# DOA ESTIMATION IN THE PRESENCE OF UNKNOWN COLORED NOISE, THE GLOBAL MATCHED FILTER APPROACH.

*Jean-Jacques Fuchs*

IRISA / Université de Rennes 1
Campus de Beaulieu - 35042 Rennes Cedex - France
phone: + (33) 299847223, email: fuchs@irisa.fr
web: www.irisa.fr/temics/staff/fuchs/

**ABSTRACT**

The problem of the localization of multiple narrow band sources in the presence of arbitrary noise of unknown spatial spectral density is addressed. The array geometry can be arbitrary but must be known. The spatial noise spectrum is described using a sufficiently rich class of models that somehow covers the set of rational spectra. The Global Matched Filter is used to identify the characteristics of the sources that are present and to get a approximate model of the unknown colored noise. It is a technique that can be seen as a model-fitting or sparse representation approach in which the observations are decomposed on the association of different bases of candidate models. The computational complexity is reasonable and the performance are quite good and compare favorably with other methods.

## 1. INTRODUCTION

Estimating the directions of arrivals (DOA) of narrowband sources impinging on an array of sensors has applications in many different fields. Numerous investigations have been performed to investigate the performance of the more or less sophisticated methods developed to detect and locate the sources. Most of them assume and quite often require that the additive noise be spatially white (i.e. uncorrelated between sensors) or the noise covariance matrix be known up to a multiplicative constant. In most practical situations this is not the case and important degradations do appear, in terms of bias, poor resolution, spurious peaks, non detection of weak sources.

Many solutions have already been proposed. One possibility is to bypass or avoid the difficulty by using either higher order statistics [1] or instrumental variable like methods [2]. In the first case, one assumes that the noise is Gaussian while the source signals are non-Gaussian, in the second case the assumption is made that the spatial correlation length of the sources is much larger than that of the noise. Even if these assumptions are valid, they lead to a loss in information. For a given number of available samples, the estimation of higher order moments is less accurate than the second order moments and the same remark holds within the covariance sequence, the initial terms are better identified as well.

The second and by far the most investigated possibility is to introduce a parametric model for the noise covariance and to identify the parameters of that model together with the parameters of interest from the observations. Basically two types of models are considered. A number of contributions [3], impose a parametric model on the noise in the spatial domain (i.e. along the array), usually autoregressive (AR) models are used. A more abstract model consists in using a pre-specified set of array-geometry-dependent matrices to model the noise covariance matrix [4]. In both models, the goal is to approximate the true noise covariance matrix by a parametrized model whose order (the number of components, the number of weights to be adapted) has to be fixed a priori. Indeed, since -in general- maximum likelihood (ML) approaches are then used to identify all the unknown parameters: the sources DOA's, the sources powers and the noise parameters, the precise orders have to be fixed a priori.

Direct optimization of the resulting multimodal nonlinear likelihood function is a difficult task and the main contribution of the proposed methods lies in the reduction of dimensionality and/or the tentative separation of the parameters of interest from the others. A full separation seems unfeasible and a high-dimensional optimization problem must be solved.

Global optimization algorithms such as simulated annealing (SA) [5] , genetic algorithms (GA) [6] or particle swarm optimization (PSO) algorithm [7] have been considered to solve this complex, multimodal, highly non-linear optimization problem.

Now, of course optimizing the ML functional is a valuable approach only if the underlying pre-specified parametric model is adequate. Strictly speaking it should be the exact model. Since over-parametrization induces ill-conditioning and local minima and under-parametrization induces a lack of adequate model, a prior difficulty with all these approaches lies in the choice of the model order, of the number of free parameters.

The contribution is organized as follows. Section II describes the model of the signals while in section III two different noise models are introduced. The Global Matched filter is briefly described in section IV and applied to the present context in section V. Simulations and concluding remarks follow in section VI and VII.

## 2. PROBLEM FORMULATION

We consider the problem of estimating the direction of arrivals (DOA) of $P$ narrowband sources impinging on a array of $N$ sensors. To simplify the exposition we limit ourselves to the one dimensional localization problem (the azimuth angle $\theta$), i.e., we assume that the sources and sensors are coplanar and that the sources are in the far field. We denote $Z_k$ the k-th snapshots, an N-dimensional vector of the array outputs (after Fourier transformation and selection of the appropriate frequency bin). This vector can be modeled as

$$Z_k = As_k + n_k$$

with $A$ the $N \times P$ matrix with columns the steering vectors $a(\theta_p)$ for $p = 1$ to $P$, $s_k$ the $P$-dimensional signal vector with components $s_k(p)$ and $n_k$ the $N$-dimensional additive spatial noise vector. The signals $s_k(p)$ and noises are wide-sense stationary complex valued random processes with zero mean.

