Proceedings of the Institutue of Acoustics

N A LARGE ARRAY OF SENSORS

J.P. LE CADRE*, P.Y. ARQUES**

* : IRISA, CAMPUS DE BEAULIEU 35042 RENNES CEDEX FRANCE
** : GERDSM LE BRUSC 83140 SIX-FOURS-LES-PLAGES, FRANCE

SUMMARY

We present here a new method for estimating the noise model received on a large array of sensors (high number of sensors). With this aim in view, we present an algorithm which uses only the outputs of beamforming as observations. The asymptotic properties of the method are detailed, an extension to the multifrequency estimation is then proposed. Finally, simulation results confirm its interest.

The developed method proves robust and efficient. Its low computation cost should allow to include it easily into classical array processing and to improve its performances.

1. INTRODUCTION

The problem of noise correlation estimation is crucial for high resolution (H.R) methods. Especially in the underwater acoustic area where the additive noise may be highly correlated (spatially) ; even an isotropic noise can lead to strong correlations (between the sensors of the array).

For the sequel, we call additive noise the sum of all the noises received on the array, i.e. the sum of ambient noise, traffic noise, flow noise... We shall simply consider that it is constituted by the part of the received signal which is not fully spatially coherent. This definition is not without ambiguity but it is well suited to array processing.

The spatial correlations of the noise result in spatio-temporal properties (noise directivity) which are the major tools for array processing. The noise directivity is generally non constant (in spatial frequency) for a (spatially) correlated noise. The consequences of these variations may be dramatic for high resolution methods. The classical beamforming is fortunately much more robust, however the detection of weak sources can be seriously affected by the noise directivity.

In order to remedy these problems, heuristic algorithms have been developed but they are essentially local and this fact can lead to severe drawbacks. On the opposite, our method consists in estimation of a global method (i.e. valid for all the spatial frequencies). The proposed method uses as observation the beamforming outputs which are the basic quantities for all sonarists. Furthermore, it enjoys the following properties otherwise its practical interest should be null : robustness, low cost of computation, convergence ensured.

The method relies upon the definition of a functional named relative entropy functional. Maximizing that functional requires iterative methods (gradient's like). After a study of the properties of such algorithms, the derived estimates of the noise model are carefully studied. Then we present an extension of the method to multifrequency estimation of the noise model whose practical interest is evident.

Simulation results illustrate the practical interest of our method for classical array processing.

2. RELATIVE ENTROPY FUNCTIONAL (REF). DEFINITION AND PROPERTIES

Consider an array of sensors constituted by n_s equispaced sensors. Then at a given frequency (f_0), omitted for the sequel, the only available observations are the array outputs whose an exhaustive estimate is the estimated interspectral matrix $\hat{R}$. The problem consists now in separation of sources and noise part and amounts to solving the following ill posed problem :

$$\hat{R} = S + B \qquad (1)$$

(S : cross-spectral matrix of sources, B of noise).

Without supplementary hypotheses, the above problem has no meaning. At this step, it is necessary to add hypotheses which are :

- sources number majorized
- B corresponds to a short correlation hypothesis.

This last hypothesis is fundamental. Using the fact that many vectorial samples of noise are available along the array and the concept of mutual information, a functional (called Relative entropy functional) is derived and given by the formula below [1] :

$$H(B) = \text{Log}\ \det(\hat{R} - B) + L \,.\, \text{Log}\ \det B \qquad (2)$$

(det : meaning determinant of a square matrix).

In formula (2), L is a scalar factor called redundancy factor (relatively to the correlation length of the noise). The numerical problem consists now in determining the matrix B (definite positive) which maximizes the REF H.

Note that maximization is achieved only w.r.t. the parameters defining B. For instance, B may be parametrized by a spatial MA model, an AR model, etc...

Replacing $\hat{R}$ by its exact value (i.e. : R) and calling B_0 the exact noise matrix and λ_i^w the eigenvalues of the matrix $B_0^{-1}.B$ then the following fundamental property is available [1,2].

Prop. 1 : Let $\{\lambda_i^w\}$ the eigenvalues of the matrix $B_0^{-1}.B$, then if $\hat{B}$ is the matrix maximizing H (under the constraints : B and R-$\hat{B}$ definite positive), then these eigenvalues satisfy the following inequalities :

$$L/L+1 \leqslant \hat{\lambda}_i^w \leqslant 1 \qquad (\hat{\lambda}_i^w \text{ : eigenvalues of } B_0^{-1}\,.\,\hat{B})$$

The proof of this property needs elementary but rather tedious calculations. Obviously as L increases, the eigenvalues $\hat{\lambda}_i^w$ tend towards 1 and therefore $\hat{B}$ (maximizing H) tends towards B_0 (for any matricial norm).

