



**HAL**  
open science

# Subspace SNR Maximization: The Constrained Stochastic Matched Filter

Bruno Borloz, Bernard Xerri

► **To cite this version:**

Bruno Borloz, Bernard Xerri. Subspace SNR Maximization: The Constrained Stochastic Matched Filter. *IEEE Transactions on Signal Processing*, 2011, 59 (4), pp.1346 - 1355. 10.1109/TSP.2010.2102755 . hal-01823637

**HAL Id: hal-01823637**

**<https://amu.hal.science/hal-01823637v1>**

Submitted on 26 Jun 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Subspace SNR maximization: the constrained stochastic matched filter

Bruno Borloz<sup>1,2</sup> - Bernard Xerri<sup>1,2</sup>

<sup>1</sup> Université du Sud Toulon Var, IM2NP, Equipe "Signaux et Systèmes"

<sup>2</sup> CNRS, IM2NP (UMR 6242)

Bâtiment X, BP 132, F-83957 La Garde Cedex (FRANCE)

Tel: 33 494 142 461 / 33 494 142 565

Fax: 33 494 142 598

borloz@univ-tln.fr, xerri@univ-tln.fr

## Abstract

In this paper, we propose a novel approach to perform detection of stochastic signals embedded in an additive random noise. Both signal and noise are considered to be realizations of zero mean random processes whose only second-order statistics are known (their covariance matrices).

The method proposed, called *Constrained Stochastic Matched Filter* (CSMF), is an extension of the *Stochastic Matched Filter* itself derived from the *Matched Filter*. The CSMF is optimal in the sense that it maximizes the *Signal-to-Noise Ratio* in a subspace whose dimension is fixed *a priori*.

In this paper, after giving the reasons of our approach, we show that there is neither obvious nor analytic solution to the problem expressed. Then an algorithm, which is proved to converge, is proposed to obtain the optimal solution.

The evaluation of the performance is completed through estimation of *Receiver Operating Characteristic* curves. Experiments on real signals show the improvement brought by this method and thus its significance.

*Keywords*— **detection ; subspace method ; reduced-rank method ; signal-to-noise ratio maximization ; matched filter ; matched subspace.**

## I. INTRODUCTION

This paper deals with the problem of detecting a stochastic signal (like a transient signal for example) embedded in an additive random noise.

Throughout this paper, all the signals will be real and discrete (time samples, pixels of images, ...) and represented with vectors of  $\mathbb{R}^N$ .

The method proposed here consists in a linear filtering called (for reasons explained later) "*Constrained Stochastic Matched Filter*" (CSMF). This method gives, for an integer value  $p$  ( $1 \leq p < N$ ), among all the  $p$ -dimension subspaces, the one where the *Signal-to-Noise Ratio* (SNR) is maximum: the CSMF is optimal for this criterion. This is a reduced-rank method (a projection) under constraint (the constraint being the *a priori* knowledge of the dimension  $p$ ) [1].

The SNR is invariant in a  $p$ -dimension subspace: it does not depend on the basis chosen to describe the subspace. An important consequence of this invariance of the SNR

w.r.t. the basis is that the simplest basis, say an orthonormal one, can usefully be chosen: moreover in such a basis the mathematical expression of the SNR is simple to obtain and will simplify later calculations.

In this paper we show that there is neither immediate nor obvious way to find the optimal  $p$ -dimension subspace: then we propose an algorithm and its the convergence to the good solution is proved.

The performances of the method and the comparisons with other methods are performed through the *Receiver-Operating-Characteristic* (ROC) curves giving the *Probability of Detection*  $P_D$  w.r.t. the *Probability of False Alarm*  $P_{FA}$ . Nevertheless, this paper gives no demonstration that ROC curves are better for a predicted value of  $p$ : we only observe, with results obtained from numerical simulations, that there exists a value of  $p$  for which the ROC curve is the best one.

Let us also note that our model is not a parametric one. The only knowledge is the covariance matrices of the random signals .

### A. Problem Formulation

Let us consider an observation  $\mathbf{x} \in \mathbb{R}^N$ . Two hypotheses can be formally stated (detection problem): this measurement was produced by ambient noise  $\mathbf{n}$  alone or by a signal  $\mathbf{s}$  embedded in this noise, respectively:

$$H_0: \mathbf{x} = \mathbf{n}$$

$$H_1: \mathbf{x} = \mathbf{s} + \mathbf{n}$$

The objective is to decide between these hypotheses. Our model will not be a parametric one.

The assumptions of our model are the following :

1)  $\mathbf{s}$  and  $\mathbf{n}$  are realizations of zero mean random processes.

2) The covariance matrices of  $\mathbf{s}$  and  $\mathbf{n}$ , respectively  $\mathbf{A}$  and  $\mathbf{B}$ , are supposed to be known, full rank and different.

3)  $\mathbf{s}$  and  $\mathbf{n}$  are uncorrelated, not necessarily Gaussian, and their *Probability Density Functions* (PDF) are unknown.

Two kinds of error are possible: the missing of the signal and the false alarm. A trade-off (highlighted by the ROC curves) must be found between a small average number of misses and a small average number of false alarms.

When the PDF of the signals are known, the key quantity to compute is the *Likelihood Ratio* (LR)  $\mathcal{L}(\mathbf{x})$  which must be compared to a threshold determined according to a criterion such as the minimization of the probability of error, the maximization of  $P_D$  when  $P_{FA}$  is fixed *a priori* (Neyman-Pearson criterion) [2]-[5].

When the PDF are unknown,  $\mathcal{L}(\mathbf{x})$  cannot be calculated. This is why we take into consideration methods based on SNR maximization. Furthermore the CSMF method takes place among numerous currently known reduced-rank techniques which have been proposed (Section II is a survey of some SNR maximization and reduced-rank methods justifying the approach of our method).

### B. Why the CSMF ?

When the PDFs are known (e.g. Gaussian), the *Likelihood Ratio Test* (LRT) is an optimal test which leads to compare a value to a threshold. For Gaussian signals, the test can easily be written as a sum of  $N$  terms (then onto the whole space of the used signals) depending on the observation vector  $\mathbf{x}$ :

$$\log \mathcal{L}(\mathbf{x}) = \Lambda(\mathbf{x}) = \sum_{i=1}^N \frac{\lambda_i}{1 + \lambda_i} (\mathbf{u}_i^t \mathbf{x})^2. \quad (1)$$

where the  $\lambda_i$  ( $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$ ) and  $\mathbf{u}_i$  are respectively the eigenvalues and corresponding eigenvectors of  $\mathbf{B}^{-1}\mathbf{A}$  (see Section II-C). In fact, the eigen-elements of  $\mathbf{B}^{-1}\mathbf{A}$  naturally appear when trying to maximize the output SNR of a linear filter  $\mathbf{h}$ : this output SNR can be written  $\rho = \frac{\mathbf{h}^t \mathbf{A} \mathbf{h}}{\mathbf{h}^t \mathbf{B} \mathbf{h}}$ .

The maximal value of  $\rho$ , noted  $\rho_{max}$ , is obtained for  $\mathbf{h} = \mathbf{u}_1$  the eigenvector associated with the largest eigenvalue  $\lambda_1$  of  $\mathbf{B}^{-1}\mathbf{A}$ : this filtering consists in a projection of the signal  $\mathbf{x}$  onto  $\mathcal{E}_{\mathbf{u}_1}$ , and then it is easy to verify that  $\rho_{max} = \lambda_1$ . This method is called "*Stochastic Matched Filter*" (SMF) [6].

When signals are not Gaussian, we can continue to use  $\Lambda(\mathbf{x})$  which is no longer the log of the LR. This expression has no reason to be optimal, and experimental results show, first, that a truncation of this sum to  $p$  terms can improve the ROC curves and next that there exists an optimal value of  $p$  for which the ROC curve is the best one. This truncation ( $p < N$ ) is expressed as follows

$$\sum_{i=1}^p \frac{\lambda_i}{1 + \lambda_i} (\mathbf{u}_i^t \mathbf{x})^2,$$

and can be seen as a projection of  $\mathbf{x}$  onto a  $p$ -dimension subspace  $\mathcal{E}_p^\dagger$ :  $\mathcal{E}_p^\dagger$  is spanned by  $\{\mathbf{u}_i\}_{i=1,\dots,p}$ . This method, called "*Extended Stochastic Matched Filter*" (ESMF) could be wrongly interpreted as a SNR maximization method: in fact it does not maximize the SNR in a  $p$ -dimension subspace but a weighted sum of output SNRs, each of them after a projection onto  $\mathcal{E}_{\mathbf{u}_i}$  for  $i = 1, \dots, p$  (see Section II-C).

The method proposed in this paper is naturally inferred from these remarks concerning the output SNR maximization and the projection onto a subspace of dimension two

or higher; therefore its aim is to maximize the SNR in an aptly chosen subspace with an *a priori* given dimension  $p$ . The choice of  $p$ , and then the dimension of the optimal subspace searched for, is a constraint: this is why the name of "*Constrained Stochastic Matched Filter*" (CSMF) was given to this optimal filter. We will clearly see that the CSMF is not a simple extension of the ESMF and that the CSMF can no more be inferred from the ESMF.

