# Regularized Semi-blind Estimation of spatio-temporal Filter coefficients for mobile radio communications

**A.M.Kuzminskiy *†, L.Féty*, P.Forster*, S.Mayrargue‡**

* Laboratoire d'Electronique, CNAM, 292, rue Saint-Martin, 75141 Paris Cedex 03, FRANCE,
tel: 33 01 40 27 24 85, fax: 33 01 40 27 27 79, e-mail:kuzminsk@cnam.fr

‡FRANCE TELECOM CNET 38-40 rue du Général Leclerc 92131 Issy-les-Moulineaux, FRANCE,
tel. 33 01 45 29 60 51, fax. 33 01 45 29 41 94, e-mail Sylvie.Mayrargue@issy.cnet.fr

†Odessa State Polytechnic University, DB "Diskret", 270044, av.Shevchenko 1, Odessa, UKRAINE,
tel: 38 04 82 28 87 25, e-mail:kuzm@diskret.opu.odessa.ua

RÉSUMÉ

On étudie l'adaptation des coefficients du filtre spatio-temporel utilisé à la station de base pour l'égalisation du canal et la réjection des brouilleurs, quand la longueur de la séquence d'apprentissage est insuffisante pour faire appel aux estimateurs classiques. Un critére semi-aveugle et l'algorithme de traitement correspondant sont proposés, en exploitant la présence d'une séquence d'apprentissage courte et la propriété de module constant des signaux. En fait, c'est une nouvelle version de l'estimateur régularisé, adapté aux particularités des radiocommunications mobiles numé riques.

ABSTRACT

The solved problem is the adjustment of the spatio-temporal filter at the base station, for channel equalization and jammer rejection, when the length of a training sequence is not sufficient to use standard estimators. A semi-blind criterion and processing algorithm are proposed, which exploit the presence of a short training sequence and the constant modulus property of the signals. It is a new version of the regularized estimator, which is adapted to the features of digital mobile radio communications.

## Introduction

In digital mobile radio communications the data are transmitted in bursts and a training sequence of short duration is attached to each burst. The length of this sequence (26 symbols for the GSM) may not be sufficient to adjust the coefficients of the spatio-temporal filter to equalize the channel and reject the jammers when an array of sensors together with temporal filters are used in the base station. This difficulty can be overcome with the help of suboptimal [1] or blind algorithms [2, 3]. Both approaches have known disadvantages. For example, a global convergence of the popular blind fractionally spaced constant modulus (CM) algorithm is established only for an infinite number of data (mathematical expectation in CM criterion) [3 and others]. It has been pointed out in [4,5] that combining training and blind techniques can be effective, and semi blind criterions and algorithms for single input multiple output channel identification based on maximum likelihood (ML) principle have been proposed. The necessity of complete modelling for ML approach limits the aplication area of these algorithms.

A semi-blind approach based on a least squares (LS) criterion regularized by means of the CM function is proposed here to find the coefficients of the spatio-temporal filter in the general case, with jammers and unknown lengths of propagation channels. It is shown that it allows to reduce the dimension of the optimized CM function by the length of the training sequence. An algorithm for the minimization of the proposed criterion is derived. Its efficiency is demonstrated by simulations, in situations where the length of the training sequence is not sufficient to use the standard regularized LS estimator and the limited volume of data in one burst is not sufficient to use the least squares CM algorithm (LSCMA) [6].