We assume the source signals to be uncorrelated and uncorrelated from the spatial noise and denote $Q$ the covariance matrix of the spatial noise that is an arbitrary positive definite unknown hermitian matrix. It follows that

$$R = E(Z_k Z_k^*) = ASA^* + Q, \text{ with } S = E(s_k s_k^*). \quad (1)$$

with $S = \text{diag}(\alpha_p)$ the diagonal matrix of the source powers.

The steering vector $a(\theta)$ is of the form

$$a(\theta) = [\, e^{-2j\pi d_1} \;\; e^{-2j\pi d_2} \; ... \; e^{-2j\pi d_N} \,]^T \quad (2)$$

with $d_k$, the distance, expressed in wavelengths, between sensor $k$ and the reference sensor projected on the direction $\theta$ of the potential source. This means that some calibration has been performed so that each sensor has nominal gain 1 and nominal phase 0.

 ISSN 2076-1465 

## 3. MODELING THE NOISE FIELD

The noise in the receiving system is composed of internal and external noise. The internal noise mainly consists in thermal noise in the receivers but other contributors may exist. If the thermal noise is dominant, a scaled identity matrix is a adequate model for the covariance matrix of internal noise.

The external noise is the result of the combination of all the unmodeled and unwanted signals that are intercepted by the sensors. It is difficult to assume that this ambient noise is uncorrelated from sensor to sensor. One can however assume that, over the considered time-spans, this noise is stationary with respect to time and one can then consider modeling it as the contribution of infinitesimal independent sources in the far-field of the array with azimuth dependend power $p(\theta)$ that varies smoothly with $\theta$. Sharp changes in $p(\theta)$ would be associated with strong localized noise sources that cannot be distinguished from the point sources of interest that are to be localized. It is then natural to represent $p(\theta)$ by its Fourier series expansion and to expect that, due to its smoothness, only few terms will be needed to obtain a quite adequate representation.

The ambient noise covariance matrix $Q$ is then given by

$$Q = \gamma I + \int_{-\pi}^{\pi} p(\theta) a(\theta) a(\theta)^* d\theta,$$

where the first part represents the internal noise contribution. Its $q$-th order expansion is then

$$Q \simeq \gamma I + \sum_{m=-q}^{q} c_m R_m \tag{3}$$

with

$$R_m = \int_{-\pi}^{\pi} e^{im\theta} a(\theta) a(\theta)^* d\theta$$

and

$$c_m = \frac{1}{2\pi} \int_{-\pi}^{\pi} p(\theta) e^{-im\theta} d\theta.$$

For a linear array the range is limited to $]-\pi/2,\ \pi/2[$ and a similar model with the corresponding modifications applies.

Another way to model the noise and its covariance matrix is to use complex autoregressive/autoregressive moving average (AR/ARMA) models along the array. And, since complex AR(1) models can be seen as the building block of the more general ARMA models, a sum of complex AR(1) processes is a sufficiently rich model that allows to represent with any desired accuracy, any covariance sequence and associated spatial spectral density function. In the present context, the use of these models *along the array* amounts to assume that the noise present at one sensor can somehow be predicted from the noise present at the neighboring sensors. The corresponding noise covariance matrix is then the sum of hermitian full rank hermitian Toeplitz matrices, the Toeplitz matrices whose first column is the partial covariance sequence associated with a complex AR(1) process. A complex AR(1) model satisfies $z_n = \beta z_{n-1} + e_n$ where $\beta = \rho e^{i\varphi}$ with $\rho \in [0,\ 1[$ and $\varphi \in [0,\ 2\pi[$ and $e_n$ is zero complex circular Gaussian noise with variance $\sigma^2$. This means that the real and imaginary parts of $e_n$ are independent white Gaussian noises with variance $\sigma^2/2$ each. The associated covariance sequence satisfies then $r_k = E(z_n \bar{z}_{n-k}) = \beta^k r_0$ for $k \geq 0$ with $r_0 = \sigma^2/(1-\rho^2)$ and the associated covariance function, needed for arbitrary geometry arrays, is $r(\tau) = r(0)\beta^\tau$ with $\tau$ expressed in half-wavelengths. The noise covariance matrix will thus be modeled as, compare with (3)

$$Q \simeq \gamma I + \sum_m \gamma_m T(\beta_m), \tag{4}$$

where $T(\beta)$ is the full rank, positive definite hermitian Toeplitz matrix associated with AR(1) process with parameter $\beta$ and $r_0 = r(0) = 1$ on the diagonal, to fix ideas.