However, when  $p = 1$ , the CSMF and the SMF are identical. But when  $p > 1$ , it is proved in this paper (*cf.* Section III-F) that the optimal space  $\mathcal{E}_p^*$  cannot be simply deduced from the knowledge of either  $\mathcal{E}_p^\dagger$  or  $\mathcal{E}_{p-1}^*$ . Hence, it is necessary to propose an algorithm that finds the solution: this algorithm is given and is proved to converge to the solution (*cf.* Section IV).

### Organization of the paper

In Section I we formulate the mathematical model, present the basic assumptions. Section II describes existing methods and introduces those proposed in the paper. The method is detailed in Section III and useful properties are highlighted. Section IV is dedicated to the practical determination of this subspace: an algorithm is proposed to find the optimal subspace  $\mathcal{E}_p^*$  and the proof of its convergence is given. Then experimental results are presented in Section V.

In this paper, we apply the method to detection, but it could be obviously used for compression, filtering or estimation problems.

## II. OVERVIEW OF SOME EXISTING METHODS

The model is those given in Section I-A.

### A. The Karhunen-Loève Transform

The *Karhunen-Loève transform* (KLT) is a *Principal Component Analysis* used to tackle this model [6]-[8] when noise is white ( $\mathbf{B} = \sigma_n^2 \mathbf{I}$ ) or absent; it provides the best approximation, in the sense that it minimizes a *mean square error* (MSE) for a stochastic signal under the condition that its rank is fixed and is used for example for data compression or filtering. When noise is white, it determines the  $p$ -dimension subspace where the SNR is maximum.

But it does not consider colored noise, and therefore is not optimum even when it is used with a noise suppression filter such as the *Wiener filter* (which is not a reduced-rank method). The SMF, a *Generalized Eigen Decomposition* (GED) introduced by Cavassilas-Xerri [6] will be detailed in Section II-C: it performs a two-stage operation (pre-whitening and KLT), but is shown to be not optimal in terms of maximization of the SNR. GED is a major problem in many modern information processing applications (adaptive filtering, blind source separation [7], ...) and fast algorithms to estimate and track the principal generalized eigenvectors have been developed [8].

We will show that the CSMF can be seen like an extension of the KLT and the SMF for the problems we are interested in.

Some authors want to find an optimal linear data compression method in the presence of noise: the *Proportional KLT* (PKLT) applies an oblique projection operator onto a subspace  $\mathcal{S}$  ( $\dim(\mathcal{S}) = p$ ) along a subspace  $\mathcal{L}$  ( $\mathcal{S}$  and  $\mathcal{L}$  are both unknown) [9]. This operator  $\mathcal{P}$  naturally maximizes a ratio of powers. Solving this problem without any constraint concerning the rank of  $\mathcal{P}$  naturally leads to an impossibility.

The first part of their work shows how justifiable it is to take an interest in the maximization of the SNR in a subspace.

The CSMF proposed in this paper solves this problem by adding a constraint, the rank of the subspace to project data onto.

### B. Parametric Models

For detection with reduced-rank methods, many authors have worked on a parametric model of the following form:  $\mathbf{x} = \mathbf{H}\boldsymbol{\theta} + \mathbf{n} = \sum_{i=1}^p \theta_i \mathbf{h}_i + \mathbf{n}$  ( $\mathbf{H}$  is a  $N \times p$  matrix). The useful signal  $\mathbf{s} = \mathbf{H}\boldsymbol{\theta}$  is a stochastic signal constrained to lie in the signal subspace, the  $p$ -dimension subspace spanned by the known modes  $\mathbf{h}_i$ , with mode weights or gains, the entries  $\theta_i$  of  $\boldsymbol{\theta}$ . This model in an extension of those used for the *Matched Filter* (MF) detector, matched to a signal that is assumed to lie in a 1-dimensional subspace (*i.e.*  $\mathbf{H} = \mathbf{h}_1$  is the deterministic signal to detect).

It is noteworthy that our model is very different, because the covariance matrix  $\mathbf{A}$  is full rank and then the signal subspace has dimension  $N$ .

When the noise is Gaussian, the output of the MF provides a sufficient statistic for any LRT for detection. The knowledge of  $\mathbf{s}$  and the second-order statistics of  $\mathbf{n}$  is necessary to derive the corresponding MF. For  $p > 1$  the MF detector is no more efficient and is extended to the *Matched Subspace Detector* (MSD) [10]-[17] assuming prior knowledge of  $\mathbf{B}$ . The MF is also named *coherent MSD*.

When the gains are unknown, the *Generalized Likelihood Ratio Test* [5] takes the form of a ratio of two quadratic forms of prewhitened observations using orthogonal projections onto suitable subspaces. The statistic obtained has natural invariances (the energy of the subspace signal and the SNR are unchanged).

When  $\mathbf{B}$  is unknown, it is obtained from training data (*Adaptive Subspace Detectors*) [13][14].

Numerous papers deal with the MF detector and its extensions: several reasons may imply that signal or/and noise are not exactly known (channel nonlinearities, timing jitter, non-stationarities, model uncertainties ...) [18][19].

Another problem studied (e.g. in digital communications [20]) is those of detecting a transmitted signal when one out of several known signals is transmitted. When the additive noise is white and Gaussian, the optimal detector consists of a bank of MF followed by a detector which chooses as the detected signal the one with the maximal output value. Improvements have been observed in many cases [21].

### C. The Stochastic Matched Filter and the Extended SMF

#### C.1 Introduction

The SMF was first introduced to detect a random signal not supposed to lie in a known subspace: furthermore, the second-order statistics of both  $\mathbf{s}$  and  $\mathbf{n}$  are supposed to be known [6]. It can be seen like an extension of the MF (it provides an optimal filter since it maximizes the SNR), but also of the KLT. This problem is a generalized eigenvalue problem using the covariance matrices  $\mathbf{A}$  and  $\mathbf{B}$ ; the filtering is a projection onto the optimal subspace spanned by the eigenvector of  $\mathbf{B}^{-1}\mathbf{A}$  with maximum eigenvalue which is also the value of the maximum output SNR. The output SNR  $\rho$  of a linear filter  $\mathbf{h}$  can be written like a Rayleigh quotient:  $\rho = \frac{\mathbf{h}^t \mathbf{A} \mathbf{h}}{\mathbf{h}^t \mathbf{B} \mathbf{h}}$  (if  $\mathbf{A}$  and  $\mathbf{B}$  have unit trace,  $\rho$  can be interpreted like a gain on the SNR). This problem is equivalent to solve the following generalized eigenvalue problem:  $\mathbf{A} \mathbf{h} = \lambda \mathbf{B} \mathbf{h}$ . The maximal value  $\rho_{max}$  is obtained for  $\mathbf{h} = \mathbf{u}_1$  the eigenvector associated with the largest eigenvalue  $\lambda_1$  of  $\mathbf{B}^{-1}\mathbf{A}$ ; then  $\rho_{max} = \lambda_1 > 1$ : this filtering performs a projection of the signal onto  $\mathcal{E}_1^* = \mathcal{E}_{\mathbf{u}_1}$ .

If  $\{\lambda_i\}$  and  $\{\mathbf{u}_i\}$  are the eigenvalues and eigenvectors of  $\mathbf{B}^{-1}\mathbf{A}$  with  $\lambda_1 \geq \dots \geq \lambda_N$ , we easily prove that:

►  $\lambda_i \geq 0$  can be interpreted like a gain on SNR after projection onto  $\mathcal{E}_{\mathbf{u}_i}$ .

►  $\{\mathbf{u}_i\}$  is a non-orthogonal basis of  $\mathbb{R}^N$  performing simultaneous diagonalization of  $\mathbf{A}$  and  $\mathbf{B}$ . If  $\mathbf{U} = [\mathbf{u}_1 \dots \mathbf{u}_N]$ , an appropriate normalization of the  $\mathbf{u}_i$  gives, with  $\Delta$  positive diagonal matrix [22]:

$$\begin{cases} \mathbf{U}^t \mathbf{A} \mathbf{U} = \Delta \\ \mathbf{U}^t \mathbf{B} \mathbf{U} = \mathbf{I} \end{cases}$$

The interpretation of the  $\lambda_i$  naturally leads to take into account directions of projection that could statistically contribute to a better detection, that means growing the dimension of the subspace to project data onto. Actually, it has been shown [6] that a few other eigenvectors, the dominant one, can statistically contribute to improve ROC curves. To take a decision, we have to propose a function of  $\mathbf{x}$  and the  $\mathbf{u}_i$ .

