Towards Globally Convergent Blind Equalization of Constant Modulus Signals: A Bilinear Approach

Constantinos B. PAPADIAS † and Dirk T. M. SLOCK ††

†, †† *Institut EURECOM, 2229 route des Crêtes*
B.P. 193, 06904, Sophia Antipolis Cedex, FRANCE
† *Tel/Fax: +33 93002618/ 93002627, E-Mail: papadias@eurecom.fr*
†† *Tel/Fax: +33 93002606/ 93002627, E-Mail: slock@eurecom.fr*

Abstract. We consider the problem of blind equalization of a constant modulus signal. One of the most popular classes of algorithms in this context is the Godard family of blind equalizers [1], which includes among others the Constant Modulus Algorithm (CMA) [2]. A common drawback of these algorithms is that they may converge to undesired equalizer settings if not properly initialized. This is known as the problem of *ill convergence* and is primarily due to the non-convex form of the cost function of algorithms of this class with respect to the equalizer parameters. We propose a different approach to the problem, namely, a *bilinear* one, which leads to a different parameterization and to the construction of a convex cost function with respect to the parameters introduced. In a perfectly parameterized case (the equalizer's order matches exactly the order of the channel inverse), the solution to the problem is unique and permits for a direct calculation of the optimal equalizer. In over-parameterized cases however, there exist multiple solutions to our cost function. However, we propose a method that still allows to determine the channel inverse in this case. Different adaptive schemes are proposed to adaptively compute the solution of our criterion and the influence of additive noise is also discussed.

1. The Constant Modulus Algorithm

The classical setup for adaptive blind-equalization is shown in figure 1. We denote by a_k, x_k and y_k the transmitted symbol, received sample and equalizer output, respectively, all at time instant k. The equalizer (a FIR filter of order $N-1$) at the same time instant is denoted by a $N \times 1$ vector $W_k = [w_0 \cdots w_{N-1}]^T$ and its output can be written as $y_k = X_k^H W_k$, where $X_k = [x_k \; x_{k-1} \cdots x_{k-N+1}]^H$, ^{T} denotes transpose and ^{H} complex conjugate transpose. The Godard algorithms are stochastic gradi-

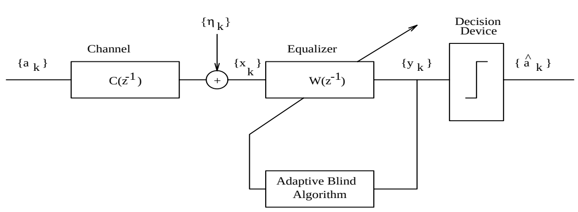

Figure 1: An adaptive blind equalization setup

ent algorithms that attempt to solve the following minimization problem:

$$\min_W J_p(W) = \frac{1}{2p} E(|y|^p - l_p)^2 \quad p \in \{1, 2, \ldots\} \ , \tag{1}$$

where E denotes statistical expectation and $l_p = \frac{E|a_k|^{2p}}{E|a_k|^p}$. The popular CMA 2-2 corresponds to the particular choice $p = 2$ and if one assumes that the constellation modulus equals one ($|a_k| = 1$), is given by:

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

where μ denotes the algorithm's stepsize. Due to the non-convex form of the cost function $J_p(W)$ w.r.t. W, the problem (1) admits more than one solution which can be found by setting the gradient of $J_p(W)$ w.r.t. W equal to zero. This gives for $p = 2$:

$$E((|y_k|^2 - 1) y_k X_k) = 0 \ . \tag{3}$$

The stochastic gradient algorithm described by (2) may therefore converge to one of the solutions of (3) which does not correspond to the global but to a local minimum of $J_2(W)$ if it is initialized close to it and if a small stepsize μ (necessary for the stability of (2)) is used [3]. This is the problem of ill-convergence of CMA's and in the sequel we will present a different approach that tries to circumvent it.