Indeed since one is not interested in the model of the noise per se, since the noise is considered as a nuisance, all one needs is a model that is sufficiently rich. This is the case of both models introduced above.

## 4. THE GLOBAL MATCHED FILTER

Let us briefly sketch the DOA estimation procedure that will be used in the sequel.

### 4.1 Introduction

It is a high resolution DOA estimation scheme for which no preliminary decision has to be made as to its complexity. In the present context, this means that it is not necessary to fix a priori neither the number of point sources that are present nor the number of contributors in the model of the noise covariance matrix (3,4). The scheme is called the Global Matched Filter (GMF) in [10, 11] and it can be seen as a sparse representations technique, a model-fitting approach or an inverse-problem solver. It works whenever one wants to decompose a vector of observations into the sum of a small number of vectors belonging to a known parametrized family of vectors. This is the case in the present source localization context when, for instance, one considers as vector of observations a set of beamformer outputs.

The beamformer output at azimuth $\phi$ is defined to be

$$y(\phi) = a(\phi)^* R a(\phi) \tag{5}$$

with $a(\phi)$ in (2), and with the notations introduced above (1) and using (3) to model $Q$ the noise covariance matrix, one gets

$$y(\phi) = \sum_{p=1}^{P} \alpha_p |a(\phi)^* a(\theta_p)|^2 + \gamma N + \sum_{m=-q}^{q} c_m a(\phi)^* R_m a(\phi). \tag{6}$$

It can indeed be seen as being the sum of the contributions of the $P$ sources and of the different spatial noise components. Since $R$ is not available estimated beam outputs, say $\hat{y}(\phi)$ are obtained by replacing $R$ by its estimate

$$\hat{R} = \frac{1}{T} \sum_{1}^{T} Z_k Z_k^*, \tag{7}$$

the so-called snapshot covariance matrix. The GMF uses as input an $L$-dimensional vector with components $\hat{y}_k = \hat{y}(\phi_k)$. We will denote $\hat{Y}$ this vector filled with these $L$ beams. The value retained for $L$ is equal to the number of real degrees of freedom in $R$ so that there is no information loss in replacing $\hat{R}$, the usual input to most DOA estimation schemes, by $\hat{Y}$. The bearings $\phi_k \in \Phi$ depend upon the array geometry. For a uniform linear array (ULA), $L = 2N-1$ and the $\phi_k$'s are equispaced in spatial frequencies.

The decomposition (6) of $y(\phi)$ extends to the vector $Y$

$$Y = \sum_{p=1}^{P} \alpha_p\ f(\theta_p)\ +\ \gamma\, N\mathbf{1} + \sum_{-q}^{q} c_m g_m,$$

where $f(\theta)$ denotes the $L$-dimensional vector of the contribution of a source with bearing $\theta$ and unit power to the beams in $Y$, $\mathbf{1}$, a vector of ones, that allows to model the contribution of the spatially white noise to $Y$ and $g_m$ the vector with the contribution to the beams of the $R_m$ matrices in (3).

The aim of the GMF is to recover this sparse exact representation of $Y$ from the observation of its noisy estimate $\hat{Y}$.

One therefore introduces a set of $M$ $L$-dimensional vectors $f_m = f(\psi_m)$ with $\psi_m \in \Psi$, a set of $M \gg L > P$ bearings representing the positions of all the potential sources. One builds the $L \times M_F$ matrix $F$ with $M_F = M + 2 + 2q$ columns: the $f_m$'s, $N\mathbf{1}$ the contribution of the spatially white noise and the $g_m$'s. A sparse representation of $\hat{Y}$ is then of the form $FX$, with $X$ a vector of weights having just a few non-zero components. Since the true bearings $\theta_p$ do generically not belong to the discretization grid points $\Psi$, two columns of $F$ will, in general, be needed to approximatively model the contribution of each true source. A typical sparse $X$ representing $\hat{Y}$ will thus have about $2P$ non-zero components to represent the sources plus a number of components to model the noise contribution.