When the signals are Gaussian, the calculation of the logarithm of the LR leads easily to equation (1). This expression has no reason to be optimal when the signals are not Gaussian, but, according to the remarks above, this summation is shortened to  $p$  terms corresponding to significant eigenvalues and the function chosen is then:

$$\Lambda(\mathbf{x}; p) = \sum_{i=1}^p \frac{\lambda_i}{1 + \lambda_i} (\mathbf{u}_i^t \mathbf{x})^2, \quad (2)$$

which is a weighted sum of  $(\mathbf{u}_i^t \mathbf{x})^2$  the power of the observation  $\mathbf{x}$  after projection onto each direction  $\mathbf{u}_i$ , each weight being linked to the SNR on this direction. This method, called "*Extended SMF*" (ESMF), does not maximize the SNR in a  $p$ -dimension subspace, but a weighted sum of output SNRs, each of them after a projection onto  $\mathcal{E}_{\mathbf{u}_i}$  for  $i = 1, \dots, p$ . We will denote  $\mathcal{E}_p^{\dagger}$  the subspace spanned by  $\{\mathbf{u}_i\}_{i=1, \dots, p}$ .

## C.2 Illustration: a practical example

To illustrate the interest of taking into account a subspace of dimension higher than one, let us apply this method on a narrow-band signal embedded in simulated underwater acoustics: the central frequency is  $f_0 = 3131$  Hz and the spectral bandwidth of noise is  $B = 1260$  Hz. The sampling frequency is here  $F_e = 15423$  Hz. Experiments are performed on  $N_r = 1000$  realizations of signal denoted  $\mathbf{s}_i$  ( $N = 21$ , see appendix A) and the initial SNR is about -14 dB. Envelop detection techniques could be used but, in practice, they give lower quality results.

$\mathbf{A}$  is calculated as follows:  $\mathbf{A} = \sum_{i=1}^{N_r} \mathbf{s}_i \mathbf{s}_i^t$ . For these simulations, the areas of presence or not of a signal are obviously known. A detection (which can be a false alarm) is decided each time there are at least 5 consecutive points of the result function done by equation (2) above the threshold.

ROC curves are shown in figure 1, first for  $p = 1$  ( $\mathcal{E}_1^\dagger$ ) and for the optimal value of  $p$ , say 3 ( $\mathcal{E}_3^\dagger$ ):

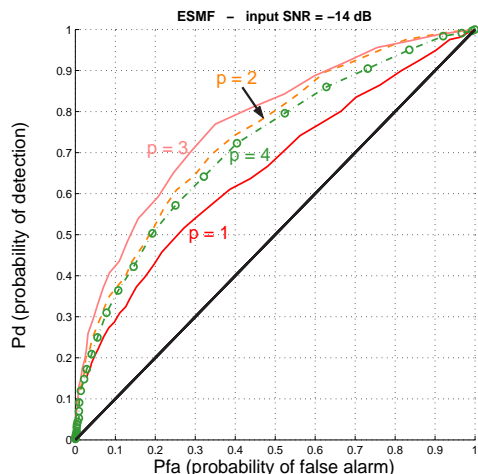


Fig. 1. ROC curves (input SNR=-14 dB) for different values of  $p$

The improvement brought by projecting data onto  $\mathcal{E}_3^\dagger$  is obvious. This example shows how the make of decision can be greatly improved by taking into consideration more than one eigenvector. Even if the SNR in  $\mathcal{E}_3^\dagger$  is smaller than in  $\mathcal{E}_1^\dagger$ , we observe that increasing the projection subspace dimension  $p$  brings an improvement that is not counterbalanced by the decrease of the SNR.

We can also see that there is a worsening of the ROC curve for  $p = 4$ . The projection onto  $\mathcal{E}_4^\dagger$  will give statistically worse results than those onto  $\mathcal{E}_3^\dagger$ . This result confirms that there is an optimal value of  $p$  for the chosen criterion.

## C.3 Conclusion

The SMF can be proved to be a two-stage operation: data whitening and then maximisation of SNR in a  $p$ -dimensional subspace  $\mathcal{E}_p^\dagger$  (KLT). But, as whitening is not an optimal operation in terms of SNR, the global operation has no reason to maximize the SNR in  $\mathcal{E}_p^\dagger$ . Hence, methods

which try to maximize a SNR while performing a reduced-rank operation are natural when PDF are unknown.

Thus, it seems natural and legitimate to ask oneself if there exists a  $p$ -dimensional subspace ( $p > 1$ ) in which the SNR is maximal, with the hope that ROC curves would be improved again, and then if it is possible to find it.

## III. THE CONSTRAINED STOCHASTIC MATCHED FILTER

The method proposed in this paper has been called *Constrained SMF* (CSMF) because it can be seen like an extension of the SMF, naturally inferred from the remarks in previous section concerning a projection onto a subspace of dimension two or higher where the SNR is maximum.

## A. Preliminary remarks and notations

A random signal  $\mathbf{s}$ , such as in our model, is a vector of  $\mathbb{R}^N$  and can always be expressed as follows:

$$\mathbf{s} = \sum_{i=1}^N \alpha_i \mathbf{x}_i = \mathbf{X} \mathbf{a}$$

where  $\mathbf{a} = [\alpha_1 \dots \alpha_N]^t$  is a vector of random variables and  $\mathbf{X} = [\mathbf{x}_1 \dots \mathbf{x}_N]$  a basis of  $\mathbb{R}^N$ .

In this paper we are interested only in powers in subspaces. Let us denote  $\mathbf{A} = \mathbf{E}(\mathbf{s} \mathbf{s}^t)$  the covariance matrix and  $P_s$  the power of  $\mathbf{s}$ . Of course,  $P_s$  depends only on the subspace and not on the basis used to describe it, and with no loss of generality, it is possible to consider only orthonormal bases to describe any subspace: hence  $\mathbf{X}^t \mathbf{X} = \mathbf{I}$ .

It readily follows that:

$$P_s = \sum_{i=1}^N \mathbf{x}_i^t \mathbf{A} \mathbf{x}_i = \text{tr}(\mathbf{X}^t \mathbf{A} \mathbf{X}) = \text{tr}(\mathbf{A} \mathbf{X} \mathbf{X}^t) = \text{tr} \mathbf{A}. \quad (3)$$

Moreover, as  $P_s = \text{tr} \mathbf{A}$ , we can consider, without loss of generality, only covariance matrices of trace 1.

 B. SNR in a  $p$ -dimension subspace  $\mathcal{E}_p$ 

Considering an integer  $p$  chosen *a priori* in  $[1, \dots, N-1]$ , let us denote  $\mathcal{E}_p = \mathcal{E}_{\{\mathbf{x}_1, \dots, \mathbf{x}_p\}}$  the  $p$ -dimension subspace spanned by the  $p$  orthonormal vectors  $\mathbf{x}_1$  to  $\mathbf{x}_p$ . We will also denote  $\mathcal{E}_p^\perp = \mathcal{E}_{\{\mathbf{x}_{p+1}, \dots, \mathbf{x}_N\}}$ . Then, with  $\mathbf{X}_p = [\mathbf{x}_1, \dots, \mathbf{x}_p]$ , the projection of  $\mathbf{s}$  onto  $\mathcal{E}_p$  along  $\mathcal{E}_p^\perp$  has power  $P_p$ :

$$P_p = \text{tr}(\mathbf{X}_p^t \mathbf{A} \mathbf{X}_p) = \sum_{i=1}^p \mathbf{x}_i^t \mathbf{A} \mathbf{x}_i,$$

and the expression of the SNR in  $\mathcal{E}_p$  takes the form:

$$\rho = \frac{\sum_{i=1}^p \mathbf{x}_i^t \mathbf{A} \mathbf{x}_i}{\sum_{i=1}^p \mathbf{x}_i^t \mathbf{B} \mathbf{x}_i} = \frac{\text{tr}(\mathbf{X}_p^t \mathbf{A} \mathbf{X}_p)}{\text{tr}(\mathbf{X}_p^t \mathbf{B} \mathbf{X}_p)}. \quad (4)$$

The objective is to find the unknowns  $\{\mathbf{x}_i\}$  in order to maximize this term. The optimal subspace will be denoted  $\mathcal{E}_p^*$  and the corresponding SNR  $\rho^*$ .  $\frac{\text{tr} \mathbf{A}}{\text{tr} \mathbf{B}}$  is the input SNR



and if covariance matrices  $\mathbf{A}$  and  $\mathbf{B}$  have unit trace,  $\rho$  is in fact a gain on the SNR (and no longer the output SNR) which can be proved to be necessarily lower than the largest eigenvalue of  $\mathbf{B}^{-1}\mathbf{A}$ , say  $\lambda_1$  [1].

We see here the important difference with the SMF which maximizes the following expression:

$$\rho_{smf} = \sum_{i=1}^p \frac{\mathbf{v}_i^t \mathbf{A} \mathbf{v}_i}{\mathbf{v}_i^t \mathbf{B} \mathbf{v}_i},$$

where the  $\mathbf{v}_i$  does not form an orthonormal basis.

Throughout the following section, we will focus our attention on equation (4) and try to find  $\mathcal{E}_p^*$ .

