DOA ESTIMATION IN THE PRESENCE OF MODELING ERRORS, 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

Taking into account the modeling errors on the array manifold in direction of arrival (DOA) estimation is a crucial issue in practice. It is regularly considered in the literature but no satisfactory solution is known and only sensitivity or performance analyses of existing methods (usually subspace-based) are proposed mostly in the exact data case. We show that projecting the observations on the nominal array manifold is a solution to this problem that greatly reduces the degradations in performance induced by modeling errors. Except for some very specific array geometries, this projection is unfeasible in practice, hence the interest of the Global Matched Filter (GMF) which, when using as inputs a finite number of beamformer-outputs, benefits from the projection that is implicitly performed by the conventional beamformer. Its performance is quite remarkable.

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 a perfect match between the assumed model and the reality. Unfortunately, even after a calibration procedure, in an operational context the assumed model is always different from the true model and it is probably fair to say that the more a method is sophisticated the less it is robust to mismatch.

Several authors have studied the performance degradation due to the modeling mismatch in terms of mean square errors, see e.g. [1, 2, 3]. First order expansions, that are valid for small errors, are proposed in, in general, the infinite data case. Assuming a perfectly known covariance matrix allows, of course, to concentrate on the degradations due to modeling mismatch and avoids to mix them with discrepancies induced by the finite number of observations. When the modeling errors are small, they introduce small errors in the DOA's and do not affect the resolution properties. As the magnitudes of the errors increase they affect the resolution itself and there appears a threshold effect, a discontinuity that depends upon the difficulty of the scenario. Recently the notion of resolution probability has been investigated in this context [4]. It aims to give guidelines for setting the calibrations requirements in a true bearing estimation system. It has been developed for a method, MUSIC [5, 6], that has poor robustness properties in its basic form but that can easily be improved upon for some array geometries, as we shall see below.

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 θ), 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 $s_p(k)$ and n_k the N-dimensional additive spatial noise vector. The signals $s_p(k)$ and noises are wide-sense stationary complex valued random processes with zero mean. For simplicity, we assume that the spatial noise to be white with variance v and uncorrelated with the signals. It follows that

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

We assume the matrix S =diag(α_p) to be diagonal, i.e., we assume the signals emitted by the P sources to be uncorrelated.

In practice a number of different modeling errors coexist, such as errors in the sensor gains, sensor phases and sensor locations but also mutual coupling coefficients between sensors or local scattering for instance. In the sequel, we will model these errors as additive perturbations on the components of the nominal steering vector $a(\theta)$, that we assume to be of the form

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

with d_k, the distance, expressed in wavelengths, between sensor k and the reference sensor projected on the direction θ of the potential source. This means that for the nominal steering vector some calibration (or normalization) has been performed so that each sensor has nominal gain 1 and nominal phase 0.

In the presence of modeling errors, the nominal steering vector $a(\theta)$ becomes $\hat{a}(\theta)$ with

$$\hat{a}(\theta) = a(\theta) + w, \quad \text{with } w \in CN(0, \sigma^2 I). \quad (3)$$

The real $2N$-dimensional vector build with the real and imaginary part of w is Gaussian with zero mean and covariance matrix $(1/2)\sigma^2 I$. This perturbation model is of course quite rough but it is often considered [3, 4] and is general enough to get a good picture of the performances of an approach. We further assume that the perturbation w_p that affects the steering vector $a(\theta_p)$ is independent of the perturbation w_q that affects the steering vector $a(\theta_q)$.

3. THE GLOBAL MATCHED FILTER

We forget about the modeling errors in this section where we briefly sketch the DOA estimation that we will consider.

3.1 The elementary version

We summarize a high resolution DOA estimation scheme that simultaneously somehow detects the number of sources present. It is a sparse representations technique, we called the Global Matched Filter (GMF) in [7, 8], that can also be seen as 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

[7]. This is the case in the present source localization context when one considers as vector of observations a set of beamformer outputs. We will see that it generally outperforms MUSIC (at a higher computational cost though) and attains performances close to the Cramer-Rao bounds [8].