An equivalent property will be demonstrated much more easily,. furthermore an intuitive meaning of this property will be given.

Furthermore the following property ensures that gradient method will converge [1,2].

Prop. 2 : The REF is concave w.r.t. spatial noise correlations.

Properties 1 and 2 show the practical interest of the REF method for noise model identification. The gradient vector as well as the optimal step are straightforwardly calculated, leading to an efficient and feasible method. However its cost of computation increases greatly with the sensors number and, furthermore, it dont use the beamforming outputs. Conversely in the case of a large array it will be possible to derive (from the previous one) a method avoiding these two problems.

By using spatial frequency analysis the matricial expression of the REF will be translated into another one involving only scalar formulas.

3. REF AND SPATIAL FREQUENCY ANALYSIS

First, we shall define the basic quantities used for the rest. Considering a spatio-temporal process X(t,M) indexed by time and space with spatio-temporal correlation :

$$R(t,r) = \mathbb{E}[X(t,M) \,.\, \bar{X}(t - \tau,\, M - r)]$$

its spatio-temporal density P(f,k) is defined by [3] :

$$P(f,k) = \int_{\mathbb{R}^4} R(\tau,r)\; e^{-2i\pi(f.\tau+k^*.r)}\, d\tau\, dr \quad . \tag{3}$$

The frequency f will be omitted for the sequel, we shall call R(k) and B(k) the spatio-temporal densities of the array outputs and of the additive noise. The array being assumed linear the vector k becomes a scalar equal to its first component.

The other ingredient is the theorem of Szegö [4] which will be precised now.

Consider f(x), a real function, its Fourier coefficients are defined as :

$$C_n = \frac{1}{2\pi} \,.\, \int_{-\pi}^{\pi} e^{-inx}\; f(x)\;(dx) \qquad n = 0,\; \pm 1,\; \pm 2, \ldots \tag{4}$$

then the quadratic (Toeplitz) associated forms defined by :

$$T_n(f) = \sum_{\mu,\nu=0,1,\ldots,n} C_{\nu-\mu}\; u_\mu \,.\, \bar{u}_\nu \quad . \tag{5}$$

or :

$$T_n(f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |u_0 + u_1\; e^{ix} + \ldots + u_n\; e^{inx}|^2\; f(x)dx \qquad n=0,1,2,\ldots \tag{6}$$

The eigenvalues of $T_n(f)$ are the roots of the characteristic equation $\det[T_n(f) - \lambda Id] = 0$ and denoted as $\lambda_1^n,\, \lambda_2^n, \ldots, \lambda_{n+1}^n$. Then the theorem of Szegö can be expressed as below :

Let $F(\ell)$ a continuous function on the interval [a,b], then :

$$\lim_{n \to \infty} \frac{F(\lambda_1^n) + F(\lambda_2^n) + \ldots + F(\lambda_{n+1}^n)}{n + 1} = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(f(x))\; dx \,. \tag{7}$$

Here, we shall consider that F is the logarithm (Log) and we deduce from Szegö's theorem :

$$\lim_{n \to \infty} \frac{1}{n+1} \sum_{i=1}^{n+1} \mathrm{Log}\ \lambda_i = \frac{1}{2\pi} \cdot \int_{-\pi}^{\pi} \mathrm{Log}\ f(x)\ dx \ . \tag{8}$$

If, furthermore, $f(x)$ is the function $P(f,k)$ defined by (3), then the following result is obtained :

$$\lim_{n \to \infty} \frac{1}{n+1} \mathrm{Log}\ \det(R_{f_o}) = \frac{1}{2W} \cdot \int_{-W}^{W} \mathrm{Log}\ P(f_o,k)\ dk \ . \tag{9}$$

(w : spatial bandwidth i.e $w = d/\lambda_o$, λ : wavelength at f_o).

