binary (-1,1) random signal. An additional Gaussian white noise with variance 10^{-2} is present at the channel output.
   (a) Apply the Godard algorithm for $p = q = 4$ with an appropriate μ and find the impulse response of an equalizer with 15 coefficients.
   (b) Plot the detected equalized signal before the decision after the algorithm has converged for a number of iterations (over 50 samples) and comment on the result.

4. Repeat the problem 3 for the Sato algorithm.

5. Repeat the problem 3 for the CMA.

6. Assume a PAM signal with four symbols is transmitted through an AR channel whose transfer function is

$$H(z) = \frac{0.25z}{z+0.5}$$

   The equalizer is a first-order FIR adaptive filter. For a signal to noise ratio of 5dB, plot the error surface and contours for Godard with $p = q = 4$.

7. Assume a QAM signal with four symbols is transmitted through an AR channel whose transfer function is

$$H(z) = \frac{0.25z}{z+0.5}$$

   Simulate a blind equalization using a first-order FIR adaptive filter, for a signal to noise ratio of 10dB, using the CMA.

8. Given the channel model below whose input is a binary PAM signal.

$$\begin{aligned} H(z) &= 0.2816 + 0.5622z^{-1} + 0.2677z^{-2} - 0.3260z^{-3} - 0.4451z^{-4} \\ &\quad +0.3102z^{-5} - 0.2992z^{-6} - 0.2004z^{-7} \end{aligned}$$

Our objective is to equalize this channel with a blind affine projection CM algorithm. The equalizer has order 10 and its objective is to shorten the effective impulse response of the equalized signal. That means the channel-equalizer impulse response has most of its energy concentrated in a few samples. Simulate this experiment for a signal to noise ratio of 15dB, and comment on the channel shortening process.

9. Derive the set-membership affine projection CM algorithm.

10. (a) Show that recursion of equation (13.30) minimizes the objective function of equation (13.29).

    (b) Show that recursion of equation (13.30) also minimizes the objective function

$$\begin{array}{l} \|\mathbf{w}(k+1) - \mathbf{w}(k)\|^2 \\ \text{subject to :} \\ |\text{sgn}\left[\mathbf{x}^H(k)\mathbf{w}(k+1)\right] - |\mathbf{x}^H(k)\mathbf{w}(k+1)|^q|^p = 0 \end{array}$$

11. Derive a constrained minimum variance (CMV) affine projection algorithm for equalization, whose objective function is to minimize

$$\begin{array}{l} \frac{1}{2}\|\mathbf{w}(k+1) - \mathbf{w}(k)\|^2 \\ \text{and} \\ \frac{1}{2}\mathbf{w}^T(k+1)\mathbf{r}(k)\mathbf{r}^T(k)\mathbf{w}(k+1) \\ \text{subject to :} \\ \mathbf{w}^T(k+1)\mathbf{c} = c \end{array}$$

    where $\mathbf{r}(k)$ is a vector that in the present case represents the received signal vector, c is an arbitrary constant, and **c** is a constraint vector.

12. Assume a PAM signal with two symbols is transmitted through a noiseless AR channel whose transfer function is

$$H(z) = \frac{0.25z}{z + 0.5}$$

    Simulate a blind equalization using a first-order FIR adaptive filter, using affine projection CM algorithm as well as the stochastic gradient version CMA. Plot the convergence trajectories of $w_0(k)$ and $w_1(k)$ for 20 distinct initialization points (on the same figure) for $w_0(0)$ and $w_1(0)$ corresponding to zeros in the interior of unit circle. Interpret the results.

13. Equalize the one-input two-output channel described below using the SIMO affine projection CM algorithm. The input signal is a two PAM signal representing a randomly generated bit stream with the signal to noise ratios $\frac{\sigma_{r_i}^2}{\sigma_n^2} = 20$, for $i = 1, 2$, at the receiver end, that is, $r_i(k)$ is the received signal without taking into consideration the additional channel noise. Choose the appropriate equalizer order and the number of reuses such that the bit error rate falls below 0.01.

$$
\begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 \end{bmatrix} = \begin{bmatrix} 0.345 & -0.715 \\ -0.016 & 0.690 \\ -0.324 & 0.625 \\ 0.209 & 0.120 \\ 0.253 & 0.388 \\ -0.213 & 0.132 \\ 0.254 & -0.120 \\ 0.118 & -0.388 \\ 0.483 & 0.451 \\ -0.034 & -0.204 \\ 0.462 & 0.560 \\ -0.111 & -0.675 \\ -0.285 & 0.147 \end{bmatrix}
$$

14. Using the complex version of the SIMO affine projection CM algorithm to equalize the one-input two-output channel with the transfer function given below. The input signal is a four QAM signal representing a randomly generated bit stream with the signal to noise ratios $\frac{\sigma_{r_i}^2}{\sigma_n^2} = 10$, for $i = 1, 2$, at the receiver end, that is, $r_i(k)$ is the received signal without taking into consideration the additional channel noise. The adaptive filter has 5 coefficients.

$$H_1(z) = (0.27 - 0.34\jmath) + (0.43 + 0.87\jmath)z^{-1} + (0.21 - 0.34\jmath)z^{-2}$$

$$H_2(z) = (0.34 - 0.27\jmath) + (0.87 + 0.43\jmath)z^{-1} + (0.34 - 0.21\jmath)z^{-2}$$

(a) Run the algorithm for $\mu = 0.1$, $\mu = 0.4$, and $\mu = 0.8$. Comment on the convergence behavior in each case.
(b) Plot the real versus imaginary parts of the received signals before equalization and the single output signal after equalization.

15. Repeat problem 14 for the case the adaptive-filter order is one and comment on the results.