With the notations introduced above, the expected beamformer output power at azimuth ϕ is

$$y(\phi) = a(\phi)^* R a(\phi) = \sum_{p=1}^{P} \alpha_p |a(\phi)^* a(\theta_p)|^2 + Nv. \qquad (4)$$

It can indeed be seen as being the sum of the contributions of the P sources and the spatial noise. Its observed counterpart, say $\hat{y}(\phi)$ is obtained by replacing R by its estimate

$$\hat{R} = \frac{1}{T} \sum_{1}^{T} Z_k Z_k^*, \qquad (5)$$

the so-called snapshot covariance matrix. While $\hat{R}$ is the basic input to the MUSIC algorithm, 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 L beams evaluated at, e.g. equispaced, bearings $\phi_k \in \Phi$, where L depends upon the array geometry but is typically of the order of magnitude of N.

The decomposition of $y(\phi)$ observed in (4) extends to Y

$$Y = \sum_{p=1}^{P} \alpha_p \; f(\theta_p) \; + \; v \, N\mathbf{1},$$

where $f(\theta)$ denotes the L-dimensional vector of the contribution of a source with bearing θ and unit power to the beams in Y and $\mathbf{1}$, a vector of ones, that allows to model the contribution of the spatially white noise to Y. The aim of the GMF is to recover this sparse exact representation of Y from the observation of its noisy estimate $\hat{Y}$.

We next introduce a set of M L-dimensional vectors $f_m = f(\psi_m)$ with $\psi_m \in \Psi$ and $M \gg L \gg P$ and build the $L \times M+1$ matrix F with columns the f_m's and $N\mathbf{1}$ the contribution of the spatially white noise. A sparse representation of $\hat{Y}$ is then given by FX with X an $(M+1)$-dimensional vector of weights having just a few non-zero components. Since the true bearings θ_p do generically not belong to the discretization grid points Ψ, 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 plus one component to model the noise.

Quite specifically the most elementary version of the GMF amounts to solve the optimization problem

$$\min_X \frac{1}{2} \|F_n X - \hat{Y}\|_2^2 + h\|X\|_1, \qquad (6)$$

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 the azimuths θ_p, powers α_p, noise variance v and even source number P.

3.2 The standard version

We now take into account the statistical properties of the observations in $\hat{Y}$ to develop what we will call 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 then a sample of a complex Wishart distribution $CW(T,N,R)$ [9]. The statistical properties of the components $\hat{y}_k$ of $\hat{Y}$ are then easy to obtain, see Appendix B. With Σ, 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 (6) by $\Sigma^{-\frac{1}{2}}$ to *whiten* the observations in $\hat{Y}$. Since R and thus Σ are not known, in practice we replace R by its estimate $\hat{R}$, to get an estimate $\hat{\Sigma}$ of Σ. 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.

The more elaborate version of GMF, which is quite close to a maximum likelihood approach applied to the observations in $\hat{Y}$, amounts to replace (6) by

$$\min_X \frac{1}{2} \; \|F_{wn} X - \hat{Y}_w\|_2^2 \; + \; h \; \|X\|_1, \qquad (7)$$

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 ℓ_2 norm.

The number of waveforms P present, their powers and their bearings are then deduced from the optimum X of (7), obtained for an adequately chosen h, as explained in the following Section.

3.3 Implementation issues

For a well chosen h, the optimal X has typically about $2P+1$ nonzero components, i.e., a pair of neighboring nonzero components for each of the waveforms that are present and an additional component 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 $\psi_m \in \Psi$.

3.3.1 Setting the threshold

We consider the standard version of the proposed technique. The value of h will have a major influence upon the detection part of the whole procedure. To make this more apparent we introduce the dual of the criterion (7) which is

$$\min_X \|F_{wn} X\|_2 \quad \text{under} \quad \|F_{wn}^T (F_{wn} X - \hat{Y}_w)\|_\infty \le h \qquad (8)$$

where $\|X\|_\infty = \max_k |x_k|$. This dual which is equivalent to (7), is important because it allows to understand the role played by h [8]. One can already observe from the constraints in (8) that, as h increases, so does the size of the admissible domain. More precisely, for h larger than $\|F_{wn}^T \hat{Y}_w\|_\infty$, the point $X = 0$ becomes admissible and is then the optimum. 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 will detect sources that do not exist.

Due to limited space, we do not detail the justification of the choice of h we use in the sequel, namely $h \simeq \sqrt{2 \ln M}$ in (7) which should lead to roughly 10% of false alarms, see [8, 10] for details.