2. A bilinear approach

Consider the following expansion of the squared modulus of the equalizer's output:

$$\begin{aligned} |y_k|^2 = y_k y_k^* = (w_0 w_0^* x_k x_k^* + \cdots + w_0 w_{N-1}^* x_k x_{k-N+1}^*) + \cdots + \\ +(w_{N-1} w_0^* x_{k-N+1} x_k^* + \cdots + w_{N-1} w_{N-1}^* x_{k-N+1} x_{k-N+1}^*) \ , \end{aligned} \tag{4}$$

where * denotes complex conjugate. If now we introduce a $N^2 \times 1$ bilinear regression vector

$$\begin{aligned} \mathcal{X}_k = [x_k x_k^* \cdots x_k x_{k-N+1}^* \; x_{k-1} x_k^* \cdots x_{k-1} x_{k-N+1}^* \cdots \\ x_{k-N+1} x_k^* \cdots x_{k-N+1} x_{k-N+1}^*]^H \ , \end{aligned} \tag{5}$$

and a $N^2 \times 1$ parameter vector

$$\begin{aligned} \theta_k = [w_0 w_0^* \cdots w_0 w_{N-1}^* \; w_1 w_0^* \cdots w_1 w_{N-1}^* \cdots \\ w_{N-1} w_0^* \cdots w_{N-1} w_{N-1}^*]^T \ , \end{aligned} \tag{6}$$

that contain all the bilinear terms $x_{k-i} x_{k-j}^*$ and $w_i w_j^*$ of the expansion in (4), respectively, then the squared modulus of the output y_k can be written as:

$$|y_k|^2 = z_k = \mathcal{X}_k^H \theta_k \ , \tag{7}$$

and may be viewed as the output of a linear "filter" with impulse response θ which is excited by the "bilinear" regression vector $\mathcal{X}$, at time instant k. A constant-modulus cost function can now be constructed in terms of the new parameter vector θ and the resulting minimization problem will be:

$$\min_{\theta} J^{bil}(\theta) = \min_{\theta} E(z-1)^2 = \min_{\theta} E(\mathcal{X}^H\theta - 1)^2 \quad . \tag{8}$$

Note that the cost function $J^{bil}(\theta)$ is quadratic, and therefore convex. The principle behind this parameterization of the problem is the following: in traditional blind equalization, the received discrete-time signal is first passed through a linear filter (equalizer) and then some kind of nonlinearity is applied to its output in order to provide the higher order moments needed for the identification of the channel (or the channel's inverse) impulse response. Here we have somewhat "interchanged" the order of these two operations in that we first apply a nonlinearity to the received signal (in order to form the bilinear regression vector $\mathcal{X}_k$) and pass the resulting process through a linear filter whose output is not submitted to further nonlinearities.

The gradient of $J^{bil}(\theta)$ w.r.t. θ is given by:

$$\nabla_{\theta} J^{bil}(\theta) = 2E(\mathcal{X}(\mathcal{X}^H\theta - 1)) \ , \tag{9}$$

and therefore any solution θ to (8) should satisfy the following equation:

$$E(\mathcal{X}\mathcal{X}^H)\theta = E(\mathcal{X}) \ . \tag{10}$$

Now if the matrix $E(\mathcal{X}\mathcal{X}^H)$ is invertible, the problem (8) has the following unique solution:

$$\theta = \{E(\mathcal{X}\mathcal{X}^H)\}^{-1}E(\mathcal{X}) \ . \tag{11}$$

In this case the corresponding equalizer can be found from θ as follows: a zero-forcing (ZF) FIR equalizer of order (N-1) exists if the channel is all-pole of order $(N-1)$:

$$a_k = \sum_{i=0}^{N-1} c_i x_{k-i} = X_k^H C \ , \tag{12}$$

where $C = [c_0 \cdots c_{N-1}]^T$ contains the coefficients of the impulse response of the channel's inverse. Then the optimal (ZF) equalizer will equal the inverse impulse response ($W^{opt} = C$) and the corresponding parameter vector will be given by:

$$\begin{aligned}\theta^{opt} = [c_0 c_0^* \cdots c_0 c_{N-1}^* \ c_1 c_0^* \cdots c_1 c_{N-1}^* \cdots \\ c_{N-1} c_0^* \cdots c_{N-1} c_{N-1}^*]^T \ .\end{aligned} \tag{13}$$

It is clear that θ^{opt} solves (8) since $J^{bil}(\theta^{opt}) = 0$. Moreover, as the matrix $E(\mathcal{X}\mathcal{X}^H)$ was supposed to be invertible, (8) admits a unique solution which is given by (11). Therefore the following lemma holds:

Lemma: When the channel is all-pole of order $N-1$ and the $N^2 \times N^2$ matrix $E(\mathcal{X}\mathcal{X}^H)$ is invertible, the criterion (8) admits the unique solution θ^{opt} given by (13).