## Problem formulation

The signal model and the general spatio-temporal filter structure are shown in Fig.1.The notations are the following: $K$: number of antenna array elements; $s_n$: desired signal; $M$: number of jammers $d_{in}$, $i = 1...M$; $x_{in}$, $i = 1...K$: antenna array outputs; $\xi_{ln}$, $l = 1...K$: additive uncorrelated noise with variance $\sigma_\xi^2$; $\mathbf{G}$: propagation channels; $\hat{s}_{n-D} = \mathbf{W}^*\mathbf{X}_n$: desired signal estimator; $\mathbf{X}_n^T = \{\mathbf{X}_{1n}^T, ..., \mathbf{X}_{Kn}^T\}$: $(KL \times 1)$ input signal vector, where $\mathbf{X}_{in}^T = \{x_{in}, ..., x_{i(n-L+1)}\}$ for $i = 1...K$; $L$: number of coefficients of FIR filter in each spatial channel $\mathbf{W}_i^T = \{w_{i1}, ..., w_{iL}\}$; $\mathbf{W}^T = \{\mathbf{W}_1^T, ..., \mathbf{W}_K^T\}$: $(KL \times 1)$ vector of weight coefficients; $D$: delay. All signals assumed zero mean.

The features of this signal model in mobile radio communications application are the following:

1. Propagation channels $\mathbf{G}$ can be approximated by FIR filters of length $L_g$ [3]. Both $\mathbf{G}$ and $L_g$ are unknown.

2. The desired signal $s_n$ has the following temporal structure: data are transmitted in bursts of length $N_b$; the training sequence of length $N < N_b$ is transmitted inside each burst.

So, we have $n = 1...N_b$ for all signals in Fig.1, excluding $n = n_s...(n_s + N)$ for the training sequence, where $n_s$ is the starting training symbol. Propagation channels can be assumed to be stable during one burst.

3. The desired signal has CM property.

4. Jammers are independent of $s_n$. They may have either the same structure (multi-users) or another one.

The classical approach to the estimation of the weight vector is to minimize the LS criterion on the learning interval

$$\hat{\mathbf{W}}_N = \arg\min_{\mathbf{W}}\{N^{-1}\sum_{n=n_s}^{n_s+N} \mid s_{n-D} - \mathbf{W}^*\mathbf{X}_n \mid^2\}. \quad (1)$$

The solution of the problem (1), given $N \geq KL$ linearly independent input vectors $\mathbf{X}_n$ is

$$\hat{\mathbf{W}}_N = \hat{\mathbf{R}}^{-1}\hat{\mathbf{P}}, \quad (2)$$

where $\hat{\mathbf{R}} = N^{-1}\sum_{n=n_s}^{n_s+N}\mathbf{X}_n\mathbf{X}_n^*$, $\hat{\mathbf{P}} = N^{-1}\sum_{n=n_s}^{n_s+N} s_{n-D}^*\mathbf{X}_n$

Since [7] the standard way to find the estimator when $N < KL$ is to replace (1) by the regularized criterion

$$\hat{\mathbf{W}}_{RN} = \arg\min_{\mathbf{W}}\{\frac{1}{N}\sum_{n=n_s}^{n_s+N} \mid s_{n-D} - \mathbf{W}^*\mathbf{X}_n \mid^2 + \delta\mathbf{W}^*\mathbf{W}\}, \quad (3)$$

$$\hat{\mathbf{W}}_{RN} = (\hat{\mathbf{R}} + \delta\mathbf{I})^{-1}\hat{\mathbf{P}}, \quad (4)$$

where $\delta > 0$ is a regularization coefficient. The performance of the regularized estimator (4) is studied in [7] in the spatial processing case and for uncorrelated input vectors $\mathbf{X}_n$.

The regularized LS estimator (4) can be applied to general signals. It does not reflect the signal features pointed out above. The problem is to find another regularized estimator which takes into account the features of the digital mobile radio communications environment, and to compare its performance (needed length of the training sequence) with the known solution (4).

# Semi-blind optimization criterion

An alternative regularization of the basic LS criterion (1) is proposed below, which takes into account the CM property of the desired signal

$$\hat{\mathbf{W}}_{sb} = \arg\min_{\mathbf{W}}\{N^{-1}\sum_{n=n_s}^{n_s+N} \mid s_{n-D} - \mathbf{W}^*\mathbf{X}_n \mid^2 +$$

$$\delta_b N_b^{-1}\sum_{n=1}^{N_b}(\mid \mathbf{W}^*\mathbf{X}_n \mid -1)^2\}. \quad (5)$$

where $\delta_b$ is a regularization coefficient.