3.3.2 Getting the estimates

Having obtained the optimum of (7), denoted X below, we now explain how to deduce from X the estimates of the powers and azimuths. We first consider the case $P = 1$, where there is a single source present in the observations. For a well chosen h, the optimal X has then just two non-zero complex weights. A non-zero component associated with the noise contribution, i.e., column $N\mathbf{1}$ and another one associated with the column in F with bearing closest to the true bearing. The amplitude and bearing estimates are then simply this non-zero component (modified to take into account the normalization step of the columns of the F-matrix) and the bearing associated with the corresponding column.

The so-obtained bearing estimate being on the discretization grid may not be the best possible. A better choice of h (slightly smaller than the previous one) might yield a optimal X having three non-zero weights. Besides the 'noise' component, there would appear a couple of neighboring non-zero weights that, by simple linear interpolation, would yield a bearing estimate that does not lie on the discretization grid and may thus well be closer to the true azimuth. In case, there are more than 2 contiguous nonzero components in

X (which actually never happens) one similarly estimates the bearing by linear interpolation and the amplitude by simply adding the different weights.

If there are $P>1$ paths present in Y, one expects the optimal X to have, in addition to the noise component, $2P$ disjoints couples of non-zero weights and the same procedure applies to each couple. This, of course, corresponds to the ideal situation where the proposed algorithm does a perfect detection job.

3.3.3 *Detecting the number of paths*

An estimate of P, the number of sources present in Y, is given by the number of (significant) clusters of non-zero components in the optimal X. A cluster being typically a couple of neighboring non-zero weights. For a given scenario, the detection performance of the proposed approach depends mainly upon the choice of h in the criterion, see Section 3.3.1 above. While for easy scenarios this choice is quite robust, this is no longer true for difficult scenarios. Globally if h is fixed to get a probability of false alarms equal to P_{fa}, the procedure will detect an additional (false) bearing in P_{fa} percent of the realizations. This additional source will have a small amplitude estimate and may thus appear suspicious.

The choice of the number M of columns in F and the associated discretization step in azimuth or better, for a uniform linear array, in spatial frequency has also some importance in this context even though the optimization problems (6) or (7) are convex and have thus generically a unique optimum. If M is too large, it will essentially increase the computation time and lead to a poorly conditioned optimization problem.

If M is too small, it may happen that the two clusters associated with sources having close bearings merge which results in the non-separation of the two paths. This situation can and must be avoided by adequately choosing M. Indeed knowing roughly the resolution limit associated with the array, we propose to choose L so that GMF can detect two correspondingly spaced sources. As an example, for an array with N equispaced sensors 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 at $\Delta f = 1/2N$, so that we will typically choose $M = 12N$. This choice allows for about 2 zero weights between two sources that are close in azimuths but nevertheless potentially separable and thus guarantees that two disjoint clusters will be obtained in case the two sources are detected.

It remains to explain how to choose L the number of beams formed to build the observation vector $\hat{Y}$. This vector must contain all the information present in the data. Just as $\hat{R}$, it should be a sufficient statistic and as such L should be equal to the number of real degrees of freedom in the array manifold. We will detail this point below and in any case indicate the value taken for L in the simulations.

3.3.4 *Unbiasing the estimates*

Since even for $\hat{Y} = Y$, the optimal X will never be such that $FX = Y$, for $h > 0$, it is obvious that the presence of the term $h\|X\|_1$ in (6) (7) induces bias into the estimates. This bias while concerning quite directly the amplitude estimates, also affects sightly the azimuth estimates. Again we will not elaborate on this, but one can for instance use the estimates furnished by GMF to initialize a ML algorithm that will then converge in a few steps.

3.3.5 *Summary of the algorithm.*

Let us summarize the algorithm. For a given array, we build the F matrix of dimension $(L, M+1)$ with $M \simeq 12N$ and column vectors the contribution of sources at equispaced bearings and the noise to the vector Y, see Section 3.1 , [8]. With the T observed snapshots we build an estimate $\hat{R}$ of their covariance matrix and form $L \simeq 2N$ beams stored in the vector $\hat{Y}$. To implement the standard version (7), one further has to whiten $\hat{Y}$ and normalize the columns of the whitened F-matrix. We then solve (7) with h as defined in Section 3.3.1, to deduce the estimates from its optimum.