Quite specifically a preliminary version of the GMF amounts to solve the optimization problem

$$\min_X \frac{1}{2}\|F_nX - \hat{Y}\|_2^2 + h\|X\|_1, \tag{8}$$

with $\|X\|_1 = \sum|x_k|$, $\|X\|_2^2 = x_k^2$, $F_n$ the $F$ matrix with its columns normalized to one in Euclidean norm and $h$ a positive real to be fixed by the user. This is a convex program, for which fast dedicated algorithms are available, and one deduces from its unique optimum the different estimates of interest: the azimuths $\theta_p$, powers $s_p$ and also the source number $P$ since this number as well as the order of the noise model has not to be fixed a priori.

### 4.2 The standard version

We now take into account the statistical properties of the observations in $\hat{Y}$ to develop the standard version of GMF. Under the current assumptions $\hat{R}$ is an estimate of the covariance matrix of $Z_k \in CN(0,R)$ and it follows that $T\hat{R}$ is a sample of a complex Wishart distribution $CW(T,N,R)$ [12]. The statistical properties of the components $\hat{y}_k$ of $\hat{Y}$ are then easy to obtain. With $\Sigma$, denoting the covariance matrix of $\hat{Y}$ one has

$$\Sigma_{k,l} = \frac{1}{T}|a(\phi_k)^* R a(\phi_\ell)|^2.$$

It is then natural to premultiply both $F$ and $\hat{Y}$ in (8) by $\Sigma^{-\frac{1}{2}}$ to *whiten* the observations in $\hat{Y}$. Since $R$ and thus $\Sigma$ are not known, in practice one replaces $R$ by its estimate $\hat{R}$, to get an estimate $\hat{\Sigma}$ of $\Sigma$. The components in the resulting observation vector, say, $\hat{Y}_w = \hat{\Sigma}^{-\frac{1}{2}}\hat{Y}$ are then, asymptotically in $T$, uncorrelated and of unit variance.

This version of GMF, which is quite close to a maximum likelihood approach applied to the observations in $\hat{Y}$, amounts to replace (8) by

$$\min_X \frac{1}{2}\ \|F_{wn}X - \hat{Y}_w\|_2^2\ +\ h\ \|X\|_1, \tag{9}$$

with $\hat{Y}_w = \hat{\Sigma}^{-\frac{1}{2}}\hat{Y}$, $F_w = \hat{\Sigma}^{-\frac{1}{2}}F$ and $F_{wn}$ represents the matrix $F_w$ with columns normalized to one in $\ell_2$ norm. Just as (8), this is a convex program and all the estimates are deduced from its generically unique optimum.

### 4.3 Implementation issues

Let us sketch briefly how to implement in practice (9) or (8), further details can be found in [11] or more recently in [13].

The number of parameters to be tuned is quite small. The choice of an adequate number $L$ of components in $\hat{Y}_w$ has already been discussed above. The choice of $M$ the number of columns in $F$ devoted to the potential point sources and the associated discretization step in bearing should be fixed according to the resolution possibilities of the array. As an example, for a uniform linear array (ULA) with $N$ sensors a half wavelength apart and signal to noise ratios around 0 dB, the resolution limit (Rayleigh limit) in spatial frequency is about $\Delta f = 1/N$ and a high resolution method able to separate sources at $\Delta f = 1/2N$, so that one takes $M = 10N$. This choice allows for about 2 zero weights between two sources that are close in bearings but nevertheless potentially separable and thus guarantees that two disjoint clusters on nonzero weights in $X$ will be obtained in case the two sources are detected.

For a well chosen $h$, the optimal $X$ will have ideally about $2P$ nonzero components among its first $M$ components. A pair of neighboring nonzero components for each of the sources that are present. Other nonzero components will be present in the latter components of the optimal $X$ to model the noise contribution. The estimate of the number of waveforms present is then given by the number of (significant) clusters of nonzero components, the power of a source is estimated by the sum of the weights in its associated cluster and the azimuth estimate is obtained by linear interpolation of the associated indices of $\psi_m \in \Psi$.

The choice of $h$ is crucial, it varies linearly with the standard deviation of the (estimation) noise affecting the components in $\hat{Y}$. For a standard deviation equal to one, as is the case in $\hat{Y}_w$, one should take $h \simeq \sqrt{2\ln 2L}$. Roughly speaking, the larger $h$, the sparser the optimal $X$ and vice versa.

If the value of $h$ is too large, the procedure may not detect weak sources and if $h$ is too small, there may appear many false alarms, i.e., the procedure might detect sources that do not exist.