### C. Characterization of the optimal subspace $\mathcal{E}_p^*$

Let us consider a  $p$ -dimension subspace  $\mathcal{E}_p$  spanned by a set of  $p$  orthonormal vectors  $\mathbf{X} = [\mathbf{x}_1 \dots \mathbf{x}_p]$ . The expression of the SNR  $\rho$  in  $\mathcal{E}_p$  is given by the equation (4).

The constraints can be expressed by the  $p^2$  relationships “ $\mathbf{x}_i^t \mathbf{x}_j = \delta_{ij}$ ”. Clearly,  $p$  is given and the unknowns of our problem are  $\rho^*$  and the  $p$  vectors  $\mathbf{x}_i$  which must be calculated so as to maximize  $\rho$ . We are faced with an optimization problem with constraints which is usually solved using a Lagrange multipliers method. Let us define the following function:

$$L(\mathbf{X}, \mathbf{\Omega}) = \rho + \sum_{i=1}^p \sum_{j=1}^p \omega_{ij} (\mathbf{x}_i^t \mathbf{x}_j - \delta_{ij}). \quad (5)$$

This equation can be written:

$$L(\mathbf{X}, \mathbf{\Omega}) = \frac{\text{tr}(\mathbf{X}^t \mathbf{A} \mathbf{X})}{\text{tr}(\mathbf{X}^t \mathbf{B} \mathbf{X})} + \text{tr}(\mathbf{\Omega} (\mathbf{X}^t \mathbf{X} - \mathbf{I})), \quad (6)$$

where  $\mathbf{\Omega} \equiv [\omega_{ij}]$  is a  $p \times p$  symmetric matrix.

A necessary condition for this value to be maximum is  $\frac{\partial L}{\partial \mathbf{X}} = 0$ , which means that for  $\rho = \rho^*$ :

$$\frac{(\mathbf{A} - \rho^* \mathbf{B}) \mathbf{X}}{\text{tr}(\mathbf{X}^t \mathbf{B} \mathbf{X})} + \mathbf{X} \mathbf{\Omega} = 0.$$

As  $\mathbf{B}$  is positive definite,  $\text{tr}(\mathbf{X}^t \mathbf{B} \mathbf{X}) > 0$  and this equation becomes

$$(\mathbf{A} - \rho^* \mathbf{B}) \mathbf{X} = \mathbf{X} \mathbf{\Omega}_0 \quad (7)$$

where  $\mathbf{\Omega}_0$  is a  $p \times p$  real symmetric matrix but not diagonal. But we can find a real orthogonal matrix  $\mathbf{\Pi}$  and a real diagonal matrix  $\Delta_\mu^* \equiv [\mu_i^*]$  such that  $\mathbf{\Omega}_0 = \mathbf{\Pi} \Delta_\mu^* \mathbf{\Pi}^t$ . Then equation (7) becomes:

$$(\mathbf{A} - \rho^* \mathbf{B}) \mathbf{X} \mathbf{\Pi} = \mathbf{X} \mathbf{\Pi} \Delta_\mu^*. \quad (8)$$

As  $\mathbf{\Pi}$  is invertible,  $\mathbf{X} \mathbf{\Pi}$  and  $\mathbf{X}$  span the same subspace  $\mathcal{E}_p^*$ . Furthermore, as  $\mathbf{\Pi}$  is a real orthogonal matrix, the set of orthonormal vectors  $\mathbf{X}$  is changed to another set of orthonormal vectors  $\mathbf{X} \mathbf{\Pi}$ . Noting  $\mathbf{X} \mathbf{\Pi} = \mathbf{T}^* = [\mathbf{t}_1^* \dots \mathbf{t}_p^*]$ , equation (8) can be written

$$(\mathbf{A} - \rho^* \mathbf{B}) \mathbf{T}^* = \mathbf{T}^* \Delta_\mu^* \quad (9)$$

which is an eigenvalue problem. Note that for any value of  $\rho$ ,  $(\mathbf{A} - \rho \mathbf{B})$  is always real symmetric and then diagonalizable through a  $N \times N$  unitary real eigenvector matrix. That means that  $\mathcal{E}_p^*$  is spanned by a set of  $p$  orthonormal vectors which are eigenvectors of  $(\mathbf{A} - \rho^* \mathbf{B})$ . Nevertheless, equation (9) is not simple to solve because, if the  $\mathbf{t}_i^*$  and  $\mu_i^*$  are naturally unknown,  $\rho^*$  is unknown too.

For any value  $\rho > 0$  let us denote:

$$(\mathbf{A} - \rho \mathbf{B}) \mathbf{t}_i = \mu_i \mathbf{t}_i, \quad i = 1, \dots, N. \quad (10)$$

$(\mathbf{A} - \rho \mathbf{B})$  is always real symmetric and the  $\{\mathbf{t}_i\}$  naturally form an orthonormal basis.

We note that the eigenvalues  $\mu_i$  depend on  $\rho$ : it is easy to show, with trivial examples, that they are non nonlinear w.r.t.  $\rho$ . A simple illustration is given on figure 2 where we can see the evolution of the eigenvalues  $\mu_i$  w.r.t.  $\rho$  for the covariance matrices  $\mathbf{A}$  and  $\mathbf{B}$  given in example 1 (cf. Section III-F).

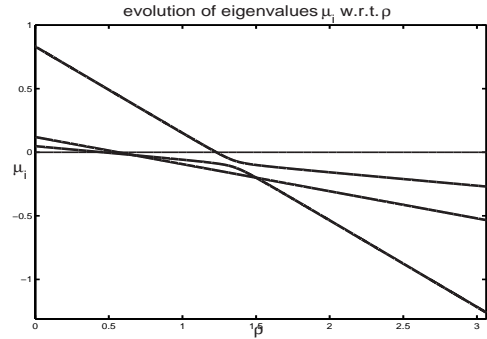


Fig. 2.  $\mu_i(\rho)$  for example 1 ( $N = 3$ )

### D. Property of the optimal subspace $\mathcal{E}_p^*$

Equation (10) implies that for any  $i, j \in \{1, \dots, N\}$ ,  $\mathbf{t}_j^t (\mathbf{A} - \rho \mathbf{B}) \mathbf{t}_i = \mu_i \mathbf{t}_j^t \mathbf{t}_i = \mu_i \delta_{ij}$ . Then for any subset  $I$  of  $\{1, 2, \dots, N\}$  with cardinality  $p$ ,

$$\sum_{i \in I} \mathbf{t}_i^t (\mathbf{A} - \rho \mathbf{B}) \mathbf{t}_i = \sum_{i \in I} \mathbf{t}_i^t \mathbf{A} \mathbf{t}_i - \rho \sum_{i \in I} \mathbf{t}_i^t \mathbf{B} \mathbf{t}_i = \sum_{i \in I} \mu_i,$$

so that

$$\frac{\sum_{i \in I} \mathbf{t}_i^t \mathbf{A} \mathbf{t}_i}{\sum_{i \in I} \mathbf{t}_i^t \mathbf{B} \mathbf{t}_i} - \rho = \frac{\sum_{i \in I} \mu_i}{\sum_{i \in I} \mathbf{t}_i^t \mathbf{B} \mathbf{t}_i}. \quad (11)$$

For the optimal subspace  $\mathcal{E}_p^*$ , i.e.  $\rho = \rho^*$ , the left expression is null, implying that

$$\sum_{i \in I^*} \mu_i^* = 0. \quad (12)$$

This property will be used to find the solution in the two trivial cases ‘ $p = 1$ ’ and ‘ $p = N - 1$ ’ (cf. Section III-E), but also to prove the convergence of the algorithm (cf. Section IV-B).

We have denoted  $I = I^*$ ; it is easy to show that  $I^* = \{1, 2, \dots, p\}$  if the eigenvalues are sorted in decreasing order, say  $\mu_1^* > \mu_2^* > \dots > \mu_N^*$ .

### E. Particular cases

In some particular cases, it is possible to reach the solution easily, without any sophisticated algorithm.

#### ► $p = 1$

The eigenvalue  $\mu_1^*$  to take into account is null. Hence,  $(\mathbf{A} - \rho^*\mathbf{B})\mathbf{t}_1^* = \mu_1^*\mathbf{t}_1^* = 0$ , *i.e.*  $\mathbf{A}\mathbf{t}_1^* = \rho^*\mathbf{B}\mathbf{t}_1^*$ .  $\rho^*$  is the largest eigenvalue of  $\mathbf{B}^{-1}\mathbf{A}$  and  $\mathbf{t}_1^*$  its associated eigenvector. Naturally we find in this case the SMF.

#### ► $p = N - 1$

As  $\sum_{i=1}^{N-1} \mu_i^* = 0$ ,  $\mu_N = 1 - \rho^*$ . Hence,  $(\mathbf{A} - \rho^*\mathbf{B})\mathbf{t}_N = (1 - \rho^*)\mathbf{t}_N$ , or  $(\mathbf{A} - \mathbf{I}_N)\mathbf{t}_N = \rho^*(\mathbf{B} - \mathbf{I}_N)\mathbf{t}_N$ ;  $\rho^*$  is the largest eigenvalue of  $(\mathbf{B} - \mathbf{I}_N)^{-1}(\mathbf{A} - \mathbf{I}_N)$  and  $\mathbf{t}_N$  its associated eigenvector. The  $N - 1$  vectors spanning  $\mathcal{E}_{N-1}^*$  are the other eigenvectors.