4. DEVELOPMENT

Let us come back to the problem we are concerned with, namely the presence of modeling errors and the ensuing performance degradation. This degradation is induced by the discrepancy between the assumed array manifold and the true one. The assumed array manifold is the set in $\Re^N$ described by the steering vector in (2) when θ varies in its domain of definition $[0, \pi]$ or $[-\pi, \pi]$. For the so-called uniform linear array (ULA) [11] with a half-wavelength separation, this vector becomes

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

where the dependency on θ is now visible and simple and, for uncorrelated sources, it is well known that this type of steering vector leads to a exact covariance matrix R (1) having a Toeplitz structure. This means that whatever the scenario, the associated exact covariance matrix belongs, just like $a(\theta)$, to a manifold which for a ULA has this simple Toeplitz structure. When the hermitian matrices of order N are seen as real vectors of dimension N^2, the manifold associated with hermitian Toeplitz matrices becomes a subspace of dimension $2N-1$, the number of real degrees of freedom in this type of matrices. The projection then consists in *averaging the diagonals* and is known as redundancy averaging [12] or sometimes spatial smoothing [13].

For uniform circular arrays (UCA), there exists similarly manifolds for the steering vector and the covariance matrix. The covariance matrix R has a more complex structure, the dimension of the manifold is much higher and is again linked to the number of components in the co-array [11].

But this is only the apparent part of the manifold which is far more complex and indeed exists for any array geometry. Projecting the estimated covariance matrix on its manifold is systematically beneficial when it is to be used in a suboptimal method such as the subspace techniques.

More importantly, while this is true in the absence of modeling errors, it becomes even more beneficial in the case of modeling errors. From (3), it follows that

$$\hat{Z}_k = \hat{A}s_k + n_k = (A+W)s_k + n_k,$$

with W and $N\times P$ matrix filled with independent $CN(0, \sigma^2 I)$ samples. The matrix in (5) then becomes, omitting the zero-mean cross-terms,

$$\hat{R} \simeq (A+W)\left(\frac{1}{T}\sum s_k s_k^*\right)(A+W)^* + \frac{1}{T}\sum n_k n_k^*,$$

where $(1/T)\sum s_k s_k^* \to S$ and $(1/T)\sum n_k n_k^* \to vI$ as the number of snapshots T increases. Averaging over the T snapshots has a beneficial effect of the spatial noise and reduces the estimation error due to this noise but is without effect on the modeling-error induced term, that are close to WSW^* in $\hat{R}$.

The beneficial effect of the projection on the manifold of modeling-error induced term is easy to perceive in the case of a ULA where the projection amounts to average the diagonals. While this will bring $(1/T)\sum n_k n_k^*$ even closer to vI, its effect on WSW^* is far more drastic since no other mitigation is performed on it. Indeed for the present additive gaussian noise model (3), it can be shown to be a first order approximation of a maximum likelihood algorithm which means that it is close to the best one can do, to take care of this type of modeling errors.

The good news for the GMF technique is that the beamformer performs the projection on the array manifold whatever the array geometry. This is established in Appendix A, where it is shown the beam-output at bearing θ: $a(\theta)^*\hat{R}a(\theta)$, see (4), remains the same if one uses $\hat{R}$ or $\hat{R}_p$, its projection on the array manifold.

This is a valuable feature of GMF, since obtaining $\hat{R}_p$ is a difficult task except for some very specific array geometries, such as ULA's. Approximate ways to perform the projection are proposed

in [14] for uniform circular arrays or in [15] for arbitrary arrays. Note that any optimal DOA estimation procedure, such as a maximum likelihood methods (ML), for instance, perform implicitly such a projection in one way or another.

5. SIMULATIONS RESULTS

To highlight the drastic benefit associated with the projection of the estimated covariance matrix $\hat{R}$ (5) on its manifold we will thus mainly consider ULA's since for them the projection is trivial. One has then to project $\hat{R}$ on the set of Toeplitz matrices and this is achieved by *averaging the diagonals*.

We essentially compare the performance in the resolution of two close sources of the basic MUSIC, the projected MUSIC, denoted P-MUSIC below, i.e. the MUSIC algorithm working on the projected covariance matrix, and the standard version of GMF described in Section 3.2. We consider first the case where there is no modeling error, the projection has already a beneficial effect on the performance of MUSIC in this case. Then, as soon as we introduce modeling errors, the basic MUSIC fails to resolve the two sources quite frequently. Even for quite moderate modeling errors MUSIC fails systematically and we thus only compare P-MUSIC and GMF. They have similar performances with a slight advantage to GMF.