Consider now the $N \times N$ matrix Θ that has as columns the N consecutive partitions of the vector θ of N elements each:

$$\Theta = \begin{bmatrix} \theta^{(0)} & \theta^{(N)} & \dots & \theta^{(N^2-N)} \\ \theta^{(1)} & \theta^{(N+1)} & \dots & \theta^{(N^2-N+1)} \\ . & . & \dots & . \\ . & . & \dots & . \\ \theta^{(N-1)} & \theta^{(2N-1)} & \dots & \theta^{(N^2-1)} \end{bmatrix} \ . \tag{14}$$

Then it follows that $\Theta^{opt} = CC^H$. The optimal equalizer setting W^{opt} can now be found from Θ^{opt} as follows:

$$W^{opt} = e^{j\phi}\sqrt{\lambda_{max}} V_{max} \ , \tag{15}$$

where λ_{max} and V_{max} denote the maximum eigenvalue and the corresponding eigenvector of Θ^{opt} and $\phi \in (0, 2\pi)$. Note that an ambiguity is inherent in the choice of the optimal equalizer, since the factor $e^{j\phi}$ cannot be determined. This is a usual phenomenon in blind equalization and can be eliminated by using differential coding at transmission.

Identification of the (generally non-minimum phase) transmission channel is therefore achievable by minimizing the cost function $J^{bil}(\theta)$ w.r.t. θ. This should not be an astonishing result, since the matrix $E(\mathcal{X}\mathcal{X}^H)$ contains 4^{th} order moments of the received signal (it is known that identification of a non-minimum phase channel is not possible at the baud rate by use of only second order statistics).

The solution of (8) can be calculated either in a batch or in an adaptive way. In the first case one just has to estimate the quantities $E(\mathcal{X}\mathcal{X}^H)$ and $E(\mathcal{X})$ and use (11) to calculate the corresponding estimate for θ. In the second case, an adaptive filtering algorithm can be employed for the parameter vector θ that uses at each iteration $\mathcal{X}_k$ as the regression vector and the scalar 1 as the "desired" sample. An LMS-like algorithm (stochastic gradient minimization) for θ_k will be as follows:

$$\begin{aligned} \epsilon_k &= 1 - \mathcal{X}_k^H \theta_k \\ \theta_{k+1} &= \theta_k + \mu \mathcal{X}_k \epsilon_k \ , \end{aligned} \tag{16}$$

where ϵ_k is the a priori error at time instant k and μ the stepsize parameter that controls both the convergence speed and steady-state error of the algorithm. As the parameter vector θ is of length N^2, the complexity of the above algorithm will be $2N^2$ multiplications/iteration. Similarly, an RLS-like algorithm can be used that will allow for faster convergence (in approximately $2N^2$ iterations) but at the expense of a higher computational complexity ($O(N^4)$ multiplications/iteration):

$$\begin{aligned} \epsilon_k &= 1 - \mathcal{X}_k^H \theta_k \\ R_k^{-1} &= \lambda^{-1} R_{k-1}^{-1} - \lambda^{-1} R_{k-1}^{-1} \mathcal{X}_k (1 + \mathcal{X}_k^H \lambda^{-1} R_{k-1}^{-1} \mathcal{X}_k)^{-1} \times \\ & \qquad \times \mathcal{X}_k^H \lambda^{-1} R_{k-1}^{-1} \\ \theta_{k+1} &= \theta_k + R_k^{-1} \mathcal{X}_k \epsilon_k \ , \end{aligned} \tag{17}$$

where R_k is $N^2 \times N^2$ and λ is the forgetting factor. A multichannel Fast Transversal Filter algorithm [4] can be used in order to reduce the complexity to $O(N^3)$ multiplications/iteration (see figure 2).