#### ► $\mathbf{B} = \mathbf{I}$

In this case,  $(\mathbf{A} - \rho\mathbf{I})\mathbf{t}_i = \mu_i\mathbf{t}_i$ , *i.e.*  $\mathbf{A}\mathbf{t}_i = (\rho + \mu_i)\mathbf{t}_i$ . For  $\rho = \rho^*$ ,  $\sum_{i=1}^p \mu_i = 0$  and then  $\sum_{i=1}^p (\rho + \mu_i) = p\rho = \sum_{i=1}^p \lambda_i^A$  (where  $\lambda_i^A$  is the  $i$ -th eigenvalue of  $\mathbf{A}$ ), that means  $\rho = \frac{1}{p} \sum_{i=1}^p \lambda_i^A$ . Hence the Karhunen-Loeve filter is a particular case of the CSMF.

### F. Examples

For the following examples, we state  $p = 2$  and search the optimal subspace  $\mathcal{E}_2^*$ . As  $p = N - 1$ , we can use the results of the previous paragraph: the SNR  $\rho^*$  obtained in  $\mathcal{E}_2^*$  is easy to calculate.

#### ► Example 1

$$\mathbf{A} = \frac{1}{3} \begin{bmatrix} 1.00 & 0.80 & 0.64 \\ 0.80 & 1.00 & 0.80 \\ 0.64 & 0.80 & 1.00 \end{bmatrix},$$

$$\mathbf{B} = \frac{1}{3} \begin{bmatrix} 1.00 & 0.60 & 0.36 \\ 0.60 & 1.00 & 0.60 \\ 0.36 & 0.60 & 1.00 \end{bmatrix}.$$

We denote  $\mathbf{U}$  and  $\Delta_\lambda$  the matrices such that  $\mathbf{A}\mathbf{U} = \mathbf{B}\mathbf{U}\Delta_\lambda$ :

$$\mathbf{U} = \begin{bmatrix} 0.6611 & -0.7071 & 0.4170 \\ 0.3549 & 0.0000 & -0.8076 \\ 0.6611 & 0.7071 & 0.4170 \end{bmatrix},$$

and  $\lambda_1 = 1.23$ ,  $\lambda_2 = 0.56$  and  $\lambda_3 = 0.46$ . The SNR obtained in  $\mathcal{E}_2^\dagger$  spanned by the eigenvectors associated with the two largest eigenvalues of  $\mathbf{B}^{-1}\mathbf{A}$  ( $\mathbf{u}_1$  and  $\mathbf{u}_2$ ) is  $\rho^\dagger = 1.06$ .

The SNR obtained in  $\mathcal{E}_2^*$  is  $\rho^* = 1.118$ . If we denote  $(\mathbf{A} - \rho^*\mathbf{B})\mathbf{T}^* = \mathbf{T}^*\Delta_\mu^*$ , then:

$$\mathbf{T}^* = \begin{bmatrix} 0.3370 & -0.7071 & 0.6216 \\ -0.8791 & 0.0000 & 0.4766 \\ 0.3370 & 0.7071 & 0.6216 \end{bmatrix},$$

and  $\mu_1^* = -0.0725$ ,  $\mu_2^* = -0.1186$  and  $\mu_3^* = 0.0725$ .  $\mathcal{E}_2^*$  is spanned by  $\mathbf{t}_1^*$  and  $\mathbf{t}_3^*$  (we verify that  $\mu_1^* + \mu_3^* = 0$  and

$\mu_2^* = 1 - \rho^*$ ). In this example, since  $\mathbf{t}_2^* = \mathbf{u}_2$ ,  $\mathcal{E}_2^*$  is spanned by two eigenvectors of  $\mathbf{B}^{-1}\mathbf{A}$ , namely  $\mathbf{u}_1$  and  $\mathbf{u}_3$  which are clearly not those associated with the two largest  $\lambda_i$ .

#### ► Example 2

Let us consider two covariance matrices of non stationary processes.

$$\mathbf{A} = \frac{1}{3} \begin{bmatrix} 0.0379 & 0.0379 & 0.1514 \\ 0.0379 & 0.0473 & 0.2650 \\ 0.1514 & 0.2650 & 2.9148 \end{bmatrix},$$

and

$$\mathbf{B} = \frac{1}{3} \begin{bmatrix} 1.2872 & 1.4658 & 0.1313 \\ 1.4658 & 1.6865 & 0.1629 \\ 0.1313 & 0.1629 & 0.0263 \end{bmatrix}.$$

Denote  $\mathbf{U}$  and  $\Delta_\lambda$  the matrices satisfying  $\mathbf{A}\mathbf{U} = \mathbf{B}\mathbf{U}\Delta_\lambda$ .

$$\mathbf{U} = \begin{bmatrix} 0.5219 & 0.7555 & -0.3972 \\ -0.5189 & -0.6549 & 0.9156 \\ 0.6770 & 0.0192 & -0.0622 \end{bmatrix},$$

and  $\lambda_1 = 1051$ ,  $\lambda_2 = 0.45$  and  $\lambda_3 = 0.01$ . The SNR obtained in  $\mathcal{E}_2^\dagger$  spanned by the eigenvectors associated with the two largest eigenvalues of  $\mathbf{B}^{-1}\mathbf{A}$  is  $\rho^\dagger = 149.64$ .

The SNR obtained in  $\mathcal{E}_2^*$  is  $\rho^* = 151.05$ . We denote  $(\mathbf{A} - \rho^*\mathbf{B})\mathbf{T}^* = \mathbf{T}^*\Delta_\mu^*$ , then:

$$\mathbf{T}^* = \begin{bmatrix} 0.7111 & -0.2531 & -0.6559 \\ -0.5845 & 0.3057 & -0.7516 \\ -0.3908 & -0.9179 & -0.0695 \end{bmatrix}.$$

and  $\mu_1^* = -0.54$ ,  $\mu_2^* = 0.54$  and  $\mu_3^* = -150.05$ .  $\mathcal{E}_2^*$  is spanned by  $\mathbf{t}_1^*$  and  $\mathbf{t}_2^*$  (we verify that  $\mu_1^* + \mu_2^* = 0$  and  $\mu_3^* = 1 - \rho^*$ ). No eigenvector of  $\mathbf{B}^{-1}\mathbf{A}$  is contained in  $\mathcal{E}_2^*$ , *a fortiori* the eigenvector associated with the largest eigenvalue of  $\mathbf{B}^{-1}\mathbf{A}$  which generates  $\mathcal{E}_1^*$ . This example proves that a recursive algorithm w.r.t.  $p$ , that would calculate  $\mathcal{E}_1^*$  and then  $\mathcal{E}_2^*$ , etc... is not realistic.

#### ► Conclusion

From these simple examples, we immediately see that the optimal subspace  $\mathcal{E}_p^*$  is not necessarily spanned by eigenvectors of  $\mathbf{B}^{-1}\mathbf{A}$ , and even when this is the case, the eigenvectors are not necessarily those associated with the largest eigenvalues. It is not possible to deduce  $\mathcal{E}_p^*$  from  $\mathcal{E}_p^\dagger$ . What's more, it is not possible to find a recursive formulation on  $p$  to find  $\mathcal{E}_p^*$  from  $\mathcal{E}_{p-1}^*$ : for example, the relationship  $\mathcal{E}_1^* \subset \mathcal{E}_2^*$  is not necessarily verified. Then we have to propose an algorithm to determine  $\mathcal{E}_p^*$ . This will be performed in Section IV.

### G. Conclusion

In this section, the problem has been presented and equations have been deduced that must be solved to find the optimal  $p$ -dimension subspace. We have seen that there exists neither analytic nor obvious solution and that an algorithm must be proposed. This is the purpose of the next section.

The CSMF consists in finding the  $p$ -dimension subspace which maximizes the SNR after only a projection.

IV. ALGORITHM TO FIND THE OPTIMAL SUBSPACE  $\mathcal{E}_p^*$ 

In light of the examples of previous section, it is required to find an algorithm to determine the optimal subspace  $\mathcal{E}_p^*$  spanned by  $p$  vectors  $\mathbf{t}_i$  verifying equation (10) and maximizing  $\rho$  defined by

$$\rho = \frac{\sum_{i \in I} \mathbf{t}_i^t \mathbf{A} \mathbf{t}_i}{\sum_{i \in I} \mathbf{t}_i^t \mathbf{B} \mathbf{t}_i}, \quad (13)$$

where  $I$  is a subset of  $p$  different numbers out of  $\{1, \dots, N\}$ .

It seems natural that such an algorithm should be iterative and use, at each step, these equations alternately.

