

## Sur les propriétés statistiques des covariances spatiales estimées après rectification, applications <sup>(1)</sup>

---

### *On the statistical properties of the smoothed spatial covariances, applications*

---

Jean-Pierre LE CADRE et Olivier ZUGMEYER, IRISA, Campus de Beaulieu, 35042 Rennes Cedex

#### Résumé

Le but de cette correspondance est de donner une expression simple et générale de la variance d'estimation des corrélations spatiales estimées après rectification. On étudie ensuite les implications de ce calcul pour le traitement d'antenne.

#### Mots clés

Rectification, Toeplitz, fonction caractéristique, SVD, calcul de variance, bruit d'estimation.

#### Summary

*This correspondence deals with the calculation of the estimated spatial correlations after rectification. The consequences of this calculation, for array processing, are then studied.*

#### Key words

*Rectification, Toeplitz, Shift, Matrix, characteristic function, SVD, variance calculation, estimation noise.*

## 1. Introduction

Une large part des études récentes en traitement d'antenne et en analyse spectrale concerne l'étude des propriétés algébriques et statistiques de matrices de covariance. Dans le cadre du traitement d'antenne, l'hypothèse de stationnarité spatiale est habituellement admise. Cette hypothèse est utilisée, dans le cas d'une antenne linéaire, pour « stationnariser » spatialement les covariances ou interspectres estimés. Cette stationnarisation spatiale est désignée par divers vocables plus ou moins barbares : rectification [1], Toeplerization, Toeplitzification, ... Les deux derniers vocables font référence au mathématicien O. Toeplitz, ce qui est louable, tandis que le premier (i.e. rectification) a l'avantage de la simplicité. Nous adopterons donc pour la suite ce vocable qui provient de

l'opération réellement effectuée pour stationnariser spatialement le signal.

Le procédé de rectification consiste simplement à moyenner les corrélations spatiales le long des diagonales. On montre aisément [2] que ceci revient à accomplir (très économiquement) une projection orthogonale de la matrice de covariance estimée sur le sous-espace des matrices de Toeplitz. Il est certes tout à fait possible de perdre au passage le caractère positif de la matrice de covariance estimée, mais les propriétés statistiques (variances essentiellement) des estimateurs de paramètres des sources sont, en général, largement améliorées.

Ceci est particulièrement important pour les méthodes à haute résolution, ainsi que le confirment les études théoriques [3] et les expérimentations [4].

La donnée matricielle des covariances intercepteurs est transformée par ce procédé en un simple vecteur constitué des éléments de la première ligne de la matrice rectifiée.

<sup>(1)</sup> Ce travail a été réalisé dans le cadre de la convention de recherche IRISA-GERDSM (DCN).

Étant donnée le peu d'information dont on dispose sur les éléments propres des matrices de Toeplitz [5], [6], on peut légitimement se demander si les méthodes d'analyse en composantes principales de cette matrice rectifiée constituent l'unique ou la meilleure voie d'analyse... Le débat à ce sujet est loin d'être achevé. On peut toutefois, beaucoup plus simplement, essayer de calculer la variance d'estimation de ces corrélations spatiales rectifiées. En dépit de sa simplicité, la résolution de ce problème peut amener à considérer des questions fondamentales telles que celle-ci : existe-t-il des choix optimaux des corrélations spatiales de l'antenne et ceci pour divers critères (variance des estimateurs, détection, etc.). La réponse à ces diverses questions peut amener à modifier la vision que l'on a des divers traitements d'antenne et à se poser une question d'une importance pratique considérable : est-il possible à partir des observations (sorties de capteurs) d'optimiser la configuration spatio-temporelle utilisée.

Le but de cette correspondance est de donner une expression simple de la variance d'estimation des corrélations spatiales estimées après rectification. Pour ce faire, on montre qu'un formalisme matriciel élémentaire allège considérablement les calculs et conduit à des formules d'une grande simplicité.

Cette correspondance est présentée comme suit ; dans la section 2 les calculs de la variance sont présentés, ils sont illustrés par les résultats de la section 3 ; enfin, la section 4 présente les conséquences pratiques des calculs effectués dans la section 2 pour des cas simples. On pourra constater dans les sections 4 et 5 que les conséquences de la section 2 sont parfois surprenantes.

## 2. Calcul des covariances des corrélations spatiales estimées après rectification

On dispose d'une antenne linéaire constituée de  $p$  capteurs équirépartis, on considère un traitement bande étroite à une fréquence de référence omise pour la suite.

On dispose alors d'une matrice de covariance (plus exactement interspectrale estimée) des sorties de capteurs, soit  $\hat{R}$  cette matrice et  $\hat{R}_T$  la matrice rectifiée.

Par abus de notation, on appellera corrélations spatiales estimées rectifiées (notées  $\hat{r}_i$ ,  $0 \leq i \leq p-1$ ) les éléments de la première ligne de  $\hat{R}_T$ . On suppose  $\hat{R}$  hermitienne par construction.