Finally, assuming that the number of sensors is high, the REF takes the form below :

$$H = \int_{-W}^{W} \mathrm{Log}\ [R(k) - B(k)]\ dk + L \ . \int_{-W}^{W} \mathrm{Log}(B(k))dk \ . \tag{10}$$

The problem consists now in estimating the spatial density $B(k)$ which maximizes H. This problem needs itself an adequate parametrization of $B(k)$. Among all the parametrizations, a spatial MA model is quite convenient, it is classical defined by :

$$\left\{ \begin{array}{l} B(k) = \sigma^2 \cdot F(Z) \cdot F^*(Z^{-1}) \\ \text{with} \\ \\ F(Z) = 1 + b_1 \cdot Z^{-1} + \ldots + b_p \cdot Z^{-p} \\ Z = \exp(2i\pi kd), \quad d : \text{intersensor distance} \ . \end{array} \right. \tag{11}$$

Actually, the major part of physical noise (i.e generated by physical hypotheses about the spatial repartition of elementary sources of noise) may be modelled by a spatial MA noise model of a reasonable order.

Assuming that the received noise may be modelled by a MA model, then the coefficients of the MA model maximizing H satisfy the following property (see Appendix A).

Prop. 3 : Let $\hat{b}_{i,L}$ the coefficients of the MA model maximizing H defined by (3) then they satisfy the following inequalities :

$$|(b_i^o - \hat{b}_{i,L}) \cdot \frac{1}{b_i^o}| \leq 1 - \sqrt{\frac{L}{L+1}} \qquad i = 1,2,\ldots,p \ . \tag{12}$$

(b_i^o : exact value of the parameter).

Remark that this result does not take into account the estimation errors of R(k) otherwise L would chosen as great as possible. It's not a statistical result but it precises Prop. 1.

If b_i^o is positive, (12) becomes :

$$b_i^o \cdot \sqrt{\frac{L}{L+1}} \leqslant \hat{b}_{i,L} \leqslant b_i^o \quad . \tag{13}$$

This last property is illustrated by fig. 1 which represents the shape of the REF H when the noise is a MA model and in presence of one source.

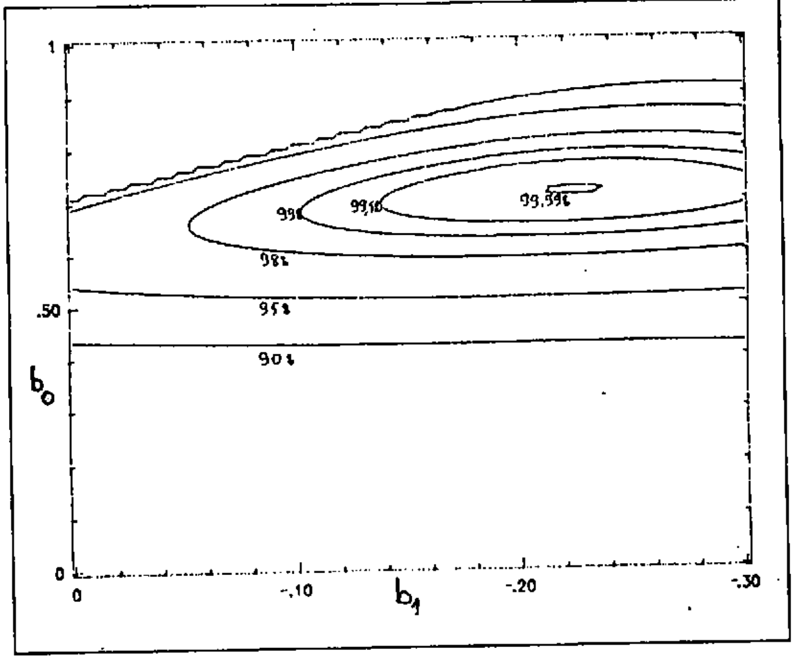


Fig. 1 : $H(b_o, b_1)$
1 source : θ : 45°, $\sigma = 0.1$
bruit MA d'ordre 2 : $b_o = 1$, $b_1 = -0.3$

We can see, on fig. 1, that H is concave and that its extremum satisfies Prop. 3 (L = 1). The estimated coefficients are respectively $\hat{b}_o = 0.71$ and $\hat{b}_1 = -0.22$; at a first glance this result seems not very good but in fact the ratio $\hat{b}_1/\hat{b}_o = -0.309$. The noise model being defined except for a multiplicative factor (for the spatial whitening), this result is quite acceptable.

The functional H being defined, we shall now consider its maximization.