This is the semi-blind criterion for the adjustment of the spatial-temporal filter shown in Fig.1 which uses both the training sequence and the information data of one burst. Using different $\delta_b$, one can get different versions of the general criterion (5):

- the choice $\delta_b \to \infty$ leads to the standard CM criterion

$$\hat{\mathbf{W}}_b = \arg\min_{\mathbf{W}}\{N_b^{-1}\sum_{n=1}^{N_b}(\mid \mathbf{W}^*\mathbf{X}_n \mid -1)^2\}; \quad (6)$$

- the choice $\delta_b \to 0$ leads to the following constrained optimization problem

$$\hat{\mathbf{W}}_{sb} = \arg\min_{\mathbf{W}\in\Omega}\{N_b^{-1}\sum_{n=1}^{N_b}(\mid \mathbf{W}^*\mathbf{X}_n \mid -1)^2\}, \quad (7)$$

where $\Omega$ is the set of all vectors $\mathbf{W}$ which satisfy the following linear equation

$$\hat{\mathbf{R}}\mathbf{W} = \hat{\mathbf{P}}. \quad (8)$$

with an infinite number of solutions when $rank(\hat{\mathbf{R}}) < KL$.

So, the two limiting cases of (5) can be used for estimating of $\mathbf{W}$. This situation makes the criterion (5) significantly different from (3) because the case $\delta \to \infty$ in (3) leads to the white noise solution for any input signals. That means at least that the strategy of selecting the regularization parameters must be different for these criteria.

The case $\delta_b \to 0$ is interesting because it helps to understand the area of applicability of the proposed approach. Indeed, all solutions of (8) can be expressed as follows

$$\mathbf{W} = \mathbf{W}_p + \sum_{i=1}^{KL-N} v_i\mathbf{U}_i, \quad (9)$$

where $\mathbf{W}_p$ is a particular solution of (8), for example $\mathbf{W}_p = \hat{\mathbf{R}}^{\#}\hat{\mathbf{P}}$, where $\hat{\mathbf{R}}^{\#}$ is the pseudoinverse of $\hat{\mathbf{R}}$; $\mathbf{U}_i$, $i = 1...KL - N$ are $(KL \times 1)$ basis vectors of the nullspace of matrix $\hat{\mathbf{R}}$; $\mathbf{V} = \{v_1...v_{KL-N}\}$ is a new $((KL - N) \times 1)$ weight vector. Substituting (9) into (7) we get the new reduced dimension CM type optimization criterion

$$\hat{\mathbf{V}}_{sb} = \arg\min_{\mathbf{V}}\{N_b^{-1}\sum_{n=1}^{N_b}(\mid y_{pn} + \mathbf{V}^*\mathbf{U}^*\mathbf{X}_n \mid -1)^2\}, \quad (10)$$

where $y_{pn} = \hat{\mathbf{P}}^*\hat{\mathbf{R}}^{\#}\mathbf{X}_n$, $\mathbf{U} = \{\mathbf{U}_1...\mathbf{U}_{KL-N}\}$. The final weight vector $\hat{\mathbf{W}}_{sb}$ can be calculated in accordance with (9)

$$\hat{\mathbf{W}}_{sb} = \hat{\mathbf{R}}^{\#}\hat{\mathbf{P}} + \sum_{i=1}^{KL-N}\hat{v}_{sb\,i}\mathbf{U}_i. \quad (11)$$

It is worth emphasizing that a similar limit situation exists for the standard regularized LS criterion (3): $\delta \to 0$ leads to $\hat{\mathbf{W}}_{RN} = \hat{\mathbf{R}}^{\#}\hat{\mathbf{P}}$. This case is adequate for both criteria (3) and (5) when $\sigma_\xi^2 = 0$. The usual choice of $\delta$ in (4) with noise is $\delta \simeq \sigma_\xi^2$ [7]. The question about the optimal choice of $\delta_b$ in (5) in the presence of noise requires a separate study. In this paper it will be considered by simulations.