On appelle  $Z$  la matrice carrée ( $p, p$ ) de décalage (shift) définie par la formule usuelle [7] :

$$Z = \begin{pmatrix} 0 & \dots & \dots & \dots & 0 \\ 1 & & & & \vdots \\ 0 & & & & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}. \quad (1)$$

La diagonale de 1 est donc décalée une fois par rapport à la diagonale principale. On appellera aussi  $Z^i$  la matrice carrée ( $p, p$ ) dont la diagonale constituée de 1 est décalée

$i$  fois relativement à la diagonale. Cette notation a un sens car on montre très aisément les résultats ci-dessous :

$$Z \cdot Z = Z^2$$

et plus généralement :

$$Z^i \cdot Z^j = Z^{i+j} \quad (i \text{ et } j \geq 0). \quad (2)$$

Par cohérence avec (2), on convient que :

$$Z^0 = \text{Id} \quad (3)$$

(Id : matrice identité ( $p, p$ )).

Les éléments  $\hat{r}_i$  de la première ligne de  $\hat{R}_T$  s'expriment très simplement avec l'aide du formalisme ci-dessus, i.e. :

$$\begin{cases} \hat{r}_0 &= \frac{1}{p} \text{tr} (\hat{R} \cdot Z^0) \\ \hat{r}_1 &= \frac{1}{p-1} \text{tr} (\hat{R} \cdot Z) \\ \vdots & \\ \hat{r}_i &= \frac{1}{p-i} \text{tr} (\hat{R} \cdot Z^i) \\ \vdots & \\ \hat{r}_{p-1} &= 1 \cdot \text{tr} (\hat{R} \cdot Z^{p-1}) \end{cases} \quad (4)$$

(tr : désigne ici la trace d'une matrice [8]).

Nous allons maintenant utiliser l'expression (4) pour calculer la covariance centrée de  $\hat{r}_i$  et  $\hat{r}_j$ . On suppose, de plus, que  $\hat{R}$  a été obtenue par la méthode du périodogramme moyenné, i.e. :

$$\hat{R} = \frac{1}{N} \sum_{k=1}^N X_k \cdot X_k^* \quad (5)$$

( $X_k$  : vecteurs de sorties de T.F. des capteurs de l'antenne. \* : signifiant transposée conjuguée).

Par conséquent, en utilisant (4) et (5), on obtient :

$$\begin{aligned} (p-i)(p-j) \cdot \mathbb{E}(\hat{r}_i \cdot \hat{r}_j^*) &= \mathbb{E}[\text{tr} (\hat{R} \cdot Z^i) \cdot \text{tr}^* (\hat{R} \cdot Z^j)] \\ &= \frac{1}{N^2} \left[ \sum_{k=1}^N \mathbb{E} \{ (X_k^* \cdot Z^i \cdot X_k)(X_k^* \cdot Z^j \cdot X_k) \} \right. \\ &\quad \left. + \sum_{k \neq k'} \mathbb{E} \{ (X_k^* \cdot Z^i \cdot X_k)(X_{k'}^* \cdot Z^j \cdot X_{k'}) \} \right]. \quad (6) \end{aligned}$$

Il reste à calculer les termes entre crochets de (6), ce qui est accompli directement par l'utilisation de la fonction caractéristique [9] de la forme quadratique  $X^* \cdot A \cdot X$ , définie relativement au vecteur  $X$  gaussien complexe circulaire (sorties de T.F.), i.e. [10] :

$$\begin{aligned} \Phi(t) \triangleq \mathbb{E}[\exp(itX^* \cdot A \cdot X)] &= \det^{-1} (\text{Id} - itA \cdot R) \\ (R = \text{cov}(X)) & \end{aligned} \quad (7)$$

d'où ; par dérivation de  $\Phi(t)$  on obtient aisément (cf. annexe 1) :

$$\begin{aligned} \mathbb{E}[(X_k^* \cdot Z^i \cdot X_k)(X_{k'}^* \cdot Z^j \cdot X_{k'})] &= \\ &= \text{tr} (Z^i \cdot R) \cdot \text{tr} (Z^j \cdot R) + \text{tr} (Z^i \cdot R \cdot Z^j \cdot R) \end{aligned}$$

et

$$\begin{aligned} \mathbb{E}[(X_k^* \cdot Z^i \cdot X_k)(X_{k'}^* \cdot Z^{j'} \cdot X_{k'})] &= \\ &= \text{tr}(Z^i \cdot R) \cdot \text{tr}(Z^{j'} \cdot R) \end{aligned} \quad (8)$$

( $R \triangleq \text{cov}(X_k)$ ),

$1 \leq k \leq N$ ,  $X_k$  et  $X_{k'}$  décorrélés si  $k \neq k'$ ).

Notons  $\tilde{r}_i$  la variable centrée :

$$\tilde{r}_i \triangleq \hat{r}_i - \mathbb{E}(\hat{r}_i).$$

Alors, on déduit directement de (6) et (8) l'expression ci-dessous :

$$\mathbb{E}(\tilde{r}_i \cdot \tilde{r}_j^*) = \frac{1}{(p-i)(p-j) \cdot N} \text{tr}(Z^i \cdot R \cdot Z^{j'} \cdot R). \quad (9)$$

On peut trouver à l'expression (9) une formulation quelque peu géométrique en remarquant que la matrice  $Z^i \cdot R$  est obtenue par sélection des  $(p-i)$  premières lignes de  $R$  (« translattées »  $i$ -fois vers le bas) et remplacement des «  $i$  » — premières par des lignes constituées de 0 ; on notera  $R^{\downarrow i}$  la matrice  $Z^i \cdot R$  et  $R^{\uparrow j}$  la matrice  $Z^{j'} \cdot R$ . Soit finalement :

$$\begin{aligned} \mathbb{E}(\tilde{r}_i \cdot \tilde{r}_j^*) &= \frac{1}{(p-i)(p-j) \cdot N} \text{tr}(R^{\downarrow i} \cdot R^{\uparrow j}) \\ 0 \leq i, j &\leq p-1. \end{aligned} \quad (10)$$

La signification des symboles  $\downarrow i$  et  $\uparrow j$  se comprend donc aisément. En dépit de sa simplicité, cette formule constitue le résultat de base de cet article. Nous allons cependant montrer que les conséquences de celle-ci ne sont pas toutes triviales.