4. MAXIMIZATION OF H

Practically R(k) is not available, two types of methods can then be used :

4.1 Simultaneous estimation of R(k) and B(k) :

$$R(k) = S(k) + B(k)$$

S(k) may be modelled by an AR model (with poles close to the unit circle), i.e :

$$S(k) = \frac{\sigma^2}{|A(Z)|^2} \quad . \tag{14}$$

The numerical problem consists now in estimating simultaneously the polynomial A(Z) and B(Z), leading to the following constrained problem :

$$\begin{cases} \text{Max } H(A,B) \\ \text{under the constraints :} \\ \int_{-W}^{W} R(k) \exp(2i\pi kjd) = \hat{r}(jd) \end{cases}$$

($\hat{r}(jd)$: cross-spectrum of two sensors spaced of jd). (15)

That formulation, rather similar to the classical maximum of entropy, does not lead to simple calculations. Therefore, a simpler approach is prefered.

4-2- R(k) is replaced by an estimate

Obviously, it is possible to replace R(k) (in(10)) by an estimate. It can be the discret Fourier transform defined by :

$$\hat{R}(k) = \sum_{j=-n_s+1}^{n_s-1} \hat{r}(jd).w(j).\exp(2i\pi kjd) \quad . \tag{16}$$

(n_s = number of sensors).

In formula (16), the {w(j)} represent the array weighting. A useful weighting is given by the classical triangle function i.e :

$$\hat{R}(k) = D_k^* \cdot \hat{R} \cdot D_k \tag{17}$$

(D_k : steering vector associated to k [5], $\hat{R}$: estimated CSM).

The main drawback of this weighting comes from enlarged lobes but its main advantage is to ensure positivy of the estimated $\hat{R}(k)$. For most of the practical cases, formula (17) is quite satisfying.

The gradient vector calculation is straightforward and described below :

1) Let B be the covariance matrix of a MA model, then [ ] :

$$B = \sigma^2 \left(\sum_{i=0}^{P} b_i Y_i\right) \cdot \left(\sum_{i=0}^{P} b_i \; Y_i^t\right) = \sigma^2 B_1 \cdot B_1^t \quad . \tag{18}$$

with Y_i a rectangular $(n_s x n_s + p)$ matrix defined by :

$$
Y_i = \begin{array}{c} \overset{i}{\longleftrightarrow} \\ \begin{bmatrix} 0 & & 0 & 1 & 0 & & 0 \\ & & & & & & \\ 0 & & 0 & 0 & 1 & & 0 \end{bmatrix} \end{array}
$$

2)
$$
\frac{\partial H}{\partial b_i} = \sigma^2 \int_0^{\pi} \frac{D_\theta^* [Y_i B_1^t + B_1 Y_i^t] D_\theta}{D_o^* (R-B) D_\theta} \sin\theta \, d\theta
$$

$$
- L \int_0^{\pi} \frac{D_\theta^* [Y_i B_1^t + B_1 Y_i^t] D_\theta}{D_o^* (B_1 - B_1^t) D_\theta} \sin\theta \, d\theta
$$

A convenient approximation of the optimal step of the gradient method is of fundamental importance in practice. Using the matricial formulation of H, this optimal step may be estimated as the value of the scalar ρ maximizing the expression below :

$$
H(\rho) = \sum_{i=1}^{n} \mathrm{Log}(1 + \rho\lambda_i^k) + L \, . \sum_{i=1}^{n} \mathrm{Log}(1 - \rho\mu_i^k) \quad . \tag{19}
$$

where λ_i^k and μ_i^k are respectively the eigenvalues of the matrices $(\hat{R}-B_k)^{-1}D_k$ and $B_k^{-1}D_k$ (B_k : noise matrix at the k-th iteration, D_k gradient matrix).

These eigenvalues are approximated by the samples values of the spatial densities $d(k)/\hat{r}(k)-b(k)$; $d(k)/b(k)$ (with samples defined by : $k_m d = m/m_s$, $1 \leqslant m \leqslant n_s$).

This approximated step size of the gradient is quite satisfying and ensures a fast convergence of the method under the positivity constraint $(\hat{r}(k)-b(k) > 0)$

5. REMARKS ABOUT THE REF

The property 2 shows the interest of our method for noise model estimation. The fig. 2 illustrates the behaviour of the methods which consists in seeking the more random noise model (i.e maximizing Log det B or $\int_{-W}^{W} \mathrm{Log}\, B(k)dk$) under the positivity constraint about R-B.

Actually the proposed method can be considered as a deconvolution method ; the aim being the estimation of the noise model seen by an array of known transfer function.

A parallel can be made with regularization methods of ill posed problem due to Tikhonov which consist in replacing the problem AZ = y by the problem below :

$$
\underset{Z}{\mathrm{Min}}\ f(Z) = \|AZ - y\|^2 + \lambda\|PZ\|^2 \ . \tag{20}
$$

In the above formula, λ is called the regularization parameter and determine a compromise between regularization and exact data proximity.