We then consider a linear array with the same number of sensors and the same aperture as previously but with non uniformly spaced sensors. GMF keeps its good performances in the presence of modeling errors, because the beamformer implicitly projects the observation on the array manifold while the P-MUSIC is no longer feasible and MUSIC fails as expected even for quite moderate modeling errors.

We simulate a linear array with $N = 10$ equispaced sensors (half wavelength) and $P = 2$ equipowered sources at bearings θ_i equal to 0 and 6 degrees with respect to broadside, at SNR=0 dB (α_1=α_2=v=1 in (1)) and $T = 150$ snapshots. Both for MUSIC and GMF we assume to know P the number of sources present. For MUSIC this means that we take a signal subspace of dimension P and localize the P main peaks in the pseudo-spectrum, for GMF this means that in the optimum X of (7) we only consider the P main clusters of non-zero components, in case there are more than P. For GMF we further take $M = 12N$, $L = 2N$ and $h = \sqrt{2\ln M}$ as suggested in Section 3. All the results presented are obtained using 1000 independent realizations.

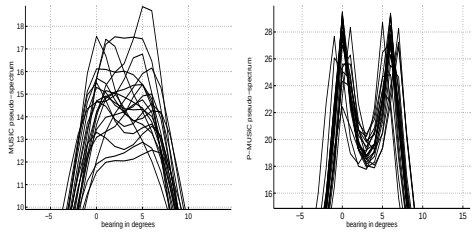
i) In the absence of modeling errors all 3 methods separate systematically the two sources. We present the results in Table 1. The observed standard deviations (st dev) are slightly below the Cramer Rao bound (CRB) for the GMF and P-MUSIC because of the granular way with which the estimates are deduced from the grid values.

Table 1: *Two equipowered sources in spatially white noise with bearing* 0 *and* 6 *degrees at SNR= 0 dB, T=150 snapshots.*

	MUSIC		P-MUSIC		GMF	
true	0	6	0	6	0	6
mean	0.19	5.80	0.01	5.98	0.01	5.99
st dev	.34	.34	.25	.25	.24	.24
CRB	.26	.26	.26	.26	.26	.26

Estimates of the mean and standard deviation of the bearings averaged over 1000 independent realizations, obtained for MUSIC, MUSIC after projection and GMF.

ii) As soon as we introduce some modeling errors, see relation (3), MUSIC fails to systematically separate the 2 sources, while P-MUSIC does so as is apparent in Figure 1., where we present the MUSIC and P-MUSIC pseudo-spectrum for 20 realizations when the modeling error in (3) has a variance equal to $\sigma^2 = .04$. Indeed while MUSIC fails to separate the 2 sources quite often even for quite small modeling-error variances, P-MUSIC just as GMF continues to resolve the two sources systematically. We present in Table 2. the results obtained for different modeling-error variances σ^2.



(a) MUSIC (b) P-MUSIC

Figure 1: *Zoom on the MUSIC pseudo-spectrum in the presence of modeling errors of variance* $\sigma^2 = 0.04$ *without (a) and with (b) projection of the estimated covariance matrix on the array manifold.*

Table 2: *Two equipowered sources in spatially white noise with bearing* 0 *and* 6 *degrees at SNR= 0 dB, T=150 snapshots and modeling-errors with variance* σ^2 *successively equal to* .02, .04 *and* .08.

	P -MUSIC		GMF	
true bearing	0	6	0	6
mean	-0.01	5.98	-0.02	5.98
std. deviation	0.47	0.48	0.48	0.49

modeling-error variance $\sigma^2 = .02$

true bearing	0	6	0	6
mean	-0.01	5.97	-0.02	6.00
std. deviation	0.62	0.65	0.66	0.66

modeling-error variance $\sigma^2 = .04$

true bearing	0	6	0	6
mean	-0.03	6.06	-0.02	6.05
std. deviation	1.26	1.40	0.94	0.95

modeling-error variance $\sigma^2 = .08$

Estimates of the mean and standard deviation of the bearings averaged over 1000 independent realizations, obtained for P-MUSIC and GMF.

iii) Keeping the two extreme sensors at their current positions, to preserve the array aperture, and moving the 8 remaining sensors slightly away from their current locations leads to a new nominal array having essentially the same performances as the initial one. In the absence of modeling error, MUSIC and GMF applied to this array will perform just as in Table 1. above. But since the projection on the new manifold is unfeasible, in the presence of modeling errors, only the use of MUSIC makes sense and it fails to resolve the two sources while GMF will have performances similar to those obtained using the ULA.