### 3. Résultats de simulations

D'après (10), la variance d'estimation du coefficient  $\hat{r}_i$  prend la forme ci-dessous :

$$\begin{aligned} \text{var}(\hat{r}_i) &= \mathbb{E}(|\tilde{r}_i|^2) = \frac{1}{(p-i)^2 \cdot N} \text{tr}(R^{\downarrow i} \cdot R^{\uparrow i}) \\ 0 \leq i &\leq p-1. \end{aligned} \quad (11)$$

On présente sur la figure 1 l'allure de  $\text{var}(\hat{r}_i)$  donnée par (11) pour une configuration assez classique. L'indice spatial des corrélations porté en abscisse est l'indice de la corrélation  $\hat{r}_i$  définie par (4). Pour une antenne constituée de 100 capteurs cet indice varie de 0 (autocorrélation) à 99. La formule théorique (10), pour les données considérées (i.e. nombre de capteurs, moyennages, rapports S/B, ...) est comparée avec l'écart type estimé sur 200 réalisations indépendantes de  $\hat{R}$  ( $N = 1$ ).

Les deux sources étant de niveaux assez faibles, le bruit est prédominant. On peut remarquer le bon accord entre les deux courbes ce qui valide (11).

On peut remarquer sur la figure 1 que cette variance

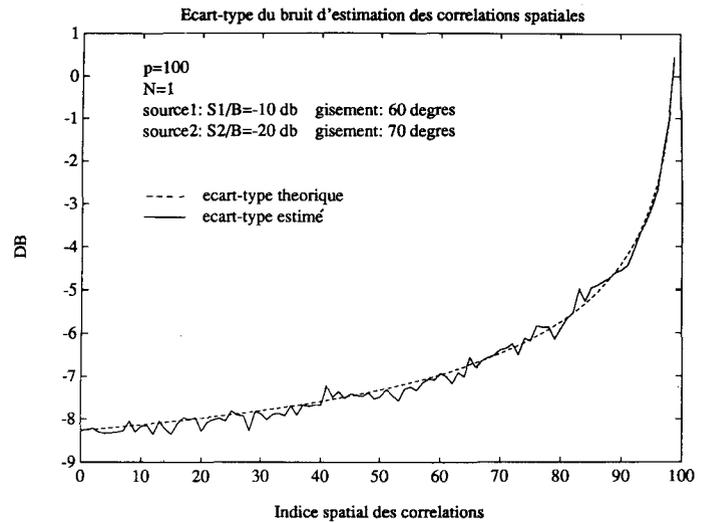


Figure 1. — Comparaison entre l'écart type théorique du bruit d'estimation de  $\hat{r}_i$  ( $0 \leq i \leq 99$ ) et son écart type expérimental ( $\sigma_1^2 = -10$  dB,  $\sigma_2^2 = -20$  dB,  $p = 100$ ,  $N = 1$ ) estimée sur 200 réalisations de  $\hat{R}$ .

augmente avec  $i$  ce qui s'explique aisément par le fait que le dernier  $\hat{r}_i$  (i.e.  $\hat{r}_{p-1}$ ) ne bénéficie d'aucun moyennage spatial alors que  $\hat{r}_0$  bénéficie de  $p$  moyennages spatiaux. Est-ce à dire que l'estimation des  $r_i$  ( $i$  fixé) bénéficie systématiquement des moyennages spatiaux, c'est à cette question que tente de répondre la section 4.

### 4. De l'influence des moyennes spatiales

Supposons que le signal soit composé d'une seule source de puissance normalisée à 1, la matrice  $R$  est alors de rang 1 et on peut supposer sans perte de généralité que tous les éléments de  $R$  sont égaux à 1, car :

$$\begin{aligned} R &= \Delta_0 \cdot R_0 \cdot \Delta_0^*, \quad \Delta_0 = \text{diag}(1, e^{i\alpha}, \dots, e^{i(p-1)\alpha}), \\ R_0(i, j) &= 1 \quad \forall (i, j) \end{aligned}$$

on obtient alors d'après (9) et (11) et des calculs élémentaires, en négligeant les termes de déphasages de  $\Delta_0$  (i.e.  $R \triangleq R_0$ ) :

$$\begin{aligned} \mathbb{E}(|\tilde{r}_i|^2) &= \frac{1}{N} \quad \text{et} \quad \text{cov}(\tilde{r}_i, \tilde{r}_j) = \frac{1}{N} \\ 0 \leq i &\leq p-1. \end{aligned} \quad (12)$$

Les moyennages spatiaux ne sont donc d'aucune utilité, ce qui est évident si on remarque que la matrice  $R$  est de Toeplitz par construction.