When the received samples x_k are real (which corresponds to a PAM modulation and a real channel), a reduction in the required number of parameters for the bilinear method is possible. Namely, in this case the expansion in (4) will have only $\frac{N \times (N+1)}{2}$ terms, and therefore the regression and parameter vectors will be defined as :

$$\mathcal{X}_k = [x_k^2 \ \ 2x_k x_{k-1} \ \cdots \ x_{k-1}^2 \ \ 2x_{k-1}x_{k-2} \cdots \ x_{k-N+1}^2]^T \ , \tag{18}$$

and

$$\theta = [\theta^{(0)} \ \ \theta^{(1)} \ \dots \ \theta^{(\frac{N \times (N+1)}{2} - 1)}]^T \ , \tag{19}$$

respectively. All the above mentioned equations are still valid in this case, the only difference being the dimensions of θ, $\mathcal{X}$ and the matrix $E(\mathcal{X}\mathcal{X}^H)$. All Hermitian transposes can also be replaced by simple transposes in this case. Figure 2 shows the setup for bilinear equalization in the real case, where the $\frac{N \times (N+1)}{2}$ entries of θ are organized in N equalizers of respective lengths $N, N-1, \dots, 1$. Note the multichannel structure that allows also for a multichannel FTF algorithm as mentioned above.

3. Over-parameterized case

In this section we consider the case where the true channel is all-pole of order $N-1$ (AP$(N-1)$) and the equalizer FIR of order $M-1$ (FIR$(M-1)$), $M > N$. This case does not seem to

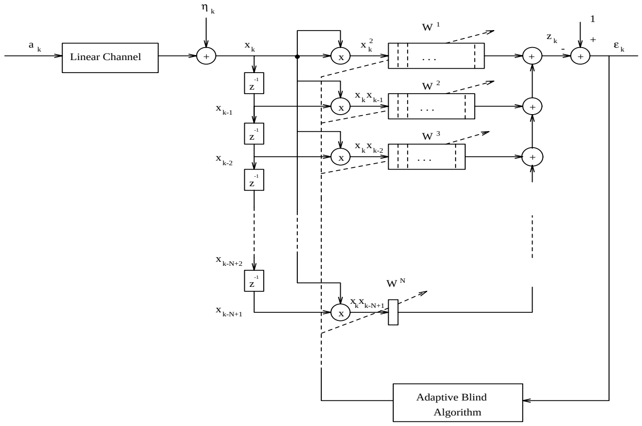

Figure 2: A bilinear blind equalization setup

be of practical interest, since in practice the transmission channel is a FIR filter and therefore an equalizer of any length will never exceed the order of the channel's inverse impulse response. However, it merits special attention because it gives insight in the behaviour of the bilinear method. Suppose that the equalizer is FIR(N). Then both equalizer settings $W^1 = [C^T \; 0]^T$ and $W^2 = [0 \; C^T]^T$ are zero forcing. If the corresponding bilinear parameter vectors are denoted by θ^1 and θ^2, respectively, then they will both satisfy (10). Moreover, any vector of the form $\alpha\theta^1 + \beta\theta^2$, where α and β are scalars such that $\alpha + \beta = 1$ will also satisfy (10):

$$E(\mathcal{X}\mathcal{X}^H)(\alpha\theta^1 + \beta\theta^2) = E\mathcal{X}, \alpha + \beta = 1 \quad . \tag{20}$$

In the general case $M = N + L$, the eq. (10) will be satisfied by any vector of the form:

$$\theta = \sum_{i=1}^{L} \alpha_i \theta^i, \quad with \quad \sum_{i=1}^{L} \alpha_i = 1 \quad , \tag{21}$$

where θ^i is the bilinear parameter vector corresponding to $W^i = [0_{1\times(i-1)} \; C^T \; 0_{1\times(L-i+1)}]^T$. This means that when the equalizer is *over-parameterized* w.r.t. the inverse of the channel impulse response, the matrix $E(\mathcal{X}\mathcal{X}^H)$ is singular, and as a consequence, the problem (8) has an infinite number of solutions. Running an adaptive algorithm like the one in (16) or in (17) will converge to one of these solutions. The same will happen if one uses a batch technique to estimate $E(\mathcal{X}\mathcal{X}^H)$ and $E(\mathcal{X})$ and a pseudo-inverse for the matrix inversion needed in (11). In all cases, the solution obtained will no longer correspond to a rank-1 matrix Θ: (14) and (15) will no longer yield the optimal ZF equalizer. Therefore the problem of ill-convergence appears (under a different form) also in the bilinear method. In the case of a FIR channel (which is a realistic one), one might think that the problem should not arise since the impulse response of the channel is theoretically of infinite length and thus an equalizer long enough should be able to approximate fairly well the ZF equalizer. However, as in practice the channel's inverse impulse response will be very close to zero out of a specific interval, if the equalizer's length is bigger than the number of samples in this interval, the same over-parameterization problem will exist and the matrix $E(\mathcal{X}\mathcal{X}^H)$ will be very ill-conditioned. This makes the choice of the equalizer length N in this case a problem of critical importance, as, unlike in conventional equalization, the longest possible length will not necessarily yield the best possible equalizer! In the next section we present a method to calculate the ZF equalizer from any solution of the form (21) when L is given.