In the REF functional, the term Log(R-B) can be seen as a "barrier functional" which avoids sources description by noise model. The role of the term L.log(B(k)) is to take into account the uncertainty about B and overall to ensure the position of the maximum of H (see fig. 2).

Another practical problem is induced by the choice of the parametric model order of noise. It may be estimated by several ways [7]. However, the methods performs very well (Prop. 3 is still satisfied) when this order is overdetermined.

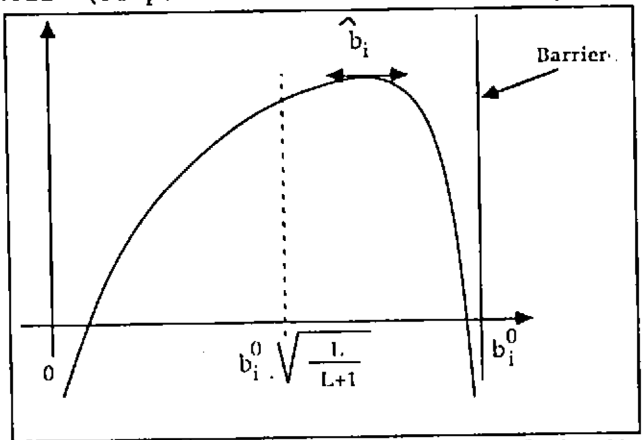


Fig. 2 : Shape of the REF functional (radially).

6. MULTIFREQUENCY ESTIMATION OF THE NOISE MODEL

We assume that the additive noise may be described by a common spatial model (with different sampling at each frequency), this hypothesis seems quite acceptable. We shall develop now a method taking advantage from this hypothesis.

Consider the frequency f_o, at that frequency assume that noise may be modelled by a MA, let be :

$$B(f_o,k) = \sigma^2 \; |1 + b_1 \cdot Z + \ldots + b_p \cdot Z^p|^2$$

$$Z = \exp(-2i\pi kd) \quad , \qquad k = \frac{f_o}{C} \cos\theta \quad (\theta \text{ : bearing}) \tag{21}$$

The wavenumber k varying from $-\frac{f_o}{C}$ to $\frac{f_o}{C}$. It can be deduced from the noise model an estimate of the cross spectrum $r_b(f_o,\ell)$, let :

$$r_b(f_o,\ell) = \int_{-1/2d}^{1/2d} B(f_o,k) \; e^{-2i\pi k\ell} \, dk \quad . \tag{22}$$

and more generally one has (unicity of the noise model)

$$r_b(f_i, nd) = \int_{-1/2d}^{1/2d} B(f_i, k)\ e^{-2i\pi k \ell}\ dk \quad . \tag{23}$$

but :

$$B(f_i, k) = \sigma^2 |1 + b_1 Z_i + \ldots + b_p\ Z_i^p|^2$$

$$Z_i = \exp(-2i\pi f_i / C\ d \cos\theta)$$

Now can also be written as :

$$Z_i = \exp(-2i\pi\ f_o / C \cos\theta\ \lambda_i) \qquad \text{with} \quad \lambda_i = d\,\frac{f_i}{f_o}$$

so that finally :

$$r_b(f_i, nd) = r_b(f_o, nd\ f_i/f_o) \quad . \tag{24}$$

Under the independancy hypothese for Fourier transforms (at different frequency), the multifrequency REF takes the following form

$$H = \sum_{f_i = f_o}^{2f_o} \left(\int_{-W}^{W} [\mathrm{Log}(\hat{R}(f_i, k) - B(f_i, k)) + L\ \mathrm{Log}\ B(f_i, k)]\ dk \right) \tag{25}$$

it is maximized by the previous gradient method.

The basic hypothesis (unique noise model) may be a little restrictive but it can be easily relaxed by use of extension models [8].

7. SPATIAL WHITENING

Assuming that a noise model have been estimated, the (numerical) problem consists now in whitening the data.

For that purpose, the classical method consists in Choleski factorization and matricial inversion of the triangular factors. For large arrays, this method is drastically costly and can present furthermore some drawbacks.