Considérons maintenant le cas où le signal est seulement composé d'un bruit blanc spatialement (i.e.  $R = \text{Id}$ ), alors d'après (11) :

$$\begin{cases} \mathbb{E}(|\tilde{r}_i|^2) = \frac{1}{N \cdot (p-i)} \\ \mathbb{E}(\tilde{r}_i \cdot \tilde{r}_j^*) = 0. \end{cases} \quad (13)$$

C'est donc dans ce cas que la rectification s'avère la plus bénéfique pour la réduction de la variance d'estimation des  $\tilde{r}_i$ . Cependant, l'utilité pratique en est limitée car s'il n'y a que du bruit...

Supposons maintenant que le signal soit la somme d'un bruit spatialement blanc et d'une source unique (i.e.  $R = \rho S + Id$ ,  $S = \Delta_\theta R_0 \Delta_\theta^*$  et  $\rho =$  rapport signal à bruit), alors toujours d'après (11) :

$$\mathbb{E}(|\tilde{r}_i|^2) = \frac{1}{N} \cdot \left( \frac{1}{p-i} + \rho^2 + 2\rho \cdot \frac{1}{p-i} \right). \quad (14)$$

On voit donc d'après (14) que si  $p$  devient grand devant  $i$  :

$$\mathbb{E}(|\tilde{r}_i|^2) \approx \frac{1}{N} \cdot \rho^2. \quad (15)$$

Il est aussi possible d'obtenir une expression développée de la corrélation et de la variance du bruit d'estimation sans trop de difficulté en considérant le cas de deux sources plus du bruit grâce à la formule (11).

La matrice interspectrale asymptotique s'écrit :

$$R = \rho_1 D_{k_1} D_{k_1}^* + \rho_2 D_{k_2} D_{k_2}^* + Id$$

où  $\rho_j$  et  $k_j$  sont respectivement le rapport signal à bruit et la fréquence spatiale de la  $j$ -ième source.

On obtient alors ( $N = 1$ ) :

$$\begin{aligned} \mathbb{E}[\tilde{r}_{m_1} \tilde{r}_{m_2}^*] &= \rho_1^2 e^{-i2\pi d(m_2 - m_1)k_1} + \\ &+ \rho_2^2 e^{-i2\pi d(m_2 - m_1)k_2} + \frac{1}{(p - m_1)} \delta_{m_1, m_2} \\ &+ \frac{2(p - \sup(m_1, m_2))}{(p - m_1)(p - m_2)} \\ &\times (\rho_1 \cos(2\pi d(m_1 - m_2)k_1) + \\ &+ \rho_2 \cos(2\pi d(m_1 - m_2)k_2)) \\ &+ 2\rho_1 \rho_2 e^{i\pi d(m_1 - m_2)(k_1 + k_2)} \\ &\times \frac{\sin(\pi d(p - m_2)(k_2 - k_1)) \sin(\pi d(p - m_1)(k_2 - k_1))}{(p - m_1)(p - m_2) \sin^2(\pi d(k_2 - k_1))} \end{aligned} \quad (16)$$

avec :

$$\begin{cases} k_j = \frac{\cos \theta_j}{\lambda} \quad (\lambda : \text{longueur d'onde}, \theta_j : \text{gisement}) \\ d : \text{distance intercepteur} \\ \delta_{m_1, m_2} : \text{symbole de Kronecker} \end{cases}$$

$$\begin{aligned} \mathbb{E}[|\tilde{r}_m|^2] &= \rho_1^2 + \rho_2^2 + \frac{(2\rho_1 + 2\rho_2 + 1)}{(p - m)} + \\ &+ 2\rho_1 \rho_2 \frac{\sin^2(\pi d(p - m)(k_1 - k_2))}{(p - m)^2 \sin^2(\pi d(k_1 - k_2))}. \end{aligned} \quad (17)$$

On remarque que dans le cas le plus favorable ( $p \rightarrow \infty$  et  $m_1$  et  $m_2$  fixés) :

$$\begin{aligned} \lim_{p \rightarrow \infty} \mathbb{E}[\tilde{r}_{m_1} \tilde{r}_{m_2}^*] &= \rho_1^2 e^{-i2\pi d(m_2 - m_1)k_1} + \rho_2^2 e^{-i2\pi d(m_2 - m_1)k_2} \\ \lim_{p \rightarrow \infty} \mathbb{E}[|\tilde{r}_m|^2] &= \rho_1^2 + \rho_2^2. \end{aligned} \quad (18)$$

En notant que :

$$r_m = \mathbb{E}[\hat{r}_m] = \rho_1 e^{i2\pi d m k_1} + \rho_2 e^{i2\pi d m k_2} + \delta_m$$

la variance du bruit d'estimation  $\tilde{r}_m$  est minorée par la puissance moyenne de  $r_m$  (sinusoïdes spatiales).

