ALGORITHMS AND STRUCTURES FOR SOURCE SEPARATION BASED ON THE CONSTANT MODULUS PROPERTY*

J.R. Cerquides, J.A. Fernández-Rubio
Signal Theory and Communications Department, Polytechnic University of Catalonia
Módulo D5, Campus Nord UPC, C/Gran Capitán s/n, 08034 Barcelona, SPAIN
Tel.:+34-3-4015938, Fax:+34-3-4016447
E-mail: ramon@tsc.upc.es

ABSTRACT

We propose two structures and theirs associated algorithms designed to solve the blind source separation problem in the presence of noise and interferences. Both structures exploit the non convexity of the Constant Modulus cost function, finding its multiple local minima. A convergence analysis shows that both schemes achieve the desired solution, separately extracting the sources of interest while rejecting noise and interferences, provided that they do not share the constant modulus property.

1. INTRODUCTION

Spatial Domain Multiplexing Access (SDMA) techniques may theoretically allow two or more users in the same cell to simultaneously employ the same frequency band and temporal slot, provided that they are not closely enough. This fact would imply an enormous increase of system capacity, specially in the mobile communications field. To accomplish this objective we need to equip the base station with the ability of generating spatial diversity, which implies to substitute the antenna by an array of sensors conveniently distributed. The problem we address here is how to separately extract the incoming signals, knowing that they share both time and frequency band.

To do this we propose two architectures, serial and parallel. Both schemes try to exploit the non convexity of the constant modulus cost function defined in [1,2]. The proposed structures are two approaches to the problem of finding the minima of the Constant Modulus cost function.

The paper is organised as follows: in section 2 we review some interesting results about the nature of the minima of the Constant Modulus cost function. Sections 3 and 4 describe the proposed structures and its associated algorithms. A brief convergence analysis shows that both structures achieve the desired signal extraction. In section 5 we test the behaviour of the proposed algorithms in an hostile environment, showing the superior performance of the parallel version.

* This work has been partially supported by the National Research Plan of Spain CICYT, under Grant TIC-95-1022-C05-01

2. THE CONSTANT MODULUS ARRAY

The Constant Modulus Array is a beamforming technique which exploits the constant envelope property exhibited by many communications signals. The algorithm uses an stochastic gradient technique to minimise a non convex cost function of the weights, J(**w**) defined as:

$$J(\mathbf{w}) = \left(\left|y[n]\right|^2 - 1\right)^2 \tag{1}$$

where y[n] is the array output, obtained as a linear combination of the input signal vector **x**[n],

$$y[n] = \mathbf{w}^H \mathbf{x}[n] \tag{2}$$

As stated before, J(**w**) is not a convex function. Thus, it is possible to find signal scenarios which lead to the presence of several local minima of the cost function. Some results about the nature of J(**w**) have been obtained by the analysis performed by Treichler et al. in restrictive signal environments[3,4,5]. In a general framework with M signals arriving together at an array with N>M elements it is difficult to extract conclusions about the number of singular points and its nature. However, from the analysis detailed in [6] we have found that the function J(**w**) exhibits, at least, as many minima as the number of incoming signals with kurtosis[1] unity and moderate to high SNR[2]. Furthermore, the weight vector associated to

[1] The kurtosis of a signal s[n] is defined as

$$k_s = \frac{E\left\{\left|s[n]\right|^4\right\}}{E\left\{\left|s[n]\right|^2\right\}^2}$$

[2] Let $\mathbf{d}_i$ the generalized steering vector of the i-th source, $s_i[n]$, and $\mathbf{R}_{xx}$ the correlation matrix of the incoming data, $\mathbf{R}_{xx}=E\{\mathbf{x}[n]\mathbf{x}^H[n]\}$. The moderate to high SNR condition is accomplished if

$$\frac{\left|\mathbf{w}^H\mathbf{d}_i\right|^2 E\left\{\left|s_i[n]\right|^2\right\}}{\mathbf{w}^H\mathbf{R}_{xx}\mathbf{w} - \left|\mathbf{w}^H\mathbf{d}_i\right|^2 E\left\{\left|s_i[n]\right|^2\right\}} > 10$$

where $\mathbf{w} = \mathbf{R}_{xx}^{-1}\mathbf{d}_i$ is the Wiener solution for the i-th source.

each minimum is proportional to the Wiener solution for the extraction of the associated signal.