To fix ideas, we keep the two extreme sensors at their current positions, and place the other sensors regularly on an arc of circle passing by these two positions. The angle that subtends the array is taken equal to 60 degrees and the bearings are measured w.r.t. the axis of symmetry of this arc. This entirely defines the shape. The array is a part of a UCA with radius 4.5 wavelengths and inter-sensor spacing $9\sin(\pi/54)$ wavelengths. We present in Table 3. the results obtained in the presence of modeling errors by GMF since MUSIC does not work and P-MUSIC is no longer feasible. We consider the same modeling-error variances as in Table 2 and indeed the results are similar and actually slightly better which may mean that the information content is larger, though the array manifold is of much higher dimension and thus the information less concentrated.

Table 3: *Arc-of circle shaped array. Two equipowered sources in spatially white noise with bearing* 0 *and* 6 *degrees at SNR= 0 dB, T=150 snapshots and modeling-errors with variance* σ^2 *successively equal to* .02, .04 *and* .08.

σ^2	0.02		0.04		0.08	
true	0	6	0	6	0	6
mean	0.01	5.98	-0.02	6.02	-0.02	6.02
st dev	.46	.46	.64	.63	.84	.85

GMF-estimates of the mean and standard deviation of the bearings averaged over 1000 independent realizations, obtained for 3 different modeling-errors variances $\sigma^2 = 0.02,\ 0.04$ *and* 0.08 *. The performance are similar to those of GMF in Table 2., P-MUSIC is no longer feasible and MUSIC does not separate the two sources systematically.*

6. CONCLUDING REMARKS

We have addressed the DOA estimation problem in the presence of modeling errors. We have shown that the projection of the observations, e.g. the snapshot-covariance matrix, on the nominal array manifold has a beneficial effect in this context. While this projection is indeed implicit in any optimal DOA estimation scheme, it is difficult to achieve in subspace-based methods such as MUSIC. Hence the benefit of using GMF, which -when based on the beam-outputs- benefits from the implicit projection performed by the conventional beamformer. The optimal DOA-estimation strategy in the presence of modeling errors is actually not clear and further investigation are needed. It certainly depends upon the modeling-error model that is used, but even for the simplest one that is used in this paper, projecting on the nominal manifold may not be globally optimal.

7. APPENDIX A : A PROPERTY OF THE BEAMFORMER

We establish that the beamformer performs the projection on the array manifold (at no cost) or more precisely that the output of the beamformer is the same when using $\hat{R}$ or $\hat{R_p}$ the projection of $\hat{R}$ on the array manifold. Since the output of the beamformer $a(\phi)^*\hat{R}a(\phi)^*$ can be rewritten $\mathrm{tr}(\hat{R}a(\phi)a(\phi)^*)$ with $\mathrm{tr}(.)$ the trace operator, one has to establish that :

$$\mathrm{tr}(\hat{R}a(\phi)a(\phi)^*) = \mathrm{tr}(\hat{R}_p a(\phi)a(\phi)^*)$$

The proof uses the fact that $\mathrm{tr}AB^*$ is a inner product on the set of square matrices that corresponds to the usual Euclidean inner product when matrices are considered as N^2- dimensional vectors. With $\boldsymbol{P}$ the operator that achieves the desired projection, one has $\boldsymbol{P}\hat{R} = \hat{R}_p$ by definition, but also $\boldsymbol{P}R = R$, and $\boldsymbol{P}a(\phi)a(\phi)^* = a(\phi)a(\phi)^*$ since R and $a(\phi)a(\phi)^*$ already belong to the manifold. To establish the result, we successively write :

$$\begin{aligned} \mathrm{tr}(\hat{R}a(\phi)a(\phi)^*) &= \mathrm{tr}(\hat{R}\,\boldsymbol{P}a(\phi)a(\phi)^*) \\ = \mathrm{tr}(a(\phi)a(\phi)^*\boldsymbol{P}\hat{R}) &= \mathrm{tr}(\hat{R}_p a(\phi)a(\phi)^*) \end{aligned}$$

where the second equality follows from the Hermitian property of the inner product.