En rappelant que :  $\hat{r}_m = r_m + \tilde{r}_m$  on définit les rapports signal à bruit des processus constituants  $\hat{r}_m$  (en se plaçant dans le cas où la variance de  $\tilde{r}_m$  est définie par (18)) :

$$\frac{S_1}{B} = \frac{\rho_1}{\sqrt{\rho_1^2 + \rho_2^2}}; \quad \frac{S_2}{B} = \frac{\rho_2}{\sqrt{\rho_1^2 + \rho_2^2}}.$$

Si  $\rho_1^2 = \rho_2^2$  (sinusoïdes de même puissance moyenne) les rapports signal à bruit  $\frac{S_2}{B}$  et  $\frac{S_1}{B}$  sont identiques et cela constitue le cas le plus favorable pour estimer le paramètre de chaque source (les fréquences spatiales  $k_1$  et  $k_2$  par exemple).

On remarquera que dès lors, que les quantités  $\rho_1^2$  et  $\rho_2^2$  deviennent différentes, ( $\rho_1^2 \gg \rho_2^2$  ou  $\rho_2^2 \gg \rho_1^2$ ), l'une des deux sources devient immanquablement pénalisée, du fait que la variance du bruit d'estimation prend en compte la totalité des puissances des sources en présence.

Il paraît donc difficile, même dans ce cas favorable, ( $p \rightarrow \infty$ ), ou après intégration temporelle ( $N \neq 1$ ), d'espérer détecter une source de puissance très faible par rapport aux autres (sans recourir au gain spatial de l'antenne).

En effet pour  $s$  sources, la  $j$ -ième, a le rapport signal à bruit :

$$S_j/B = \frac{\rho_j}{\sqrt{\sum_{p=1}^s \rho_p^2}}.$$

## 5. Application au traitement d'antenne

Considérons l'estimation de paramètres caractéristiques de sources ; soit  $\theta \triangleq (\theta_1, \theta_2, \dots, \theta_m)^t$  le vecteur de paramètres correspondant à une source, et :

$$\hat{\theta}_k = f_k(\{X_t\}_{t=1, \dots, N}) \quad 1 \leq k \leq m \quad (19)$$

( $\{X_t\}_{t=1, \dots, N}$ ) est l'observation.

La fonction  $f_k$  est de plus, supposée deux fois différentiable. Alors en suivant [12] et [13], on obtient le développe-

ment suivant de  $\hat{\theta}_k$  dans le cas où  $X = S + B$  avec  $B : \mathcal{N}(0, b^2 \cdot \text{Id})$

$$\hat{\theta}_k = f_k(S) + \left( \frac{\partial f_k(X)}{\partial X} \right)_{X=S} \cdot B + \frac{B^*}{2} \left( \frac{\partial^2 f_k(X)}{\partial^2 X} \right)_{X=S} B + \dots \quad (20)$$

avec  $f_k(S) = \theta_k$ .

Pour un fort rapport signal à bruit, on peut approximer la moyenne et la variance de  $\hat{\theta}_k$  par les formules ci-dessous [13]

$$\mathbb{E}[\hat{\theta}_k] \cong \theta_k + \frac{1}{2} b^2 \text{tr} \left\{ \left( \frac{\partial^2 f_k(X)}{\partial^2 X} \right)_{X=S} \right\}. \quad (21)$$

Si l'on suppose l'estimée de  $\hat{\theta}_k$  non biaisée ( $\mathbb{E}(\hat{\theta}_k) = \theta_k$ ) :

$$\text{var}(\hat{\theta}_k) = b^2 \left\| \left( \frac{\partial f_k}{\partial X} \right)_{X=S} \right\|^2. \quad (22)$$

Dans ces formules  $b^2$  désigne la variance du bruit d'estimation de l'observation utilisée. Cette observation est, elle-même déduite de l'observation  $\{X_t\}$  par un résumé non nécessairement exhaustif.

Ce principe général pour le calcul de la variance des paramètres est utilisé par de nombreux auteurs [12], [13]. Le calcul du membre de droite de (21) est généralement effectué au moyen de méthodes de calcul différentiel [12], [13].

Spécialisons un peu notre propos au cas où l'observation est constituée des  $\{\hat{r}_m(t)\}_{t=1, \dots, N}$ . L'indice spatial  $m$  est donc fixé pour la suite. On s'intéresse à l'estimation des paramètres d'évolutions temporelles des fréquences spatiales des sources. On considère un modèle linéaire de l'évolution temporelle des fréquences spatiales des sources, i.e. :

$$k_j(t) = k_j(0) + \dot{k}_j \cdot t. \quad (23)$$

Dans la formule (23)  $k_j(t)$  représente la fréquence spatiale de la source  $j$  à l'instant  $t$  et le paramètre  $\dot{k}_j$  la vitesse d'évolution de celle-ci. L'évolution temporelle des fréquences spatiales de la source  $j$  est donc caractérisée par les deux paramètres  $k_j(0)$  et  $\dot{k}_j$ . Ce modèle linéaire a été étudié dans [14], [15] ; il n'est, évidemment, valide que sur un intervalle de temps limité mais est suffisant pour de nombreuses applications pratiques [14], [15].

Un modèle simple des observations  $\{\hat{r}_m(t)\}$  correspondant à ce modèle de sources en mouvement est le modèle linéaire ci-dessous [15] :

$$\begin{cases} X(t+1, m) = F_1^m \cdot X(t, m) \\ \hat{r}_m(t) = h^* \cdot X(t, m) + w_m(t) \quad (m \neq 0) \end{cases}$$

avec

$$F_1 = \text{diag}(e^{i2\pi \dot{k}_1}, \dots, e^{i2\pi \dot{k}_s}), \quad h^* = (1, \dots, 1) \quad (24)$$

$w_m(t)$  : bruit blanc de variance  $b_m^2$ .