In other words, if the Wiener solution for an incoming constant modulus signal yields an output SINR greater than 10 dB, there exists a minimum of the Constant Modulus cost function J(**w**) associated to the optimal extraction of this signal. In the following paragraphs we will describe two architectures designed to find the minima of J(**w**).

3. THE SERIAL STRUCTURE

3.1 Description

The serial structure, whose scheme is shown in figure 1, was initially introduced by Shynk[7], who studied its behaviour and convergence properties. The underlying idea is to let the first beamformer to freely choose and extract one of the sources, while rejecting the rest. The output signal is then multiplied by an appropriate weight vector and subtracted from the initial snapshots. The result is that all information related to this source disappears from the snapshots used as input to the following stage. Thus, the convergence of subsequent stages over different sources is guaranteed, making unable to extract twice the same signal.

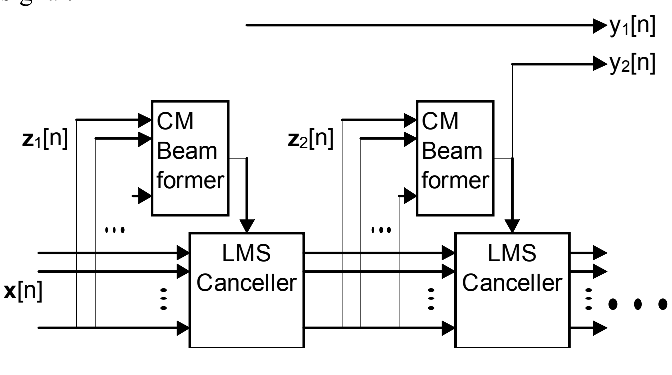

Figure 1–Scheme of the serial structure

3.2 Algorithm

Both beamformer and canceller weight vectors are updated iteratively according to normalised versions of CM and LMS algorithms. Thus, the adaptation equations for the i-th (i=1..M) stage are:

$$\mathbf{v}_i[n+1]=\mathbf{v}_i[n]-\mu_b\left(\left|y_i[n]\right|-1\right)\mathrm{sgn}\left(y_i^*[n]\right)\frac{\mathbf{z}_i[n]}{\left\|\mathbf{z}_i[n]\right\|^2} \qquad (3)$$

$$\mathbf{c}_i[n+1]=\left(1-\mu_c\right)\mathbf{c}_i[n]+\mu_c\frac{\mathbf{z}_i[n]}{y_i[n]} \qquad (4)$$

where $\mathbf{v}_i[n]$ characterises the vector associated to the i-th beamformer, $\mathbf{c}_i[n]$ is the i-th canceller weight vector, $\mathbf{z}_i[n]$ and $y_i[n]$ are the input and output of the i-th beamformer, and μ_c, μ_b are the normalised step-size parameters for both adaptive systems.

3.3 Convergence properties

In a general framework, the input snapshot **x**[n] can be described as:

$$\mathbf{x}[n]=\sum_{m=1}^{M}\mathbf{d}_m s_m[n]+\mathbf{i}[n] \qquad (5)$$

where M is the total number of signals of interest (SOIs), $s_m[n]$ is the m-th SOI, $\mathbf{d}_m$ is the generalised steering vector for this source and **i**[n] models the undesired contribution of noise and interferences. According to the results obtained in [6] the first beamformer will converge to the optimal extraction of one of the SOIs. So, once the convergence has been achieved, $\mathbf{v}_1[n]$ will take the value:

$$\mathbf{v}_1[n]=\alpha_1\mathbf{R}_{xx}^{-1}\mathbf{d}_1 \qquad (6)$$

where it is assumed, without loss of generality, that the beamformer extracts the SOI number 1. α_1 is an appropriate scaling factor. The canceller will reach its stationary state when $\mathbf{c}_1$ verifies:

$$\mathbf{c}_1=\frac{E\left\{\mathbf{z}_1[n]y_1^*[n]\right\}}{E\left\{\left|y_1[n]\right|^2\right\}}=\frac{\mathbf{d}_1}{\alpha_1^*\mathbf{d}_1^H\mathbf{R}_{xx}^{-1}\mathbf{d}_1} \qquad (7)$$