4. A subspace fitting approach for the calculation of a ZF equalizer

Our task is to try to extract a ZF equalizer from the matrix Θ that corresponds to a vector θ of the form in (21). As will be shown, this will be possible due to the known specific structure of Θ. The channel inverse is assummed to be of length N and the equalizer of length $M = N + L$. Consider the eigenvalue decomposition of the matrix Θ:

$$\Theta = \sum_{i=1}^{M} \lambda_i V_i V_i^H \quad , \tag{22}$$

where the (real) eigenvalues λ_i are in descending order of magnitude and V_i is the eigenvector corresponding to λ_i. When $M = N(L = 0)$, we saw that the ZF equalizer can be found based on a rank-1 decomposition of Θ. When $L > 0$ we will try to determine the ZF equalizer based on a rank-(L+1) decomposition of Θ by the following subspace fitting approach: we first construct an extended equalizer vector of $M + L$ entries:

$$W^e = [w_{-L} \cdots w_{-1} \; w_0 \cdots w_{M-1}]^T \quad , \tag{23}$$

where $w_{-L} \cdots w_{-1}$ and $w_N \cdots w_{M-1}$ stand for the additional coefficients. If the length of the channel inverse is indeed N, then ideally $W^e = [0 \cdots 0 \; w_0 \cdots w_{N-1} \; 0 \cdots 0]^T$.

We then create a Toeplitz matrix $\mathcal{W}$ as follows:

$$\mathcal{W} = \begin{bmatrix} w_0 & w_{-1} & \cdots & w_{-L} \\ & & \ddots & \vdots \\ w_1 & w_0 & & \\ \vdots & \vdots & \ddots & w_0 \\ \vdots & \vdots & \ddots & \vdots \\ w_{M-1} & w_{M-2} & \cdots & w_{M-1-L} \end{bmatrix} \quad . \tag{24}$$

We will try to fit the matrix $\mathcal{W}$ to a subspace of the space $\mathcal{R}^M$ created by the first L eigenvectors of Θ. This fitting may be accomplished by minimizing the following criterion:

$$\min_{Q, W^e} ||\mathcal{W} - \mathcal{V}Q||_F^2 \quad , \tag{25}$$

where $||.||_F$ denotes the Frobenius norm of a matrix ($||A||_F^2 = tr(A^H A)$), $Q \in \mathcal{R}^{(L+1)\times(L+1)}$ and $\mathcal{V}$ is a matrix containing the $L+1$ first eigenvectors of Θ:

$$\mathcal{V} = [V_1 \cdots V_{L+1}] \quad . \tag{26}$$

Minimization w.r.t. Q only, yields:

$$Q = (\mathcal{V}^H\mathcal{V})^{-1}\mathcal{V}^H\mathcal{W} \quad . \tag{27}$$

The problem (25) now becomes:

$$\begin{array}{c} \min\limits_{W^e} ||P_{\mathcal{V}}^{\perp}\mathcal{W}||_F^2 \\ s.t. ||W^e||_2^2 = 1 \end{array} = \begin{array}{c} \min\limits_{W^e} tr\{\mathcal{W}^H P_{\mathcal{V}}^{\perp}\mathcal{W}\} \\ s.t. ||W^e||_2^2 = 1 \end{array} \quad . \tag{28}$$

Noting that $tr\{\mathcal{W}^H P_{\mathcal{V}}^{\perp}\mathcal{W}\} = tr(\mathcal{W}^H\mathcal{W}) - tr\{\mathcal{W}^H P_{\mathcal{V}}\mathcal{W}\}$ the problem (28) can be approximated by the problem:

$$\begin{array}{l} \max\limits_{W^e} \; tr\{\mathcal{W}^H\mathcal{V}(\mathcal{V}^H\mathcal{V})^{-1}\mathcal{V}^H\mathcal{W}\} \\ = \max\limits_{W^e} \; F(\mathcal{W}, \mathcal{V}) = tr\{\mathcal{W}^H\mathcal{V}\mathcal{V}^H\mathcal{W}\} \end{array} \quad . \tag{29}$$