## A. Presentation of the algorithm

$\rho^*$  being unknown, it is reasonable to choose for the initial value of  $\rho$ , say  $\rho^{(0)}$ , the largest eigenvalue of  $\mathbf{B}^{-1} \mathbf{A}$ . At each step  $n \geq 0$ , we obtain the symmetric matrix  $(\mathbf{A} - \rho^{(n)} \mathbf{B})$  and calculate its  $N$  eigenvectors  $\mathbf{t}_i^{(n)}$  associated to eigenvalues  $\mu_i^{(n)}$ . Then we must choose among them the  $p$  ones  $\{\mathbf{t}_i^{(n)}, i \in I^{(n)} / \text{card}(I^{(n)}) = p\}$  for which

$$\rho^{(n+1)} = \frac{\sum_{i \in I^{(n)}} \mathbf{t}_i^{(n)t} \mathbf{A} \mathbf{t}_i^{(n)}}{\sum_{i \in I^{(n)}} \mathbf{t}_i^{(n)t} \mathbf{B} \mathbf{t}_i^{(n)}} \quad (14)$$

is maximum. These  $p$  vectors span a subspace  $\mathcal{E}_p^{(n)}$ .

Then it is easy to calculate  $(\mathbf{A} - \rho^{(n+1)} \mathbf{B})$ ,  $I^{(n+1)}$  and the new subspace  $\mathcal{E}_p^{(n+1)}$ . This process can be iterated until  $|\rho^{(n+1)} - \rho^{(n)}| < \varepsilon$  (see table 1). Of course, we have to prove that this algorithm converges to the good solution  $\rho^*$ .

## B. Study of the convergence

At step  $n$ , from (11) and (14), the variation of  $\rho$  is

$$\rho^{(n+1)} - \rho^{(n)} = \frac{\sum_{i \in I^{(n)}} \mu_i^{(n)}}{\sum_{i \in I^{(n)}} \mathbf{t}_i^{(n)t} \mathbf{B} \mathbf{t}_i^{(n)}}. \quad (15)$$

Of course, if  $\sum_{i \in I^{(n)}} \mu_i^{(n)} = 0$ , then  $\rho^{(n+1)} = \rho^{(n)}$ . Hence, as it has been proved that for  $\rho^*$ , there exists a subset  $I^*$  of cardinal  $p$  such as (12) is verified, say  $\sum_{i \in I^*} \mu_i^* = 0$ , and  $\rho^*$  is clearly a fixed-point of the algorithm.

Let us denote:

$$\sum_{i \in I^{(n)}} \mu_i^{(n)} = f(\rho^{(n)}). \quad (16)$$

Then  $f(\rho^*) = 0$ .

As for any value of  $\rho$ ,  $\mathbf{A} - \rho \mathbf{B}$  is symmetrical, the  $\{\mathbf{t}_i\}$  span an orthonormal basis: then  $\mathbf{t}_i^t \mathbf{t}_i = 1, \forall i$ . This expression can be differentiated w.r.t.  $\rho$ , leading to:

$$\mathbf{t}_i^t \frac{\partial \mathbf{t}_i}{\partial \rho} = 0, \forall i. \quad (17)$$

## Description of the algorithm

$\rho^{(0)} = \lambda_1$  the largest eigenvalue of  $\mathbf{B}^{-1} \mathbf{A}$

$n = 0$

1) Calculate  $\mathbf{M}^{(n)} = \mathbf{A} - \rho^{(n)} \mathbf{B}$

2) Compute the eigenvalues  $\mu_i^{(n)}$  and corresponding eigenvectors  $\mathbf{t}_i^{(n)}$  ( $i = 1$  to  $N$ ) of  $\mathbf{M}^{(n)}$

(it is possible to sort the eigen-elements so that

$\mu_1^{(n)} \geq \mu_2^{(n)} \geq \dots \geq \mu_N^{(n)}$ )

3) Find the  $C_N^p$  combinations of  $p$  elements out of  $\{1, 2, \dots, N\}$ :

they will be denoted  $I_k^{(n)}$  with  $\text{card}(I_k^{(n)}) = p$

and  $1 \leq k \leq C_N^p$

4) For  $k = 1$  to  $C_N^p$

$$\text{Calculate } \rho_k^{(n+1)} = \frac{\sum_{i \in I_k^{(n)}} \mathbf{t}_i^{(n)t} \mathbf{A} \mathbf{t}_i^{(n)}}{\sum_{i \in I_k^{(n)}} \mathbf{t}_i^{(n)t} \mathbf{B} \mathbf{t}_i^{(n)}} \quad (k^{\text{th}} \text{ combination})$$

combination)

5) Find the maximal value of  $\{\rho_k^{(n+1)}\}_{k=1, \dots, C_N^p}$ : this

maximum is denoted  $\rho^{(n+1)}$

6) If  $|\rho^{(n+1)} - \rho^{(n)}| < \varepsilon$

stop iterations,

$\mathbf{t}_i^* = \mathbf{t}_i^{(n)}$ ,  $\mu_i^* = \mu_i^{(n)}$ ,  $\rho^* = \rho^{(n+1)}$

Else

$n \leftarrow n + 1$ ,

Goto 1)

End If

$\mathcal{E}_p^*$  is spanned by the  $\{\mathbf{t}_i^*\}_{i \in I^* = \{1, 2, \dots, p\}}$

Table 1 : description of the algorithm

The differentiation of equation (10) w.r.t.  $\rho$  leads to

$$-\mathbf{B} \mathbf{t}_i + (\mathbf{A} - \rho \mathbf{B}) \frac{\partial \mathbf{t}_i}{\partial \rho} = \frac{\partial \mu_i}{\partial \rho} \mathbf{t}_i + \mu_i \frac{\partial \mathbf{t}_i}{\partial \rho}.$$

From equation (17) and multiplying by  $\mathbf{t}_i^t$ , it comes (using equation  $\mathbf{t}_i^t (\mathbf{A} - \rho \mathbf{B}) = \mu_i \mathbf{t}_i^t$ ):

$$-\mathbf{t}_i^t \mathbf{B} \mathbf{t}_i + \mu_i \mathbf{t}_i^t \frac{\partial \mathbf{t}_i}{\partial \rho} = \frac{\partial \mu_i}{\partial \rho},$$

or

$$\forall i, \quad \frac{\partial \mu_i}{\partial \rho} = -\mathbf{t}_i^t \mathbf{B} \mathbf{t}_i < 0. \quad (18)$$

Using equations (16) and (18), equation (15) becomes:

$$\rho^{(n+1)} = \rho^{(n)} + \frac{\sum_{i \in I^{(n)}} \mu_i^{(n)}}{-\sum_{i \in I^{(n)}} \frac{\partial \mu_i^{(n)}}{\partial \rho}} = \rho^{(n)} - \frac{f(\rho^{(n)})}{f'(\rho^{(n)})}, \quad (19)$$

Obviously  $f'(\rho^*)$  is not null: in fact, from (18) we have  $f'(\rho) = \sum_{i \in I} \frac{\partial \mu_i}{\partial \rho} < 0$  for any value of  $\rho$ .

We can use the Newton theorem that says that if :

1)  $f(\rho)$  is twice differentiable,



- 2)  $f(\rho^*) = 0$ ,
- 3)  $\rho^{(0)}$  is "close to"  $\rho^*$ ,
- 4) " $f'(\rho^*) \neq 0$ ",

then the series defined by (19) converges to  $\rho^*$  with a quadratic speed. It is easy to prove that  $I^* = \{1, \dots, p\}$  (if the eigenvalues are sorted so that  $\mu_1^* \geq \mu_2^* \geq \dots \geq \mu_N^*$ ). But we must be careful because that is not true at any step  $n$  for  $I^{(n)}$ .

Then this algorithm converges to the solution  $\rho^*$  of our problem.

### C. Uniqueness of the solution

If we denote  $\lambda_1$  the largest eigenvalue of  $\mathbf{B}^{-1}\mathbf{A}$ , then  $1 \leq \rho^* \leq \lambda_1$ . In particular cases, it may be possible to find several subspaces of dimension  $p$  for which the SNR  $\rho$  is maximal; in fact, this is not a problem. In such a case, we can take an interest in finding a subspace of higher dimension than  $p$  with the same SNR, or we can add a new criterion to choose among those subspaces.

### D. Practical remark

At each step, the identification of  $I^{(n)}$  requires a heavy calculation.  $C_p^N$  different combinations have to be tested, which can quickly increase to an unacceptable number of calculations. The algorithm proposed can be improved significantly. Instead of searching the optimal set of eigenvectors in a systematic way, one can advantageously rearrange the eigenvalues of  $(\mathbf{A} - \rho^{(n)}\mathbf{B})$  in decreasing order:  $\mu_1^{(n)} > \mu_2^{(n)} > \dots > \mu_N^{(n)}$  (note that those values can as well be positive or negative), and take, at step  $n$ ,  $I^{(n)} = \{1, 2, \dots, p\}$  (we have seen that  $I^* = \{1, 2, \dots, p\}$ ): equation (14) becomes

$$\rho^{(n+1)} = \frac{\sum_{i=1}^p \mathbf{t}_i^{(n)t} \mathbf{A} \mathbf{t}_i^{(n)}}{\sum_{i=1}^p \mathbf{t}_i^{(n)t} \mathbf{B} \mathbf{t}_i^{(n)}}$$

. In theory, there is no reason for  $\rho^{(n+1)}$  to be maximal, but in practice it so happens that  $\rho^{(n+1)}$  reaches almost systematically the maximal value, and if not, reaches a value very close to it. In the neighbourhood of the solution, the convergence is assured. Such a change of the algorithm obviously highly decreases the sum of calculation. In terms of convergence to  $\rho^*$ , there is a slight drop in the speed of convergence in terms of number of iterations. Globally, however this method converges to the solution and decreases the sum of calculation in a very significant proportion. To give a precise idea of the gain, for  $N = 21$  and  $p = 5$ ,  $C_N^p = 20349$ .