Eventually, one should mention that the presence of $h$ in (8) (9) induces bias into the estimates. This bias while concerning mainly and directly the amplitude estimates, also affects slightly the bearing estimates.

Let us summarize the algorithm. For a given array, we build the $F$ matrix whose columns model the contributions of both the point sources and noise contributors to the observations vector $Y$ build using $L$ beam outputs. The beams are computed according to (5) with $R$ replaced by the estimate $\hat{R}$ (7) of the covariance matrix of the snapshots. To implement the standard version (9), one further has to whiten $\hat{Y}$ and normalize the columns of the whitened $F$-matrix. One then solves (9) with $h$ in between $\sqrt{2\ln 2L}$ and $\sqrt{2\ln 2M_F}$, and deduces the estimates from its optimum.

## 5. DEVELOPMENT

Let us come back to the problem we are concerned with, namely the presence of arbitrary and unknown spatial noise.

To fix ideas, unless otherwise stated, a linear arrays with equispaced sensors, one half wavelength apart, with $N = 10$ sensors and $T = 100$ snapshots will be considered. The steering vector (2) associated with a source becomes then

$$a(\theta) = [\ 1\ e^{-i\pi\sin\theta}\ \ e^{-2i\pi\sin\theta}\ \dots\ e^{-(N-1)j\pi\sin\theta}\ ]^T, \tag{10}$$

where $\theta \in ]-\pi/2,\ \pi/2[$ is the bearing of the source with respect to broadside.

If the non-whiteness of the noise is not taken into account the degradation in performance can be important. While the basic beamformer, with its poor resolution performances, is quite robust, more sophisticated methods are more sensitive. For spatial noise with relatively smooth spectral densities, one will observe bias and higher variance, but as soon as the spectral density of the noise becomes slightly spiky, spurious sources will appear and the detection of weak sources will become difficult together with the determination of the *rank* of the covariance matrix of the snapshots on which most high resolution methods, such as MUSIC [8, 9], rely.

The technique of noise modeling using the Fourier series expansion (3) described at the beginning of section 3 allows to model essentially very smooth noise spectra or, at the least, requires a large number of components to represent an even only slightly resonant spectrum. It is thus more adapted to model quite smooth noise spectra that will generally essentially induce bias and additional variance in the DOA estimates.

Since the more challenging situation where colored noise may be confused with sources is to be considered here, only the complex AR modeling approach (4) will be considered in the sequel. The implementation within the proposed scheme is then slightly more complex since instead of having just $2q+1$ (see (3)) additional columns in $F$, a larger number of additional columns is required. Indeed to cover potentially all the complex AR(1) processes, one has to discretize $\beta = \rho e^{i\varphi}$ over its domain namely $\rho \in [0,\ 1[$ and $\varphi \in [0,\ 2\pi[$. For $\rho = 1$ the AR process becomes singular and indistinguishable from a point source (10), one therefore limits the domain of variation of $\rho$ to $[0,\ .9\ ]$ in the sequel. To represent all the potential noise spectra, $n_{rho} = 9$ and $n_\varphi = 12$ different values of $\rho$ and $\varphi$, i.e. $\rho =\ 0.1\ k$ for $k = 1$ to 9 and $\varphi =\ (\pi/6)\ k$ for $k = 0$ to 11. This leads to the construction of $n_\rho \times n_\varphi = 108$ columns representing the potentials noise contributions to be added in the $F$ matrix. This number is quite small and the increase in computation time it induces is negligible. It will be shown in the simulations below (see section 6.4) that this approach allows to handle smooth spectra, belonging to the other models, as well.

## 6. SIMULATIONS RESULTS

Simulation results are presented to assess the potentialities of the proposed method. A Uniform Linear Array with $N = 10$ sensors is considered and $T = 100$ snapshots are used. The matrix $F$ has 19 rows (since an hermitian Toeplitz matrix of order 10 has 19 real degrees of freedom) and 209 columns, 100 equispaced columns to model the potentials point source contributions, one to model the white noise contributions and 108 to model the complex AR(1) potentials noise contributions as described in section 5. The hyperparameter $h$ is taken equal to $\sqrt{2ln2L} \simeq 2.70$ as recommended above.