In order to remedy these problems the following approach seems much more preferable :

1) determine an AR filter equivalent to the MA model
2) the whitening filter is then $[A^{-1}(Z^{-1})]^{-1} = A(Z^{-1})$
3) compute the whitened spatial covariances :

$r_w(\ell) = A\ R_\ell A^*$ with :

$$A = \begin{pmatrix} 1 & a_1 \ldots . a_q & & \\ 0 & & & 0 \\ & & & \\ 0 & 0\ 1 & a_1 & a_q \end{pmatrix} \quad , \quad R_\ell = \begin{pmatrix} r(\ell) & r(\ell-1) \ldots\ldots . r(\ell-q) & \\ & & \\ r(\ell+q) & r(\ell+1) & r(\ell) \end{pmatrix} \quad . \tag{26}$$

Consider now the effect of that whitening on a source ; let :

$$R_w(\theta) = A\ (D_\theta\ D_\theta^*)\ A^*$$

the whitening method enjoys the following properties :

1) $\mathrm{rank}(R_w(\theta)) = 1$

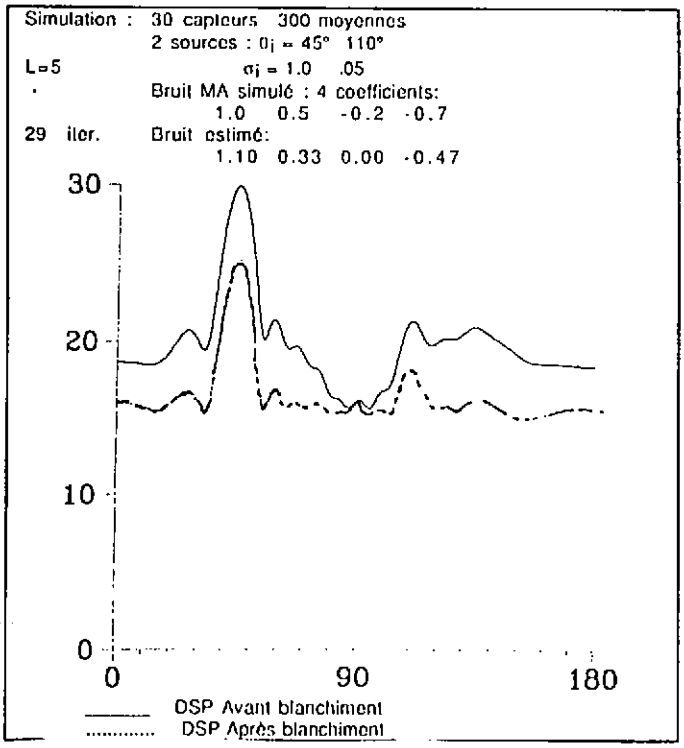
2) $A\ D_\theta = (1 + a_1\ e^{i\alpha} \ldots + a_q\ e^{iq\alpha})\ D'_\theta$. (27)

From (27), we see that whitening as defined by (26) does not modified sources properties. These formulas suggest also a fastier whitening method, i.e. compute :

$$f(\theta) = (\hat{b}^{-1}(\theta))\ .\ D_\theta^*\ .\ \hat{R}\ .\ D_\theta \qquad (28)$$

8. SIMULATION RESULTS

Fig. 3 illustrates the benefits given by the proposed method for beamforming. The quality of estimation of the noise model is quite satisfying. After whitening the weak source (100°) is clearly seen, the sidelobe at 140° completely canceled. The residual noise (after whitening) is perfectly white.



——— beamforming without whitening
-·-·- beamforming after whitening

Fig. 3 : Beamforming after whitening

For this simulation the gradient method converges in about 10 iterations.

Fig. 4 illustrates the interest of multifrequency estimation, especially in presence of a great number of sources [9] the multifrequency estimation enhances considerably the quality of the noise model estimation.