Le problème considéré ici étant l'estimation des paramètres  $\dot{k}_j$ , une méthode classique [14], [16] consiste à considérer la matrice de Hankel  $\widehat{\mathcal{H}}_m \triangleq \text{Hank}(\hat{r}_m(t))_{t=1, \dots, N}$ .

On prouve aisément la factorisation suivante de la matrice de Hankel exacte  $\mathcal{H}_m$ , i.e. :

$$\mathcal{H}_m = \begin{pmatrix} h^* \\ h^* F_1^m \\ h^* F_1^{2m} \\ \vdots \end{pmatrix} (X(1, m) X(2, m) \dots)$$

soit encore :

$$\mathcal{H}_m = \theta_m \cdot \chi_m. \quad (25)$$

La donnée de la matrice  $\theta_m$  permet d'estimer à son tour  $F_1^m$  (et donc les  $\dot{k}_j$ ) par l'intermédiaire de la classique propriété de shift, i.e. :

$$\theta_m^\dagger \cdot F_1^m = \theta_m^\dagger. \quad (26)$$

On peut noter, d'après (25), que le rang de  $\mathcal{H}_m$  est égal au nombre de sources. Une méthode classique d'estimation de  $F_1^m$  consiste en les étapes ci-dessous :

1. Former  $\widehat{\mathcal{H}}_m \triangleq \text{Hank}(\hat{r}_m(t))$ .
2. Effectuer la SVD de  $\widehat{\mathcal{H}}_m$ .
3. Sélectionner les  $s$  plus grands vecteurs singuliers (à gauche) correspondants de  $\widehat{\mathcal{H}}_m$

$$(\widehat{\mathcal{H}}_m)_s = (U_{m,s} \cdot \Delta_s^{1/2})(\Delta_s^{1/2} V_{m,s}) \quad (27)$$

$(\widehat{\mathcal{H}}_m)_s$  : approximation de rang  $s$  de  $\widehat{\mathcal{H}}_m$ .

4. Identifier  $\hat{\theta}_{m,s}$  à la matrice  $U_{m,s} \cdot \Delta^{1/2}$  et en déduire  $\hat{F}_1^m$  par l'utilisation de la propriété de shift (26).

Cette analyse peut être effectuée sur toutes les sorties  $\{\hat{r}_m(t)\}_i$ . On appellera « fil  $m$  » la donnée des observations  $\{\hat{r}_m(t)\}_i$ . On note, de plus, que ces divers fils (indice  $m$  variant) « voient » la même matrice  $F_1$  mais à des exposants différents. Ceci est illustré par la figure 2 qui présente les sorties de divers fils.

On se pose donc la question suivante : pour la procédure d'estimation de  $F_1$  définie en (27) est-il possible (à partir des observations) de déterminer le fil minimisant la variance d'estimation des  $\dot{k}_j$  ?

L'application de la formule (22) à notre problème donne pour le fil  $m$  :

$$\text{var}(\hat{k}_j) = b_m^2 \cdot \left\| \left( \frac{\partial \hat{k}_j(\hat{h})}{\partial \hat{h}} \right)_{\hat{h}=\hat{h}} \right\|^2$$

avec :