### 6.1 Example 1.

Let us first comment on the small *distance* that exists between a complex AR(1) process with $\rho = .9$ and a point source. In the beamformer output, they lead to extremely similar shapes, see Figure 1a and 2 below. Though full rank, the covariance matrix associated with such a process has essentially one large eigenvalue and makes all source number determination techniques based on the eigenvalues inoperable. In the present Global *Matched Filter* context, the correlation between the normalized column in $A$ associated with such a process and the most resembling normalized column in $A$ associated with a source is 0.986, i.e. the angle between these two columns is less than 10 degrees. It appears that nevertheless the algorithm is in general not mislead. If one simulates an complex AR(1) process with $\beta = \rho = .9$ and variance $r_0 = 1.39$ which mimics quite precisely a point source at 0 degree and unit power, the GMF systematically identifies the AR process and only in 25% of the realizations detects an additional source around 0 degree with quite small average power 0.13, about one tenth of the "true" value. This is of course without giving any prior information about the true scenario to the algorithm.

### 6.2 Example 2.

In a second couple of simulations, one considers a single resonant complex AR(1) process with $\rho = 0.9$ and $r_0 = 1$ and 2 close sources with bearings $-5$ and 0 degrees and power $s_1 = s_2 = 1$. The diagonal of the exact snapshot covariance matrix is thus equal to 3. The resonance of the noise is first (a) placed around -30 degrees ($\varphi = \pi/2$) and then (b) placed around 0 degrees $\varphi = 0$). In both cases 1000 independent realizations are performed and the estimates of the bearings and amplitudes of the two sources are presented in Table 1.

The output of the beamformer (for the exact covariance matrix) corresponding to both scenarios are presented in Figure 1.

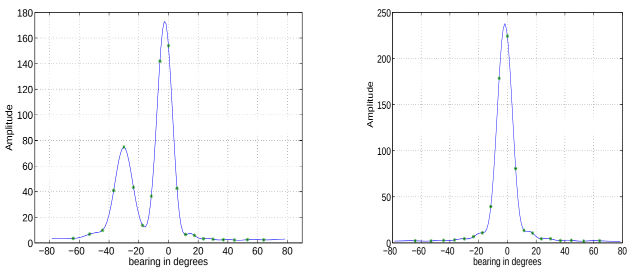


Figure 1: *The beamformer output: there is one resonant complex AR(1) processes modeling the noise and two point sources at -5 and 0 degrees. Left: the resonant noise contribution is around -30 degrees, right: it is around 0 degrees. The 19 stars are the beams used as input to the localization and detection algorithm.*

The results obtained are quite similar in both cases (a) and (b) with however a slight bias on the bearings in case (a). The performances are close to the Cramer-Rao bounds but of course at no moment one has to indicate to the algorithm that there are two sources and one AR(1) process to be estimated. The GMF decides by itself the model that best fits the observations in $\hat{Y}$. For both cases, there are, in the average, about 5.1 nonzero components in the optimal $X$ of dimension 209. Among the the first 100 components in $X$ associated with the potential point sources, 3.6 of them correspond to the two (true) sources and once they are removed, there remains one nonzero weight in about 25 % of the realizations in case (a) and less than 1 % in case (b). This nonzero weight corresponds to a false alarm (the detection of a spurious source) has average value .09 in both cases, quite weak spurious peaks that are easily discarded, since about 10 times smaller than the true sources.

Table 1: *Two equipowered sources with bearings $-5$ and 0 degrees in quite resonant AR(1) noise, T=100 snapshots: (a) noise resonance at -30$^o$, (b) noise resonance at 0$^o$.*

| | (a) noise resonance at -30$^o$ | | | | (b) noise resonance at 0$^o$ | | | |
|---|---|---|---|---|---|---|---|---|
| | bearings | | amplitudes | | bearings | | amplitudes | |
| true | -5 | 0 | 1 | 1 | -5 | 0 | 1 | 1 |
| mean | -5.14 | -.11 | .88 | .95 | -4.95 | .04 | .94 | .94 |
| st dev | .23 | .22 | .14 | 0.13 | .29 | 0.35 | .19 | .24 |

*Estimates of the mean and standard deviation of the bearings and amplitudes averaged over 1000 independent realizations.*

Let us just mention that the closeness in bearings of the two sources makes this a difficult scenario and that even in the usual spatial white noise case with unit variance, the standard MUSIC algorithm fails to separate the two sources in more than 1 fourth of the realizations.