The convergence of this modified algorithm has not been proved.

### E. Conclusion

In this section, the convergence of the algorithm proposed has been proved.

For a given value of  $p$ , we have initialized the algorithm with  $\rho^{(0)} = \lambda_1$  the largest eigenvalue of  $\mathbf{B}^{-1}\mathbf{A}$ , saying it

was reasonable to choose this value. What we are searching for is,  $p$  being fixed, the global maximum w.r.t.  $\rho$  (there exist other local maxima), and this maximum is necessarily the nearest from  $\rho = \lambda_1$  ( $1 \leq \rho^* \leq \lambda_1$ ).

Obviously, an initialization of the algorithm with any value  $\rho^{(0)} \neq \lambda_1$  can lead the algorithm to find a local maximum.

## V. EXPERIMENTAL RESULTS

Let us apply the CSMF method to the example described in Section III-F. Results are shown in Figure 3: the quality of the ROC curves increases from  $p = 1$  to  $p = 4$  (or  $p = 5$  which gives more or less the same results than  $p = 4$ ) and decreases from  $p = 6$ . The optimal result is obtained for  $p = 4$  or 5.

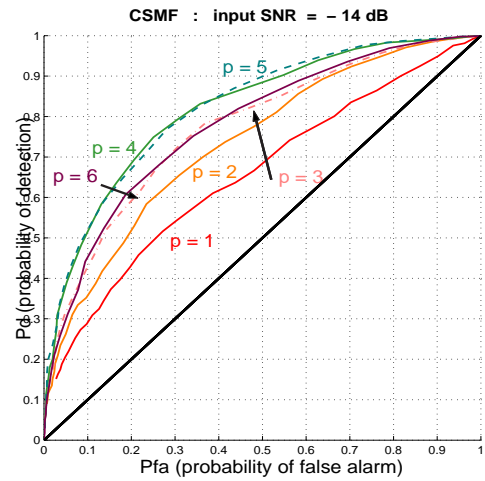


Fig. 3. CSMF : ROC curves for different values of  $p$

Figure 4 shows the best result obtained by the CSMF ( $\mathcal{E}_4^*$ ) and the best one obtained by the ESMF ( $\mathcal{E}_3^\dagger$ ). The ROC curve obtained in  $\mathcal{E}_4^*$  (the best result reachable) is obviously above those obtained in  $\mathcal{E}_3^\dagger$ . Note that  $\mathcal{E}_1^* = \mathcal{E}_1^\dagger$ .

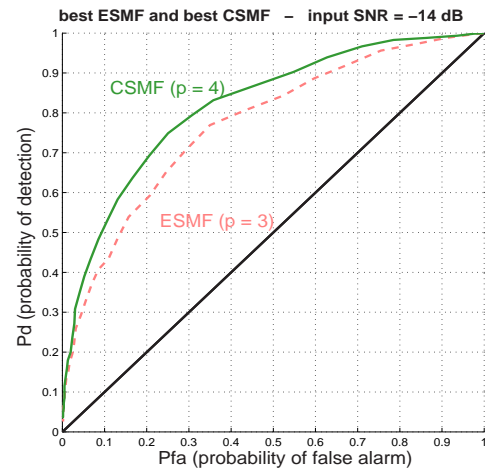


Fig. 4. Best ROC curves for CSMF ( $p = 4$ ) and ESMF ( $p = 3$ )

This example illustrates clearly the improvement that

can be brought by maximizing the SNR in  $\mathcal{E}_4^*$  instead of  $\mathcal{E}_1^*$ , but also its superiority in comparison with the ESMF method.

Now the optimal subspace  $\mathcal{E}_p^*$  (here  $p = 4$ ) has been found with the CSMF method which is a projection (a reduced-rank method). Nevertheless we did not use all the possibilities of classical filtering, and among all the basis of  $\mathcal{E}_p^*$ , we can choose one with interesting properties after a linear filtering. For example, the ESMF gives preferential treatment to directions (spanned by the  $\mathbf{u}_i$ ) with the best SNR ( $\lambda_i$ ) (like the Wiener filter): after this linear filtering, the power of noise is one in any direction  $\mathbf{u}_i$ .

Thus, we calculate the ROC curves obtained with the equation (2) where  $p$  is the subspace dimension so that all the basis vectors of  $\mathcal{E}_4^*$  are taken into account. Results are shown in figure 5.

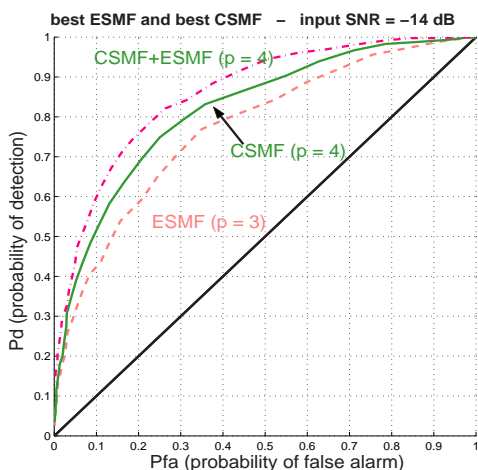


Fig. 5. ROC curves for SNR=-14 dB

A noteworthy improvement can be observed by using a simultaneous diagonalization technique in the optimal subspace calculated beforehand. We finally used a projection (to find the optimal subspace) and a linear filtering operation to improve again the detection.

## VI. CONCLUSION

The method proposed in this paper takes its place in the set of methods of decomposition of signals on appropriate basis but also in subspace methods.

When trying to detect stochastic signals with known covariance matrices but with no *a priori* knowledge on their probability density function, people usually try to project on the signal subspace (SVD,...). It is possible to take into account the structure (covariance) of the embedding noise: the SMF is used in such a point of view and in this case, a projection onto a  $p$ -dimensional subspace is made. In fact, this method is proven to be equivalent to a two-stage method: the whitening of the noise followed by the maximization of the SNR in a  $p$ -dimensional subspace. This method comes down to projecting the observation onto a subspace of dimension greater than one. However, there is no guarantee that the signal-to-noise ratio is maximum in

the subspace spanned by these vectors.

In this paper, we calculate a subspace whose dimension is chosen *a priori* and which is optimal in the sense that the SNR ratio is maximized within. We prove, through theoretical examples, that this subspace is not necessarily those spanned by the vectors calculated by the SMF. Through ROC curves, practical experimentations illustrate the interest of such an approach.

We have shown, with a practical example, that a noteworthy improvement can be reached with the ESMF applied in the optimal subspace calculated beforehand by the CSMF. It confirms that such a method is an interesting and powerful one to perform detection.

Prospects of applications of the CSMF can easily be imagined in image processing or stochastic transient signals detection (like acoustic signals). An extension to the classification problem is possible. Of course, as this method is a reduced-rank one performing SNR maximization, it can be used for data compression or estimation and filtering. Thus, the CSMF was successfully used with real signals:

- analysis of sequences of IR images (SATIR) to qualify high heat flux components (carbon bricks used in the ITER project with the CEA Cadarache): detection of defects and classification of components [23],
- detection and classification of textured images (like expanded polystyrene and textured paper or textured stones pictures for example): lot of images are texture images (forests, farming areas, ...) [1],
- detection and localization of very high energy neutrinos by a passive underwater acoustic telescope (ANTARES European project) [24],
- estimation of the sources in a specific blind source separation problem [25].

Reduced-rank estimators and filters are important for a wide range of signal processing applications, among others when data or model reduction, robustness against noise or model errors is desired. This concerns known methods like the *reduced-rank Wiener filter* (RRWF) by Scharf [17], the *reduced-rank maximum likelihood estimation* (RRMLE) by Stoica-Viberg [26], the *relative Karhunen-Loeve transform* (RKLT) by Yamashita-Ogawa [9] or the *generalized Karhunen-Loeve transform* (GKLT) by Hua-Liu [27], which is used for data compression and filtering and is in fact nothing else than the RRWF also called *low-rank Wiener filter* in [28] section 8.4.

The choice of the optimal dimension  $p$  of the subspace of projection is an important question which must be examined in more detail in the future.