The most interesting aspect in the problem formulation (7), (8) is the following: the semi-blind criterion (5) is in fact (exactly in the noiseless case when $\delta_b \to 0$) the modified CM criterion with a dimension reduced by the length of the training sequence. This dimension decrease for fixed volume of data opens the possibility to use CM criterion in situations which are adequate to mobile radio communications when spatio-temporal filter coefficients have to be estimated for each burst of data with the short training sequence inside.

## Optimization algorithm

After computation of the gradient and the Hessian of the optimization function in (5) we get the following Gauss-Newton type off-line optimization algorithm

$$\hat{\mathbf{W}}_{sb}^{k+1} = \hat{\mathbf{W}}_{sb}^{k} + (\hat{\mathbf{R}} + \delta_b \hat{\mathbf{R}}_b)^{-1}[\hat{\mathbf{R}}\hat{\mathbf{W}}_{sb}^{k} - \hat{\mathbf{P}} + \delta_b(\hat{\mathbf{R}}_b - \hat{\mathbf{R}}_{bb}(\hat{\mathbf{W}}_{sb}^{k}))\hat{\mathbf{W}}_{sb}^{k}], \quad (12)$$

where $\hat{\mathbf{R}}_b = N_b^{-1} \sum_{n=1}^{N_b} \mathbf{X}_n \mathbf{X}_n^*$,

$\hat{\mathbf{R}}_{bb}(\hat{\mathbf{W}}_{sb}^{k}) = N_b^{-1} \sum_{n=1}^{N_b} \mid \hat{\mathbf{W}}_{sb}^{k*} \mathbf{X}_n \mid^{-1} \mathbf{X}_n \mathbf{X}_n^*$.

Using zero initialization for reduced other weight vector $\mathbf{V}$ in (11) we get the following initial vector for the algorithm (12) $\hat{\mathbf{W}}_{sb}^{0} = \hat{\mathbf{R}}^{\#}\hat{\mathbf{P}}$ or for the general case with noise

$$\hat{\mathbf{W}}_{sb}^{0} = (\hat{\mathbf{R}} + \delta \mathbf{I})^{-1}\hat{\mathbf{P}}. \quad (13)$$

Consequently, the regularized semi-blind algorithm for an estimation of spatial-temporal filter coefficients is given by equations (12) and (13).

The particular cases of this algorithm are LSCMA [6] for an undefined initialization and $\delta_b \rightarrow \infty$ , and semi-blind LSCMA initialized by the training sequence (13) for $\delta_b \rightarrow \infty$ only.

## Simulation results

**Simulation scenario**. Independent random signals and jammers are of the type $(\pm 1 \pm j)/\sqrt{2}$. FIR channels filters have random complex coefficients with zero mean and unit variance. The simulation parameters are: $K = 5$, $L_g = 5$, $D = L$, $\sigma_\xi^2 = 0.08$, $n_s = 10$. Parameters $M$, $L$, $N_b$, $N$ are variable.

**Experiment 1**. Let us consider the behavior of LSCMA when the length of a burst is $N_b = 150$. It is known that Gauss-Newton type algorithms (LSCMA is this type of off-line algorithm) have a local convergence. So, it is possible to expect, that for a sufficiently large number of iterations, for some initialization the resulting weight vector for LSCMA corresponds to a local minimum of the CM criterion $J_{CM}(N_b)$. The belonging of the resulting vector to the desired (adequate to the CM signals or their delays) or to an undesired minimum can be found by the rule $J_{CM}(N_c) < \eta$, where $N_c \gg N_b$ and $\eta$ is some threshold. The estimations of the frequency of hitting on a desired CM minimum are presented in Table 1 for $N_c = 1000$, $\eta = 0.015$ and the following stopping rule $\|\hat{\mathbf{W}}^{k+1} - \hat{\mathbf{W}}^{k}\|^2/\|\hat{\mathbf{W}}^{k}\|^2 < 10^{-4}$. The number of realizations is 50 with 50 random initializations for each one.