### 6.3 Example 3.

In the third scenario the noise is modeled using 3 AR(1) processes with variance equal to 1 each and parameters $(\rho, \varphi)$ respectively equal to $(0.7, 2\pi/3)$, $(0.8, \pi/3)$ and $(0.9, 0)$ and there is one source at 20 degrees with $s_1 = 1$ in (1), the diagonal of the exact global covariance matrix is thus equal to 4. This fully characterizes the simulation and, depending upon the definition, the signal to noise ratio is slightly below 0 dB.

The output of the beamformer to this scenario is presented in Figure 2, there are essentially 4 resonances associated respectively (from left to right) with the 4 contributing components in the order they are listed above.

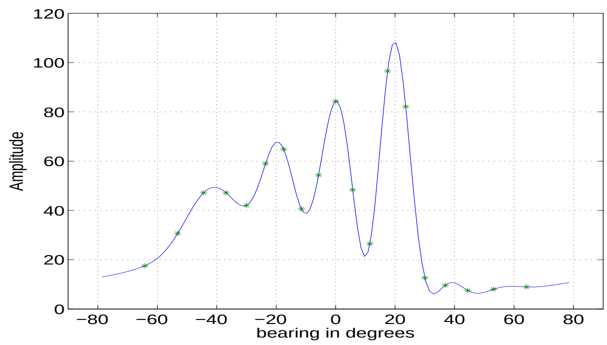


Figure 2: *The beamformer output: there are 3 more and more resonant complex AR(1) processes with contributions at -40, -20 and 0 degrees and a point source at 20 degrees. The 19 stars are the beams used as input to the localization and detection algorithm.*

Over 1000 independent realizations, the proposed algorithm always locates the unique source and, in 48% of the realizations, it locates a second point with much lower amplitude (one sixth) around 0 degrees in about half of these cases and around -20 and -40 degrees in about one quarter of these cases each. If the value of $h$ is doubled the number of false detection drops to less than one percent. The parameter $h$ can be shown to act as the threshold in a generalized likelihood ratio test and it allows to tune the probability of false alarm. The standard value recommended above corresponds to a 10% probability of false alarm in the presence of white spatial noise with unity variance.

### 6.4 Example 4.

Let us consider now a noise spectrum that does not belong to the complex AR class. The covariance function is generated using a Bessel function of the first kind and order zero, i.e. the first column of the hermitian Toeplitz matrix $Q$ in (1) is taken equal to $J_0(k\pi)$ with $k = 0$ to $N-1$. This corresponds indeed to the other type of models (3) considered above since the covariance matrix is also equal to $(1/\pi)R_0$, see (3). As in Example 2, the sources that are present are located at $-5$ and 0 degrees and have unit power $s_1 = s_2 = 1$. The estimates of the bearings and amplitudes of the two sources, obtained over 1000 independent realizations, are presented in Table 2. and the output of the beamformer (for the exact covariance matrix) is presented in Figure 3.

Table 2: *Two equipowered sources with bearings* $-5$ *and* 0 *degrees in Bessel noise , T=100 snapshots.*

| | bearings | | amplitudes | |
|---|---|---|---|---|
| true | -5 | 0 | 1 | 1 |
| mean | -5.12 | -0.44 | 0.65 | 0.64 |
| st dev | 0.31 | 0.36 | 0.15 | 0.21 |

*Estimates of the mean and standard deviation of the bearings and amplitudes averaged over 1000 independent realizations.*

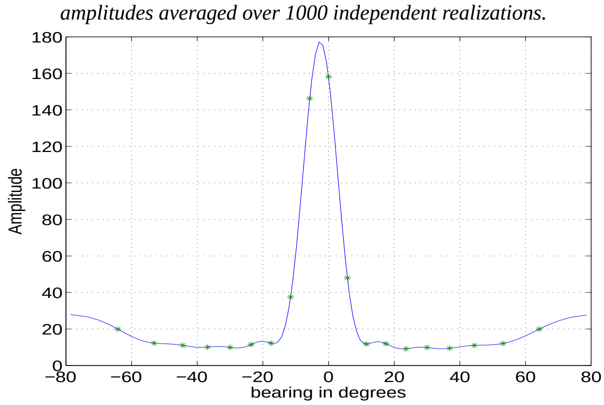


Figure 3: *The beamformer output: A spatially colored noise with Bessel covariance function and two equipowered point sources at -5 and 0 degrees. The 19 stars are the beams used as input to the localization and detection algorithm.*

The average number of non zero weights in the optimal $X$ (9) is 8.7, and 5.2 of them are used to model this noise that does not belong to the class considered, the remaining 3.5 concern the two point sources. In 15 % of the realizations there is one non-zero weight left, ones the two sources are located, but its amplitude is 0.03 in the average and it can thus be discarded and considered as a false alarm. One could add that, for this scenario, the basic Music algorithm separates the two sources in less than 10% of the realizations, although the noise spectrum seems to be close to a white noise spectrum.