Putting all together, it is possible to demonstrate that $\mathbf{z}_2[n]$ does not contain information about $s_1[n]$. To check this, it is enough to evaluate:

$$\begin{aligned}&E\left\{\mathbf{z}_2[n]s_1^*[n]\right\}=E\left\{\left(\mathbf{x}[n]-\mathbf{c}_1y_1[n]\right)s_1^*[n]\right\}=\\&=\left(\mathbf{d}_1-\mathbf{c}_1\alpha_1^*\mathbf{d}_1^H\mathbf{R}_{xx}^{-1}\mathbf{d}_1\right)E\left\{\left|s_1[n]\right|^2\right\}=\mathbf{0}\end{aligned} \qquad (8)$$

So, the second beamforming stage does not receive any information related with $s_1[n]$, avoiding the possibility of extracting again this signal. Extending this analysis to the M stages of the structure will lead us to the previously claimed result: the ability of separately extract the M incoming SOIs.

4. THE PARALLEL STRUCTURE

4.1 Description

In the sequential scheme shown in figure 1, all paths propagate in a forward manner. This allows the i-th stage of the system to take benefit from the task carried out by previous stages. However, on the contrary it is not true. It seems reasonable to introduce additional feedback paths to get a symmetric structure, as shown in figure 2.

The relations between the variables that describe the structure are now more complex than in the serial version. The input to the i-th beamformer, $\mathbf{z}_i[n]$ can be written as:

$$\mathbf{z}_i[n]=\mathbf{x}[n]-\sum_{\substack{j=1\\ j\neq i}}^{M}\mathbf{c}_j y_j[n] \qquad (9)$$

where **x**[n] is the received snapshot, $\mathbf{c}_j$ is the canceller vector for the j-th signal, and $y_j[n]$ is the output of the j-th stage:

$$y_j[n] = \mathbf{v}_j^H \mathbf{z}_j[n] \quad (10)$$

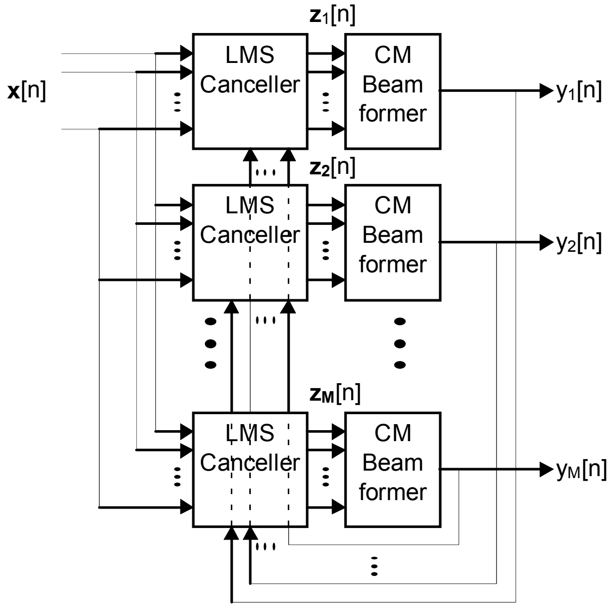

Figure 2– Scheme of the parallel structure

Imposing the additional condition:

$$\mathbf{v}_i^H \mathbf{c}_i = 1 \quad \forall i = 1..M \quad (11)$$

it is possible to simplify the solution of the system of M equations that results when we substitute (9) into (10). This solution takes the following form:

$$\mathbf{y}[n] = \left(\mathbf{V}^H \mathbf{C}\right)^{\#} \mathbf{V}^H \mathbf{x}[n] \quad (12)$$

where $\mathbf{y}[n]$ is the M×1 output vector,

$$\mathbf{y}[n] = \begin{bmatrix} y_1[n] & y_2[n] & \cdots & y_M[n] \end{bmatrix}^T \quad (13)$$