One can see from Table 1 that $N_b = 150$ is practically sufficient for a global convergence only for a few adjustable coefficients in very simple environment without jammers.

**Experiment 2**. The other possibility to get one of the desired weight vectors by LSCMA is to increase the number of processed symbols. We get 9.7% for $N_b = 500$ and 98.1% successful trials for $N_b = 10000$ in the conditions of the last column in Table 1 (with 0.1% successful trials for $N_b = 150$).

The constellation pictures in Fig.2 - 4 for this environment reflect the typical situation for presented experiments when most of the resulting weight vectors are "good" CM solutions without any relations with initial CM signals. One can see that the "quality" of spurious CM solutions decreases with increasing of the data volume $N_b$; for some undefined large $N_b$, spurious solutions have practically disappeared.

Conditionally speaking the known theory about globally converged fractionally spaced CMA [3 and others] deals with Fig.4 and the first column in Table 1. But our problem corresponds to Fig.2 and the last column in Table 1: in that case, LSCMA fails while our algorithm is successful as shown below.

**Experiment 3**. The same data as in the last column of Table 1 are used in Fig.5 for the algorithm (12), (13) with variable $N$. The performance of the proposed algorithm in this environment is presented in Fig.6 by averaging 50 realizations of signals and propagation channels, in Fig.7, 8 a typical realization is displayed for different $\delta_b$ when $\delta = 0.08$. We use the same stopping rule as previously with LSCMA. For comparison, the performance of (4) with $\delta = 0.08$ is plotted in these figures too.

These simulations show that, as soon as the training sequence length $N$ is greater then 22, our algorithm yields almost the same MSE as the standard regularized LS estimator with $N \gg KL$. The curves in Fig.7 show that the the choice of the intermediate $\delta_b$ (curve 3) gives better performance compared with the extreme cases (curve 1 and 2) for variable $N$ with given $N_b$.

## Conclusion

The semi-blind criterion for the adjustment of the spatial-temporal filter is presented. It reduces a dimension of the CM criterion by the length of the training sequence. This allows to use CM criterion when spatio-temporal filter coefficients have to be estimated for each burst of data with the short training sequence inside.

## Acknowledgement

The authors wish to thank Prof.M.G.Bellanger, from Laboratoire d'Electronique of CNAM Paris and Prof.Y.I.Abramovich from Cooperative Research Centre for Sensor Signal and Information Processing, Australia for stimulating discussions and helpful advices.

## References

[1] E.Lindskog, A.Ahlén, M.Sternad, "Combined spatial and temporal equalization using an adaptive antenna array and a decision feedback equalization scheme", in Proc. ICASSP, pp.1189-1192,1995.

[2] A.J.Paulraj, "Space-time processing for wireless communications", in Proc. ICASSP, pp.1-4, 1997.

[3] S. Mayrargue, "A blind spatio-temporal equalizer for a radio-mobile channel using the constant modulus

algorithm (CMA)", in Proc. ICASSP, pp. 3-317 - 3-320, 1994.

[4] A.Gorokhov, P.Loubaton, “Semi-blind second order identification of convolutive channels", in Proc. ICASSP, pp.3905-3908, 1997.

[5] E. De Carvalho, D.T.M.Slock, “Maximum-likelihood blind FIR multi-channel estimation with Gaussian prior for the symbols", in Proc. ICASSP, pp.3593-3596, 1997.

[6] B.G.Agee, “The least-squares CMA: a new technique for rapid correction of constant modulus signals ", in Proc. ICASSP, pp.19.2.1-19.2.4, 1986.