$$\hat{h} \triangleq (\hat{r}_m(1), \dots, \hat{r}_m(N)). \quad (28)$$

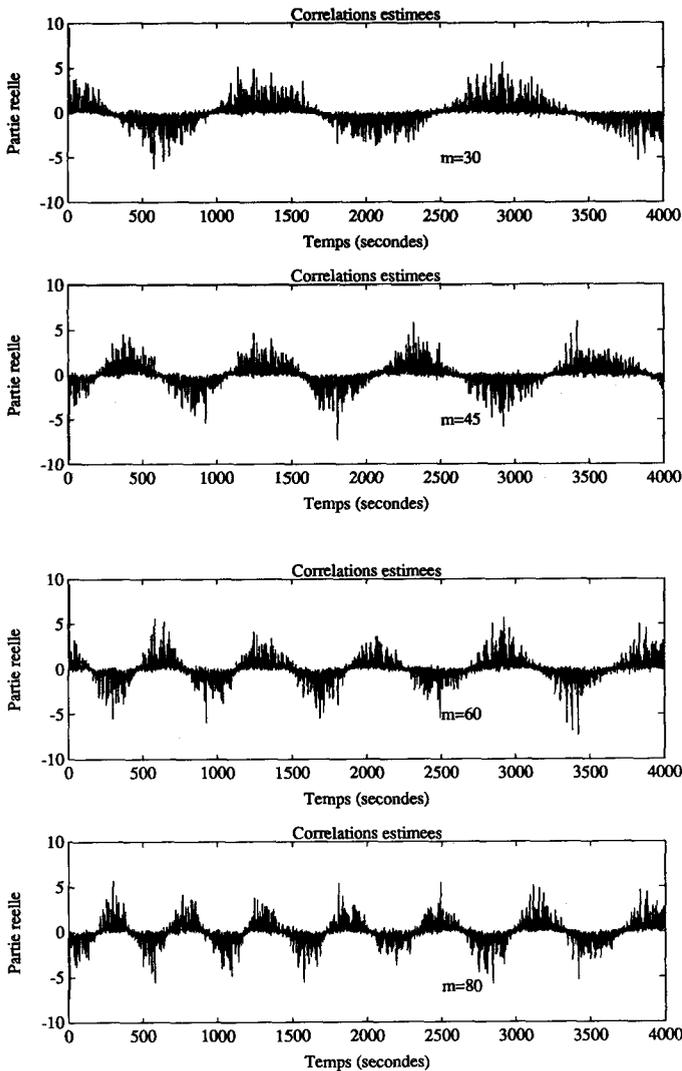


Figure 2. — Sorties des fils  $m = 30, 45, 60, 80$  pour la configuration ci-dessous :  $p = 100$ .

La variance du bruit  $b_m^2$  est quant à elle donnée par la formule fondamentale (11).

Le calcul de la variance de  $\hat{k}_j$  passe par celui des composantes du vecteur ci-dessous :

$$\left( \frac{\partial \hat{k}_j(\hat{h})}{\partial \hat{h}} \right)_{\hat{h} = h} = \left( \frac{\partial \hat{k}_j(\hat{h})}{\partial \hat{h}_1}, \dots, \frac{\partial \hat{k}_j(\hat{h})}{\partial \hat{h}_N} \right)_{\hat{h} = h}$$

avec :

$$\hat{h}_i \triangleq \hat{r}_m(i). \quad (29)$$

On rappelle que la matrice de transition estimée  $\hat{F}_1^m$  est obtenue par pseudo-inverse de la matrice d'observabilité estimée notée  $\hat{U} = \hat{U}_m$  (eq. 27), soit :

$$\hat{F}_1^m = (\hat{U}^\dagger * \hat{U}^\dagger)^{-1} \cdot \hat{U}^\dagger * \cdot U^\dagger. \quad (30)$$

Les fréquences spatiales estimées  $\hat{k}_j$  sont les valeurs propres de la matrice  $\hat{F}_1^m$  définie par (30). Le calcul des dérivées partielles définies en (29) nécessite donc deux étapes principales : l'une constituée du calcul des dérivées partielles des vecteurs propres ; l'autre des dérivées partielles des valeurs propres de  $\hat{F}_1^m$  (30), cette dernière étape étant assurée par l'utilisation de l'algorithme de Leverrier. Ces calculs étant classiques [13], [17] nous nous contenterons de donner ci-dessous les étapes du calcul des dérivées

partielles  $\frac{\partial \hat{k}_j}{\partial \hat{h}_i}$ , soit : Notons  $\hat{\lambda}_1, \dots, \hat{\lambda}_s$  les valeurs propres de  $\hat{F}_1^m$  (30), alors on a :

1.

$$\frac{\partial \hat{k}_j}{\partial \hat{h}_i} = \frac{1}{2 \pi dm |\hat{\lambda}_j|^2} \cdot \left( \text{Im}(\hat{\lambda}_j) \cdot \text{Re} \left( \frac{\partial \hat{\lambda}_j}{\partial \hat{h}_i} \right) - \text{Re}(\hat{\lambda}_j) \cdot \text{Im} \left( \frac{\partial \hat{\lambda}_j}{\partial \hat{h}_i} \right) \right).$$

2.

$$\frac{\partial \hat{\lambda}_j}{\partial \hat{h}_i} = \frac{1}{\prod_{\ell=1, \ell \neq j}^s (\hat{\lambda}_j - \hat{\lambda}_\ell)} \cdot \sum_{\ell=1}^s \hat{\lambda}_j^{s-\ell} \cdot \frac{\partial \hat{a}_\ell}{\partial \hat{h}_i}.$$

3.

$$\frac{\partial \hat{a}_\ell}{\partial \hat{h}_i} = \frac{1}{\ell} \text{tr} \left( \frac{\partial \hat{F}_1^m}{\partial \hat{h}_i} \cdot D_{\ell-1} + \hat{F}_1^m \frac{\partial D_{\ell-1}}{\partial \hat{h}_i} \right).$$

Les matrices D étant définies par les équations suivantes

$$\left. \begin{aligned} D_0 &= \text{Id} \\ D_\ell &= \hat{F}_1^m \cdot D_{\ell-1} - \hat{a}_\ell \cdot \text{Id} \\ &\vdots \\ D_{s-1} &= \hat{F}_1^m D_{s-2} - \hat{a}_{s-1} \cdot \text{Id} \\ D_s &= \hat{F}_1^m \cdot D_{s-1} - \hat{a}_s \cdot \text{Id} = 0 \end{aligned} \right\}$$

$$\frac{\partial D_\ell}{\partial \hat{h}_i} = \frac{\partial \hat{F}_1^m}{\partial \hat{h}_i} \cdot D_{\ell-1}$$

$$+ \hat{F}_1^m \frac{\partial D_{\ell-1}}{\partial \hat{h}_i} - \frac{\partial \hat{a}_\ell}{\partial \hat{h}_i} \cdot \text{Id}$$

$$\hat{a}_\ell = \frac{1}{\ell} \text{tr} (\hat{F}_1^m \cdot D_{\ell-1})$$

4.