$\mathbf{V}$ is an N×M matrix containing information about the beamformers,

$$\mathbf{V} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_M \end{bmatrix} \quad (14)$$

$\mathbf{C}$ is another N×M matrix containing information about the cancellers,

$$\mathbf{C} = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_M \end{bmatrix} \quad (15)$$

and the superscript $^{\#}$ denotes the pseudoinverse. As the values of $\mathbf{v}_i$ and $\mathbf{c}_i$ will be continuously changing we cannot always guarantee the existence of the inverse of $\mathbf{V}^H\mathbf{C}$, so we use its pseudoinverse instead. However, in practice, we have not found problems related with the inversion of this matrix. Let's describe now the algorithm employed to update both beamformer and canceller weights.

4.2 Algorithm

For each new sample $\mathbf{x}[n]$ we now propose the following steps:

1. Compute $\mathbf{y}[n]$ using equation (12).
2. Compute $\mathbf{z}_i[n]$, i=1..M, using equation (9) .
3. Update $\mathbf{v}_i$, i=1..M following a normalised CM adaptation equation (as described in eq. (3)).
4. Update $\mathbf{c}_i$, i=1..M using:

$$\mathbf{c}_i[n+1] = (1-\mu_c)\mathbf{c}_i[n] + \mu_c \frac{\mathbf{x}[n]}{y_i[n]} \quad (16)$$

5. Normalise $\mathbf{c}_i$ to verify eq. (11):

$$\mathbf{c}_i[n+1] = \frac{\mathbf{c}_i[n+1]}{\mathbf{v}_i^H[n+1]\mathbf{c}_i[n+1]} \quad (17)$$

6. Back to step 1.

4.3 Convergence properties

As all the operations performed by the elements of the structure are linear, it is possible to describe the relation between $\mathbf{x}[n]$ and $\mathbf{z}_i[n]$ by means of a matrix $\mathbf{M}_i$. Thus,

$$\mathbf{z}_i[n] = \mathbf{M}_i \mathbf{x}[n] \quad (18)$$

As before we can assume, without loss of generality, that the first beamformer will converge when vector $\mathbf{v}_1$ verifies:

$$\mathbf{M}_1 \mathbf{R}_{xx} \mathbf{M}_1^H \mathbf{v}_1 = \alpha_1 \mathbf{M}_1 \mathbf{d}_1 \quad (19)$$

Rearranging terms in eq. (9) it is possible to relate the first and j-th path through:

$$\mathbf{z}_j[n] - \mathbf{c}_j y_j[n] = \mathbf{z}_1[n] - \mathbf{c}_1 y_1[n] \quad (20)$$

Computing now $E\{\mathbf{z}_i[n]s_1^*[n]\}$ and substituting $\mathbf{c}_1$ and $\mathbf{c}_j$ by its final values (eq. 7), we get the following equation, which must be accomplished when convergence is attained:

$$\mathbf{M}_j \mathbf{d}_1 - \frac{\mathbf{M}_j \mathbf{R}_{xx} \mathbf{M}_j^H \mathbf{v}_j \mathbf{v}_j^H \mathbf{M}_j}{\mathbf{v}_j^H \mathbf{M}_j \mathbf{R}_{xx} \mathbf{M}_j^H \mathbf{v}_j} \mathbf{d}_1 = \mathbf{0} \quad (21)$$

which has only a valid solution, given by:

$$\mathbf{M}_j \mathbf{d}_1 = \mathbf{0} \quad (22)$$

So, signal $s_1[n]$ is extracted by the first channel and rejected by the rest. Applying iteratively these results to the remaining M-1 dimensional problem we can conclude that the final state of the structure will be reached when the M channels converge to the separate extraction of the M incoming SOIs.