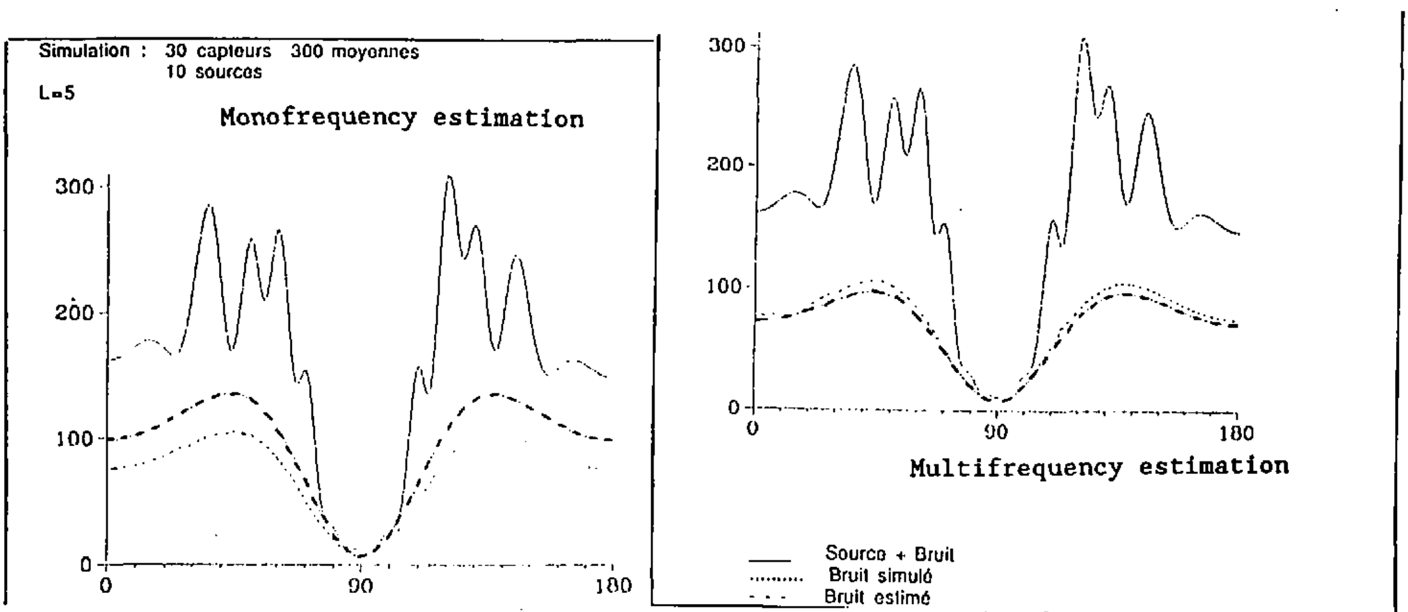


Fig. 4 Multifrequency estimation

9. CONCLUSION

A novel method for noise model estimation in presence of sources has been developed. It uses beamforming outputs as data and is therefore well suited to large arrys. Its computation cost does not depend on the sensor number.

Furthermore, its robustness w.r.t. physical hypotheses has been proved both by theoretical considerations and simulations. A simple extension to multifrequency analysis has been presented.

The proposed method can be directly applied to beamforming outputs and would enhance sonar performances with a low computation cost.

REFERENCES

[1] J.P. LE CADRE, 'Contribution à l'utilisation de méthodes paramétriques en traitement d'antenne'. Thèse de Doctorat es Sciences, INPG-USMG, Grenoble, Octobre 1987.
[2] J.P. LE CADRE et P. RAVAZZOLA, 'Practical use of parametric methods for array processing'. Lectures Notes in Control and Inforamtion Sciences, M. Bouvet and G. Bienvenu editors, Springer-Verlag, 1990.
[3] T.E. BARNARD, 'Two maximum entropy beamforming algorithms for equally spaced line arrays'. IEEE Trans. on ASSP, Vol. ASSP-30, n° 2, pp. 175-188, April 1982.
[4] U. GRENANDER and G. SZEGO, 'Toeplitz forms and their applications'. University of California Press, San Francisco, 1958.
[5] W.S. BURDIC, 'Underwater acoustic system analysis'. Prentice Hall, 1984.
[6] B. FRIEDLANDER, 'On the computation of the Cramer-Rao Bound for ARMA parameter estimation'. IEEE Trans. on ASSP, Vol. ASSP-32, No 4, pp. 721-727, August 1984.
[7] J.J. FUCHS, 'Estimation du nombre de sinusoïdes dans du bruit coloré'. Actes du 11ème Colloque GRETSI, Nice, pp. 197-200, Mai 1987.
[8] Y. GRENIER, 'Model expansion applied to speech analysis and source location'. Proc. of the ICASSP, Glasgow, pp. 2302-2305, Mai 1989.

APPENDIX A

The aim of this appendix is to prove the basic Property 3 of the REF ; with the notations of (12) one obtains :

$$\frac{\partial H}{\partial b_k} = - \int_{-W}^{W} \frac{\mathrm{Re}\ (Z^i\ \bar{F}(k))\ [LR(k) - (L+1)\ B(k)]}{[R(k) - B(k)]\ B(k)}\, dk \tag{A1}$$

If no source is present, a direct consequence of (A1) is that partial derivatives $\partial/\partial b_i H$ are nulls when $\hat{b}_i$ is equal to $b_i^o \sqrt{\frac{L}{L+1}}$ (b_i^o exact value of the parameter). H being a concave functional w.r.t the $\{b_i\}$ one deduces that :

$$\hat{B}_{i,L} = b_i^o \sqrt{\frac{L}{L+1}} \qquad \text{for } i = 1,\dots,P\ . \tag{A2}$$