## 7. CONCLUSIONS

This paper presents an approach that allows to localize sources in the presence of additive noise of unknown spectrum. It is indeed one way to apply the Global Matched Filter [11, 13] to the unknown colored noise situation. It has been presented for the case of a uniform linear array but can be adapted to any array geometry.

The task that is considered is of great practical interest and has been considered for at least two decades. Competing algorithm do in general rely on a complete and precise parametric description of both the signals and noise covariances and aim to solve the associated Maximum Likelihood criterion. This is of course a difficult optimization problem and most papers are concerned with smart ways to make this problem solvable in practice. Suboptimal strategies whose aim is to decouple the estimation of the signal parameters from the noise parameters are often considered and sensitivity to the initialization is in general an important problem. This explains that new types of optimization routines, such as simulated annealing, genetic algorithms or particle swarm optimizations are being considered.

All these approaches nevertheless heavily rely on the validity of the model that is used, and not only upon the model itself but also quite drastically upon the precise tuning of the model, i.e. the number of assumed sources and the order (complexity) of the noise model has to be precisely fixed a priori. This is of course a penalizing situation because somehow the detection problem and the order determination has to be done beforehand and these (probably more complex) issues are seldom addressed.

The method proposed in this paper is only marginally concerned by these preliminary tuning issues. Their is no need to know a priori neither the number of sought sources, nor the noise type and noise model order. Some sort of detection scheme has to be added to the proposed procedure though, to discard occasional weak spurious peaks that are to be considered as false alarms. Indeed the only parameter whose tuning has a drastic effect is $h$ in (9) that is indeed strictly comparable to the tuning of the threshold in a detection test [14]. According to the performances observed in the simulation section, the value recommended for this parameter at the very end of section 4 seems to be quite adequate.

## REFERENCES


[1] J.F. Cardoso and E. Moulines, " Asymptotic performance analysis of direction finding algorithms based on fourth order cumulants," *IEEE-T-SP*, vol. 43, 1, pp. 214-224, 1995.

[2] P. Stoica et al, " Instrumental variable approach to array processing in spatially correlated noise fields," *IEEE-T-SP*, vol. 42, 1, pp. 121-133, 1994.

[3] J.P. Le Cadre, "Parametric methods for spatial signal processing in the presence of unknown correlated noise fields," *IEEE-T-ASSP*, vol. 37, 7, pp. 965-983, 1989.

[4] B. Friedlander and A. Weiss, "Dierction finding using noise covariance modeling," *IEEE-T-SP*, vol. 43, 7, pp. 1557-1567, 1995.

[5] S. Kirkpatrick et al., "Optimization by simulated annealing," *Science*, vol. 220, 5, pp. 671-680, 1983.

[6] D. Goldberg, *Genetic algorithms in search, optimization and machine learning.* Kluwer, Boston , 1989.

[7] M. Li and Y. Lu, "ML DOA estimation in unknown colored noise fields," *IEEE-T-AES*, vol. 44, 3, pp. 1079-1090, 2008.

[8] G. Bienvenu and L. Kopp, "Optimality of high resolution array processing using the eigensystem approach," *IEEE-T-ASSP*, vol. 31, 10, pp. 1235-1248, 1983.

[9] R.O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE T-AP* , vol. 34, 3, pp. 276-280, 1986.

[10] J.J. Fuchs, "Detection and estimation of superimposed signals," In *Proceedings ICASSP*, Seattle, 4, pp. 1649-1652, 1998.

[11] J.J. Fuchs, "On the Application of the Global Matched Filter to DOA Estimation with Uniform Circular Arrays," *IEEE-T-SP*, vol. 49, 4, pp. 702-709 , 2001.

[12] K.S. Miller, *Complex Stochastic Processes. An introduction to theory and applications.* Addison-Wesley, 1974.

[13] J.J. Fuchs, "DOA estimation in the presence of modeling errors, the Global Matched Filter approach." EUSIPCO, Glasgow, pp. 1963-1967, August 2009.

[14] J.J. Fuchs, "Joint detection and estimation using a sparse representation approach." submitted to ICISP, January 2010.