The quantity $F(\mathcal{W}, \mathcal{V})$ can be written as:

$$\begin{array}{l} F(\mathcal{W}, \mathcal{V}) = W^{eH} \begin{bmatrix} 0_{L\times(L+1)} \\ \mathcal{V} \end{bmatrix} \begin{bmatrix} 0_{L\times(L+1)} \\ \mathcal{V} \end{bmatrix}^H W^e + \\ + W^{eH} \begin{bmatrix} 0_{L-1\times(L+1)} \\ \mathcal{V} \\ 0_{1\times(L+1)} \end{bmatrix} \begin{bmatrix} 0_{L-1\times(L+1)} \\ \mathcal{V} \\ 0_{1\times(L+1)} \end{bmatrix}^H W^e + \cdots + \\ + W^{eH} \begin{bmatrix} \mathcal{V} \\ 0_{L\times(L+1)} \end{bmatrix} \begin{bmatrix} \mathcal{V} \\ 0_{L\times(L+1)} \end{bmatrix}^H W^e \quad , \end{array} \tag{30}$$

which gives the following expression for (29):

$$\max_{W^e} \; W^{eH}(\sum_{i=1}^{L+1} \Omega_i \Omega_i^H) W^e \quad , \tag{31}$$

where $\Omega_i = [0^T_{(L-i+1)\times(L+1)} \ \ \mathcal{V}^T \ \ 0^T_{(i-1)\times(L+1)}]^T$. The solution to (31) is:

$$W^e = max \ \ eigenvector \ \ of \ \ \Lambda = \sum_{i=1}^{L+1} \Omega_i \Omega_i^H \ \ . \qquad (32)$$

When L is known, (32) will give the optimal ZF equalizer (in the absence of additive noise).

5. The influence of additive noise

When additive noise is present in the received signal, it will corrupt both the quantities $E(\mathcal{X}\mathcal{X}^H)$ and $E(\mathcal{X})$ and therefore the solution (11) will be biased and no longer correspond to a ZF equalizer (even in the absence of order mismatch). We will study the influence of additive white zero-mean Gaussian noise n_k in the real case (a_k, x_k, y_k are all real):

$$x_k = x'_k + n_k \ \ , \qquad (33)$$

where x'_k is the noiseless channel output. The fourth order moments involved in $E(\mathcal{X}\mathcal{X}^H)$ are:

$$\begin{array}{l} E(x^4_{k-i}) = m_{4,x'}(0,0,0) + 3\sigma^4_n + 6m_{2,x'}(0)\sigma^2_n \\ E(x^3_{k-i}x_{k-j}) = m_{4,x'}(0,0,|i-j|) + 3m_{2,x'}(|i-j|)\sigma^2_n \\ E(x^2_{k-i}x^2_{k-j}) = m_{4,x'}(0,|i-j|,|i-j|) + 2m_{2,x'}(0)\sigma^2_n + \sigma^4_n \\ E(x^2_{k-i}x_{k-j}x_{k-l}) = m_{4,x'}(0,|i-j|,|i-l|) + m_{2,x'}(|i-j|)\sigma^2_n \\ E(x_{k-i}x_{k-j}x_{k-l}x_{k-m}) = m_{4,x'}(|i-j|,|i-l|,|i-m|) \ \ , \end{array} \qquad (34)$$

If the noise variance σ^2_n is known, then the noise-free fourth order moments can be calculated as follows:

$$\begin{array}{l} E((x^2_{k-i} - 3\sigma^2_n)^2) - 6\sigma^4_n) = m_{4,x'}(0,0,0) \\ E((x^2_{k-i} - 3\sigma^2_n)x_{k-i}x_{k-j}) = m_{4,x'}(0,0,|i-j|) \\ E((x^2_{k-i} - \sigma^2_n)(x^2_{k-j} - \sigma^2_n)) = m_{4,x'}(0,|i-j|,|i-j|) \\ E((x^2_{k-i} - \sigma^2_n)x_{k-j}x_{k-l}) = m_{4,x'}(0,|i-j|,|i-l|) \\ E(x_{k-i}x_{k-j}x_{k-l}x_{k-m}) = m_{4,x'}(|i-j|,|i-l|,|i-m|) \ \ , \end{array} \qquad (35)$$