[7] Y.I.Abramovich, “Controlled method for adaptive optimization of filters using the criterion of maximum SNR", Radio Engineering and Electronic Physics, vol.26 (3), pp.87-95, 1981.

| $M=0, L=1$ $(KL=5)$ | $M=1, L=5$ $(KL=25)$ | $M=2, L=7$ $(KL=35)$ |
|---|---|---|
| 88.9% | 9.7% | 0.1% |

Table.1. Estimations of a frequency of hitting to a local CM minimum adequate to CM signals or their delays with random initialization of LSCMA when $N_b = 150$

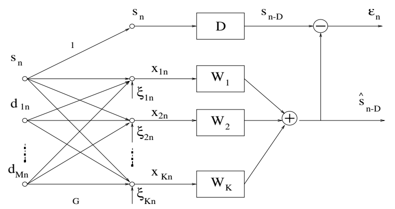


Figure 1 — Signal model and spatial-temporal filter

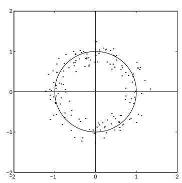

Figure 2 — Constellation picture for spurious CM local minimum when $N_b = 150$. There are 0.1% detected successful CM minimums with random initialization of LSCMA

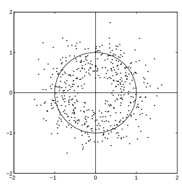

Figure 3 — Constellation picture for spurious CM local minimum when $N_b = 500$. There are 9.7% detected successful CM minimums with random initialization of LSCMA

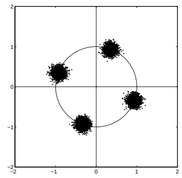

Figure 4 — Constellation picture for desired CM local minimum when $N_b = 10000$. There are 98.1% detected successful CM minimums with random initialization of LSCMA

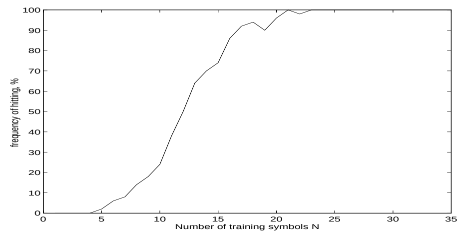


Figure 5 — Estimations of a frequency of hitting to the desired CM minimum for the proposed algorithm for $\delta_b = 1$ when $N_b = 150$, $M = 2$, $L_g = 5$, $K = 5$, $L = 7$

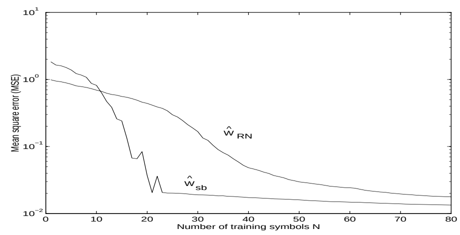


Figure 6 — Averaged performance of the proposed algorithm for $\delta_b = 1$ when $N_b = 150$, $M = 2$, $L_g = 5$, $K = 5$, $L = 7$

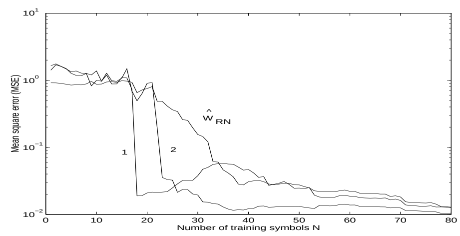


Figure 7 — Performance of the proposed algorithm for a typical realization for $\delta_b = 0.01$ (1 - constrained CM criterion (10)), $\delta_b = 1000$ (2 - LSCMA with initialization (13)), when $N_b = 150$, $M = 2$, $L_g = 5$, $K = 5$, $L = 7$

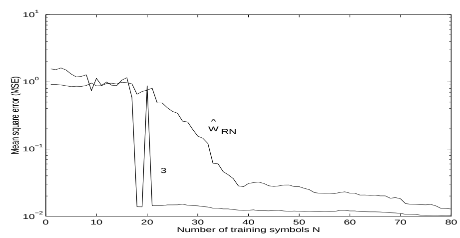


Figure 8 — Performance of the proposed algorithm at the experiment in Fig.7 when $\delta_b = 1$ (3)