Things are less clear in the presence of sources. Consider now the weighted sum of partial derivatives, i.e. :

$$\sum_{i=1}^{P} b_i \ \frac{\partial H}{\partial b_i}$$

using (A1) one obtains straightwardly :

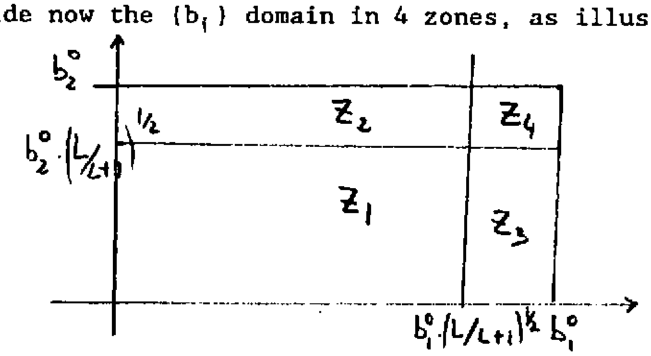
$$\sum_{b_i}^{P} b_i \ \frac{\partial H}{\partial b_i} = L \underbrace{\int_{-W}^{W} \frac{S(k)}{R(k) - B(k)}\, dk}_{(1)} + L \underbrace{\int_{-W}^{W} \frac{B_o(k) - (L+1)/L\ B(k)}{R(k) - B(k)}\, dk}_{(2)}\ . \tag{A3}$$

(with $R(k) = S(k) + B_o(k)$; $\delta(k)$ signal density).

Let us examine now the respective signs of quantities (1) and (2) of (A3).

The term (1) is positive since S(k) and R(k) - B(k) are positive (whatever k) by assumption.

Divide now the $\{b_i\}$ domain in 4 zones, as illustrated below :



We shall first show that the maximum of H cannot be in the zone Z_1. More precisely, when the coefficients $\{b_i\}$ satisfy the following inequalities :

$$|b_i| \leqslant |b_i^o| \sqrt{\frac{L}{L+1}} \qquad \text{for } i = 1,2,\ldots P \,. \tag{A4}$$

one obtains easily that :

$$\int_{-W}^{W} (B_o(k) - \frac{L+1}{L} B(k))dk = \sum_{i=o}^{P} [(b_i^o)^2 - \frac{L+1}{L} b_i^2] \,. \tag{A5}$$

Moreover, one can assume that R(k) - B(k) is greater that a scalar α strictly positive (otherwise if $R(k)-B(k) \to 0$ on a non zero-measure set, then $H \to -\infty$), so that finally :

$$\sum_{i=1}^{P} b_i \frac{\partial H}{\partial b_i} \geqslant L \int_{-W}^{W} \frac{S(k)}{R(k) - B(k)} dk + \frac{1}{\alpha} \int_{-W}^{W} [B_o(k) - \frac{(L+1)}{L} B(k)] \, dk$$

Hence :

$$\sum_{i=1}^{P} b_i \frac{\partial H}{\partial b_i} > 0 \quad \text{on } (Z_1) \quad .$$

The only one hypothesis which has been used till is the positivity of S(k), but actually S(k) corresponds to the spatial density of sources.

For instance, assume that only one source is present :

$$S(k) = \frac{\sigma^2}{|Z - Z_o|^2} \quad ; \qquad Z_o \text{ : source pole}$$

therefore :

$$\int_{-W}^{W} \frac{S(k)}{R(k) - B(k)} dk \geqslant \frac{1}{\alpha} \frac{\sigma^2}{(1 - |Z_o|^2)}$$

If the coefficients $\{b_i\}$ belong the Z_2 or Z_3, the term (2) is not necessarily positive but is bounded. On the other hand, the term (1) is great (Z_o near the unit circle : source pole). Therefore the sum $\sum_{i=1}^{P} b_i \frac{\partial}{\partial b_i} H$ is asymptotically positive on Z_2 and Z_3.

If the coefficients $\{b_i\}$ tend towards their exact values, H tends towards $-\infty$. Finally, the maximum is in the zone Z_4 itself determined by the conditions :

$$|b_i - b_i^o| \leqslant |b_i^o| \left(1 - \sqrt{\frac{L}{L+1}}\right) \qquad \text{for } i = 1,2,\ldots,p \tag{A6}$$