## APPENDIX A : MATLAB<sup>®</sup> CODE.

```

► In the main program
B = 1260 % (Hz) noise bandwidth
f0 = 3131 % (Hz) central frequency
fe 15324 % (Hz) sampling frequency
Z = reponse(f0,0.25,fe); % generation of a narrow-band signal
signal = create(Ls,Z); % Ls = length of the signal

```

Z = reponse(f0,B,fe); % generation of the noise  
 noise = create(Ln,Z); % Ls = length of the noise

```

► subroutine 1
function y = create(lg,filtre)
n = length(filtre);
X = randn(1,lg+n);
X1 = conv(X,filtre);
X1 = X1/std(X1);
y = X1(n:length(X1)-n);
end

► subroutine 2
function Z = reponse(F0,B,fe)
Te = 1 / fe;
n0 = round( 1/(2*B*Te));
n = 1 : (2 * n0 + 1);
Z = cos(2*pi*F0*(n-1-n0)*Te).*sin(2*pi*B*(n-1-n0)*Te)./(pi*(n-1-n0)*Te);
Z(n0+1) = 2 * B;
end

```

The realizations  $\mathbf{s}_i$  ( $i = 1, \dots, N_r$ ) are generated by taking  $N$  consecutive points in the narrow-band signal, the first point being chosen randomly.

## REFERENCES

- [1] B. Borloz, *Estimation, détection et classification par maximisation du rapport signal-à-bruit : le filtre adapté stochastique sous contrainte*, Ph.D. Thesis, University of Toulon-France, Jun. 2005.
- [2] H.L. Van Trees, *Detection, Estimation and Modulation Theory*, Radar-Sonar Signal Processing and Gaussian Signals in Noise, Wiley, Part III. 2001
- [3] H.V. Poor, *An Introduction to Signal Detection and Estimation*, Springer-Verlag, 1988.
- [4] T. Kailath, H.V. Poor, *Detection of Stochastic Processes*, IEEE Transactions on Information Theory, Vol.44, no.6, pp.2230-2259, Oct. 1998.
- [5] A. Hero, *Signal Detection and Classification*, in 'The Digital Signal Processing Handbook' by Vijay K. Madisetti. CRC Press LLC, Series: Electrical Engineering Handbook, 1997. <http://www.eecs.umich.edu/~hero/Preprints/crc.article.ps.Z>
- [6] J.F. Cavassilas, B.Xerri, *Extension de la notion de filtre adapté. Contribution à la détection de signaux courts en présence de termes perturbateurs*, Revue Traitement du Signal vol.10 no.3, pp.215-221, 1992.
- [7] P.P.Pokharel, U. Ozertem, D. Erdogmus and J.C.Principe *Recursive complex BSS via generalized eigendecomposition and application in image rejection for BPSK*, Signal Processing, vol.88, pp.1368-1381, 2008.
- [8] J. Yang, H.S. Xi, F. Yang and Y. Zhao *RLS-based adaptive algorithms for generalized eigendecomposition*, IEEE Trans. Signal Processing, vol.54, no.4, pp.1177-1188, Apr. 2006.
- [9] Y. Yamashita, H. Ogawa, *Relative Karhunen-Loeve transform*, IEEE Trans. Signal Processing, vol.44, pp. 371-378, Feb. 1996.
- [10] L.L. Scharf and B. Friedlander: *Matched Subspace Detectors* IEEE Trans. Signal Processing, vol.42, no.8, pp. 2146-2157, Aug. 1994.
- [11] Mukund N. Desai and Rami S. Mangoubi, *Robust Gaussian and Non-Gaussian Matched Subspace Detection*, IEEE Trans. Signal Processing, vol.51, no.12, pp.3115-3127, Dec. 2003.
- [12] Mukund N. Desai and Rami S. Mangoubi, *Robust Subspace Learning and Detection in Laplacian Noise and Interference*, IEEE Trans. Signal Processing, vol.55, no.7, pp.3585-3595, Jul. 2007.
- [13] S. Kraut, L.L. Scharf and L. Todd McWhorter *Adaptive Subspace Detectors* IEEE Trans. Signal Processing, vol.49, no.1, pp.1-16, Jan. 2001.
- [14] S. Kraut and L.L. Scharf *The CFAR Adaptive Subspace Detector is a Scale-Invariant GLRT* IEEE Trans. Signal Processing, vol.47, no.9, pp.2538-2541, Sept. 1999.
- [15] L.L. Scharf and S. Kraut *Geometries invariances, and SNR interpretations of matched and adaptive subspace detectors* Traitement du Signal, vol.15, no.6, pp.527-534, 1998.
- [16] L.L. Scharf, Shawn Kraut, Michael L. McCloud *A Review of Matched and Adaptive Subspace Detectors* Proc. Symp. on Adaptive Systems for Signal Process. Commun. and Control, Lake Louise, Alta., Canada, Oct. 2000.
- [17] L.L. Scharf and M.L. McCloud *Matched and Adaptive Subspace Detectors When Interference Dominates Noise* Proc. Asilomar '00, Monterey, CA, Oct. 2000 (invited).
- [18] S. Verdu, H.V. Poor, *Signal Selection for Robust Matched Filtering*, IEEE Trans. on Communications, vol.COM-31, no.5, pp.667-670, May. 1983.
- [19] S. Verdu, H.V. Poor, *Minimax Robust Discrete-time Matched Filters* IEEE Trans. Communications, vol.COM-31, no.2, pp.208-215, Feb. 1983.
- [20] J.G. Proakis, *Digital Communications*, McGraw-Hill, Inc., third edition, 1995.
- [21] Y.C. Eldar, A.V. Oppenheim, *Orthogonal Matched Filter Detection*, Proceedings on the International Conference Acoustics, Speech, Signal Processing (ICASSP-2001) Salt-Lake, UT, May. 2001.
- [22] R.A. Horn and C.R. Johnson, *Matrix analysis*, Cambridge University Press, ISBN 0-521-38632-2, 1999.
- [23] F. Cismondi, B. Xerri, C. Jauffret, J. Schlosser, N. Vignal and A. Durocher *Analysis of SATIR Test for the Qualification of High Heat Flux Components: Defect Detection and Classification by Signal-to-Noise Ratio Maximization* Physica Scripta, T. 128, pp 213-217, Mar. 2007.
- [24] N. Juennard, C. Jauffret, B. Xerri *Detection and Localization of Very High Energy Neutrinos by a Passive Underwater Acoustic Telescope* Passive'08, IEEE OES, Hyères, 14-17, Oct. 2008.
- [25] B.Xerri, B.Borloz *Detection by SNR maximization: application to the blind source separation problem*, 5th International Conference On Independent Component Analysis and Blind Signal Separation. Sept. 22-24 2004. Grenade (Spain)
- [26] P. Stoica and M. Viberg *Maximum Likelihood Parameter and Rank Estimation in Reduced-Rank Multivariate Linear Regressions*, IEEE Trans. Signal Processing, vol.44, pp. 3069-3078, Dec. 1996.
- [27] Y. Hua and W. Liu, *Generalized Karhunen-Loeve Transform*, IEEE Signal Processing Lett., vol.5, pp.141-142, Sept. 1994.
- [28] L.L. Scharf *Statistical Signal Proces.: Detection, Estimation and Time Series Analysis* Reading MA; Addison Wesley, 1991.

## ANNEX : NOTATIONS

$\mathbf{C} \equiv [c_{ij}]$	:	matrix of entries $c_{ij}$
$\mathbf{C}^t$	:	transpose of $\mathbf{C}$
$\mathbf{C}^{-1}$	:	inverse of nonsingular $\mathbf{C}$
$\text{tr}\mathbf{C}$	:	trace of $\mathbf{C}$
$\mathbf{I}$	:	$N \times N$ unity matrix
$\mathbf{0}$	:	$N \times N$ null matrix
$\Delta_\lambda \equiv [\lambda_i]$	:	diagonal matrix of entries $\lambda_i$
$\mathbf{v}$	:	column vector
$E[\cdot]$	:	expected value of $[\cdot]$
$\delta_{ij}$	:	Kronecker delta of rank 2
$\mathbf{s}$	:	signal of interest
$\mathbf{n}$	:	corrupting noise
$\mathbf{A}$	:	$N \times N$ full-rank covariance matrix of $\mathbf{s}$
$\mathbf{B}$	:	$N \times N$ full-rank covariance matrix of $\mathbf{n}$
$\lambda_i$	:	eigenvalue of $\mathbf{B}^{-1}\mathbf{A}$ (with $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N \geq 0$ )
$\mathbf{u}_i$	:	eigenvector of $\mathbf{B}^{-1}\mathbf{A}$
$\mu_i$	:	eigenvalue of $(\mathbf{A} - \rho\mathbf{B})$
$\mathbf{t}_i$	:	eigenvector of $(\mathbf{A} - \rho\mathbf{B})$
$\mathcal{E}_p$	:	subspace of dimension $p$
$\mathcal{E}_p^\dagger$	:	subspace of dimension $p$ spanned by $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$
$\mathcal{E}_\mathbf{u}$	:	subspace of dimension 1 spanned by $\mathbf{u}$
$\mathcal{E}_p^*$	:	optimal subspace of dimension $p$