5. SIMULATION

To compare the behaviour of the serial and parallel versions of the Multiple Constant Modulus Algorithm (MCMA) we have chosen the environment shown in table 1, where 4 signals arrive at a 6 sensor, linear, equally spaced array (separation=$\lambda/2$). Three of them are SOIs and the fourth signal is a strong interference.

Signal	Type of signal	Input SNR	Angle of arrival
#1	4-PSK	5 dB	-30º
#2	4-PSK	5 dB	20º
#3	4-PSK	5 dB	40º
#4	Tone (f=0.1)	20 dB	-10º

Table I – Description of the environment employed in the simulation.

The normalised step-size parameters μ_b=0.01, μ_c=0.001 were chosen. Figure 3 shows the evolution of the SINR ratio for both structures in the three channels.

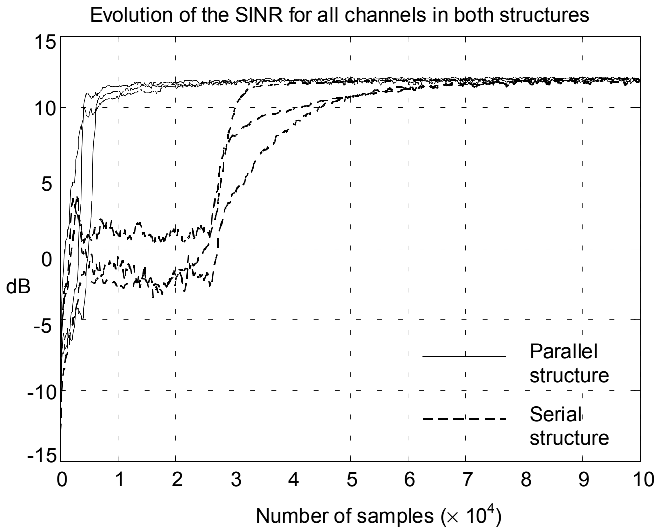

Figure 3 – Comparison of the evolution of the output SINR

The comparison offers no doubts. The parallel structure is about ten times faster than its serial version. This enormous improvement in the convergence speed compensates the amount of computational load added to the algorithm. But there is not only this difference between both schemes. Although it cannot be appraised in the figure, the output SINR offered by the parallel algorithm is slightly better. And there is also an improvement in the rejection of co-channel interferences. In a forthcoming paper we will show other interesting features of the parallel structure: its enhanced resolution and the capability of extract and follow the original sources when the signal environment is nonstationary.

6. CONCLUSIONS

Two structures based on the multiple minima property of the Constant Modulus cost function have been proposed. Both adaptive schemes are composed by CM arrays and signal cancellers, whose respective weight vectors are updated following stochastic steepest descent algorithms. A convergence analysis shows the ability of both structures to achieve the desired source separation without a great amount of computational load. Finally, simulations demonstrate that the parallel structure, where all stages work on a collaboration basis, exhibits a faster convergence and better performance than its sequential version.

7. REFERENCES

[1]Godard, D.N. "Self-Recovering equalization and carrier tracking in two-dimensional data communication systems", *IEEE Transactions on Communications*, vol. COMM-28, pp. 1867-1875, November, 1980.

[2]Treichler, J.R., Agee, B.G., "A new approach to multipath correction of constant modulus signals", *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-31, pp. 459-472, April, 1983.

[3]Treichler, J.R., Larimore, M.G., "Convergent behavior of the constant modulus algorithm", *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, pp. 13-16, Boston, 1983

[4]Smith, J.O., Friedlander, B., "Global Convergence of the Constant Modulus Algorithm", *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, pp. 30.5.1-30.5.4, Tampa, 1985.

[5]Agee, B.G., "Convergent behavior of modulus restoring adaptive arrays in Gaussian interference environments", *Proceedings of the Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, November 1988.

[6] Cerquides, J.R., *Contribución a la conformación ciega de haz en comunicaciones*, *Ph. D., Appendix A*, Polytechnic University of Catalonia, Barcelona, May, 1996.

[7] Shynk, J.J., Gooch, R.P., "Convergence properties of the Multistage CMA Adaptive Beamformer", *Proceedings of the Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, November 1993.