8. APPENDIX B : STATISTICAL PROPERTIES OF THE BEAMS

We obtain the statistical properties of the beams in order to evaluate the covariance matrix Σ of the vector $\hat{Y}$ used as input to the GMF.

The beamformer output power $\hat{y}_k$ at direction $\phi_k \in \Phi$ is the positive real quantity

$$\hat{y}_k = a(\phi_k)^*\hat{R}a(\phi_k) = \mathrm{tr}(\hat{R}a(\phi_k)\hat{R}a(\phi_k)^*)$$

$$= \mathrm{tr}(\hat{R}a(\phi_k)a(\phi_k)^*) = \mathrm{tr}(\hat{R}P_k)$$

where we have introduced $P_k = a(\phi_k)a(\phi_k)^*$. The evaluation of the covariance of $\hat{y}_k$ is then easy when one uses the following relation [9] that holds for Wishart matrices

$$E\{\mathrm{tr}(\Delta R\, A)\mathrm{tr}(\Delta R\, B)\} = \frac{1}{T}\mathrm{tr}(RARB),$$

with $\Delta R = R - \hat{R}$, the estimation error. Taking then $A = P_k$ and $B = P_\ell$ and introducing $\Delta y_k = \mathrm{tr}(\Delta R P_k)$, the estimation error on beam y_k, it follows :

$$\Sigma_{k,l} = E(\Delta y_k \Delta y_\ell) = \frac{1}{T} tr(RP_kRP_\ell)$$

which by substitution becomes

$$\frac{1}{T} tr(Ra(\phi_k)a(\phi_k)^*Ra(\phi_\ell)a(\phi_\ell)) = \frac{1}{T}|a(\phi_k)^*Ra(\phi_\ell)|^2.$$

For $k = \ell$, one gets

$$\Sigma_{k,k} = E(\Delta y_k \Delta y_k) = \frac{1}{T}|a(\phi_k)^*Ra(\phi_k)|^2 = \frac{1}{T}y_k^2.$$

In the GMF, to get an estimate of Σ the covariance matrix of $\hat{Y}$ one further replaces R by $\hat{R}$.

REFERENCES

[1] B. Friedlander, "Sensitivity analysis of the MUSIC Algorithm," *IEEE-T-ASSP*, vol. 38, 10, pp. 1740-1751, 1990.

[2] A Swindlehurst and T. Kailath, "A performance Analysis of Subspace based Methods in the Presence of Model Errors," *IEEE-T-SP*, vol. 40, 7, pp. 1758-1773, 1992.

[3] A Ferreol et al, "On the Asymptotic Performance analysis of Subspace DOA Estimation in the Presence of Modeling errors: Case of MUSIC," *IEEE-T-SP*, vol. 54, 3, pp. 907-919, 2006.

[4] A Ferreol et al, "On the Resolution Probability of MUSIC in Presence of Modeling Errors," *IEEE-T-SP*, vol. 56, 5, pp. 1945-1953, 2008.

[5] 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.

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

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

[8] 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.

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

[10] J.J. Fuchs, "On the use of the Global Matched Filter for DOA estimation in the presence of correlated waveforms," ASILOMAR, Monterey, CA. nov. 2008.

[11] D.H. Johnson and D.E. Dudgeon. *Array Signal Processing.* Prentice Hall, SP Series, 1993.

[12] D.A. Linebarger and D.H. Johnson, "The effect of spatial averaging on spatial correlation matrices in the presence of coherent signals," *IEEE Trans. on ASSP*, 38, p. 880-884, May 1990.

[13] J.E. Evans, J.R. Johnson and D.F. Sun, "High resolution angular spectrum estimation techniques," Proceedings of the *1st ASSP Workshop Spectral Estimation*, Hamilton, ON, p. 134-139, 1981.

[14] M. Wax and J. Sheinvald, "Direction Finding for Coherent Signals via Spatial Smoothing for Uniform Circular arrays," *IEEE Trans. on AP*, 42, 5, p. 613-620, May 1994.

[15] P. Forster, "Generalized Rectification of Cross Spectral Matrices for Arrays of Arbitrary Geometry," *IEEE Trans. on SP*, 49, 5, p. 972-978,2001.