The second order noise-free moments can also be obtained as:

$$\begin{array}{l} E(x^2_{k-i}) - \sigma^2_n = m_{2,x'}(0) \\ E(x_{k-i}x_{k-j}) = m_{2,x'}(|i-j|) \ \ , \end{array} \qquad (36)$$

(in the above expressions $m_{4,x}(i,j,l) = E(x(k)x(k+i)x(k+j)x(k+l))$ and $m_{2,x}(i) = E(x(k)x(k+i))$). Therefore, the influence of additive noise can be theoretically completely eliminated, given knowledge of its variance. Similar results can be obtained in the complex case.

6. Experimental results

We have tested the bilinear method by means of computer simulations. In the case of all-pole noiseless channels of the same order as the equalizer, as expected, the optimal (ZF) equalizer was obtained with the help of (15) in all cases (irrepsective of initialization) by using either an adaptive (16), (17) or a batch technique (11), whereas the CMA converged as well to other equalizer settings for some initializations (ill-convergence). This verifies the fact that the problem of ill convergence is avoided by the bilinear method in this case. In the case of noiseless AP($N-1$) channels and an over-parameterized equalizer FIR($M-1$), $M = N + L > N$, the optimal (ZF) equalizer was also obtained with the help of (32) instead of (15). This of course implies knowledge of L. In the case of FIR channels, the method is sensible to the choice of the equalizer length, as already explained. However, for a given length, one can still use (32) for different values of L ranging from 0 to $M-2$, and then choose the best among the $M-1$ derived equalizers by evaluating for each the constant modulus criterion and choosing the one that best satisfies it. In our simulations, there was always one among the equalizers that sufficiently opens the system's eye, provided of course that the equalizer's length is long enough to be able to approximate well the channel's inverse. It was also observed that the influence of additive white Gaussian noise resulting in an SNR up to 20dB did not in general cause serious damage to the obtained solution. For some simulation results, see also [5].

7. Further discussion

We have also adapted this bilinear approach to the framework of blind fractionally-spaced equalization (BFSE) [6]. The motivation was that in this case, ZF equalization of FIR channels is possible with FIR equalizers (!) and therefore the issue of inverting the channel exactly with an FIR equalizer does not pose a problem. However, in BFSE, multiple ZF equalizers exist not only in the over-parameterized case but even in perfect parameterization and the bilinear method will again be insensitive to a convex combination of these ZF equalizers. Moreover, the various ZF equalizers are not simply shifted versions of one another in the blind fractionally spaced setup of [6], but their dependence is more complicated and depends on the unknown channel parameters. So the same problem of singularity of $E(\mathcal{X}\mathcal{X}^H)$ that arises in the over-parameterized symbol-rate case, appears in the exactly parameterized fractionally spaced case as well. A remedy to this problem would be to choose a number of equalizer parameters slightly less (1 less for example) than exactly needed for ZF equalization. Then a unique solution is found, but is suboptimal. Further investigation of these aspects is the object of ongoing research.

Acknowledgement

After the publication [5] we realized that the principle of the bilinear cost function had already appeared in [7] in order to analyze the behaviour of CMA. We are still however not aware of any paper that discusses the acquisition of the ZF equalizer (15),(32) and the critical role of over-parameterization (21) in the behaviour of the bilinear method.

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] **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, pp. 245-250, July 14-16, 1993.

[4] **D. T. M. Slock**, *PhD thesis*, Stanford, 1989.

[5] **C. B. Papadias and D. T. M. Slock**, "A Bilinear Approach to Constant Modulus Blind Equalization," in *ATHOS, ESPRIT Basic Research Working Group 6620, System Identification and High Order Statistics*, I3S, CNRS, Sophia Antipolis, September 20-21, 1993.

[6] **D. T. M. Slock**, "Blind Fractionally-Spaced Equalization, Perfect-Reconstruction Filter Banks and Multichannel Linear Prediction", *Proc. ICASSP 94 Conf.*, Adelaide, Australia, April 1994.

[7] **H. Jamali, S. L. Wood, R. Cristi**, "Experimental Validation of the Kronecker Product Godard Blind Adaptive Algorithms," *Asilomar '92*, Pacific Grove, CA, 1992.