$$\frac{\partial \hat{F}_1^m}{\partial \hat{h}_i} = (\hat{U}^\dagger * \hat{U}^\dagger)^{-1} \left( \frac{\partial \hat{U}^\dagger *}{\partial \hat{h}_i} \cdot \hat{U}^\dagger + \hat{U}^\dagger * \cdot \frac{\partial \hat{U}^\dagger}{\partial \hat{h}_i} - \frac{\partial \hat{U}^\dagger *}{\partial \hat{h}_i} \cdot \hat{U}^\dagger \hat{F}_1^m - \hat{U}^\dagger * \cdot \frac{\partial \hat{U}^\dagger}{\partial \hat{h}_i} \cdot \hat{F}_1^m \right).$$

5.

$$\frac{\partial \hat{U}_k}{\partial \hat{h}_i} = \sum_{n=1, n \neq k}^M \frac{\hat{U}_n^* \left( \frac{\partial \hat{\mathcal{H}}_m \hat{\mathcal{H}}_m^*}{\partial \hat{h}_i} \right) \cdot \hat{U}_k}{\hat{\lambda}_k - \hat{\lambda}_n} \cdot \hat{U}_n \quad (31)$$

pour  $1 \leq k \leq s$   
 où  $(\hat{\lambda}_j)_{1 \leq j \leq M}$  sont les valeurs propres de  $\hat{\mathcal{H}}_m \hat{\mathcal{H}}_m^*$  (de dimension  $M$ ,  $M$ ).

En utilisant les formules (28), (29) et (30) il est donc possible d'estimer les termes  $\text{var}(\hat{k}_j)$ . Bien que les formules exprimées en (31) soient assez rébarbatives, il est possible d'en extraire une information simple.

Remarquons, tout d'abord, que  $\text{var}(\hat{k}_j)$  est le produit de deux termes, le terme lié au bruit d'estimation  $b_m^2$  et le terme  $\left\| \frac{\partial \hat{k}_j}{\partial \hat{h}} \right\|^2$ , terme purement géométrique.

On peut alors constater d'après (10) et (31) que  $b_m^2$  est facteur de  $\frac{1}{p-m}$  (à faible  $\frac{S}{B}$ ) tandis que le terme géométrique est facteur de  $\frac{1}{m^2}$  (d'après (31)). La variance de  $\hat{k}_j$  est donc facteur du produit de deux termes « contradictoires », l'un décroissant en fonction de  $m \left( \frac{1}{m^2} \right)$  et l'autre croissant  $\left( \frac{1}{p-m}, 0 \leq m \leq p-1 \right)$ . On peut donc penser grâce à cette analyse simpliste qu'il y a des choix de  $m$  optimaux. Considérons la fonction de  $m$  définie ci-dessous :

$$f(m) = \frac{1}{(p-m) \cdot m^2} \quad 0 \leq m \leq p-1. \quad (32)$$

Son minimum est atteint pour  $m = \frac{2p}{3}$  (ou l'entier le plus proche).

Dans le cas où le bruit n'est plus prédominant, cela conduit à remplacer  $b_m^2$  (Eq. (13)) par l'expression donnée à l'équation (14) et donc à considérer la fonction de  $m$  définie ci-dessous :

$$f_p(m) = \left( \frac{1}{p-m} + \rho^2 + 2\rho \cdot \frac{1}{p-m} \right) \cdot \frac{1}{m^2}.$$

La fonction  $1/m^2$  étant décroissante (relativement à  $m$ ), on voit donc que le choix optimal du fil  $m$  se trouve compris entre  $2p/3$  et  $p-1$  et se rapproche de  $p-1$  à mesure que le rapport  $S/B$  augmente.

Ainsi, pour l'estimation des paramètres  $\hat{k}$  on voit qu'il y a un compromis à trouver entre le facteur de dilatation de l'échelle d'analyse (pour le fil  $m$ ,  $\hat{k}$  devient  $m\hat{k}$ ) qui minimise le terme géométrique et la « qualité statistique » de l'estimation des observations constituant les sorties du fil. C'est là un problème fondamental de cette analyse multi-échelles. De plus, il est à noter que la dilatation de l'échelle des  $\hat{k}$  n'induit pas de problème de repliement car

pour les problèmes d'intérêt pratique (en sonar passif par exemple) ces paramètres sont très petits.

On voit aussi qu'il est possible de remplacer la matrice de covariance asymptotique  $R$  par une estimée  $\hat{R}$  dans la formule (10) sans modification notable de la valeur de  $\text{var}(\hat{k}_m)$ . Ainsi, à partir des seules sorties de capteurs, est-il possible de déterminer la « stratégie » (choix du ou des « fils », pour notre problème) optimale pour l'estimation des paramètres des sources.

Cette discussion est illustrée par la figure 3 où sont comparées les variances expérimentales de  $\hat{k}_m$  (indexées par  $m$ ) obtenues pour 100 réalisations de la matrice  $\hat{R}$  et théoriques d'autre part (obtenues par (31)). L'antenne, linéaire, est constituée de 100 capteurs equirépartis. Dans ce cas (source unique à faible  $S/B$ ) on peut constater un fort bon accord entre les deux courbes, ce qui valide (31).

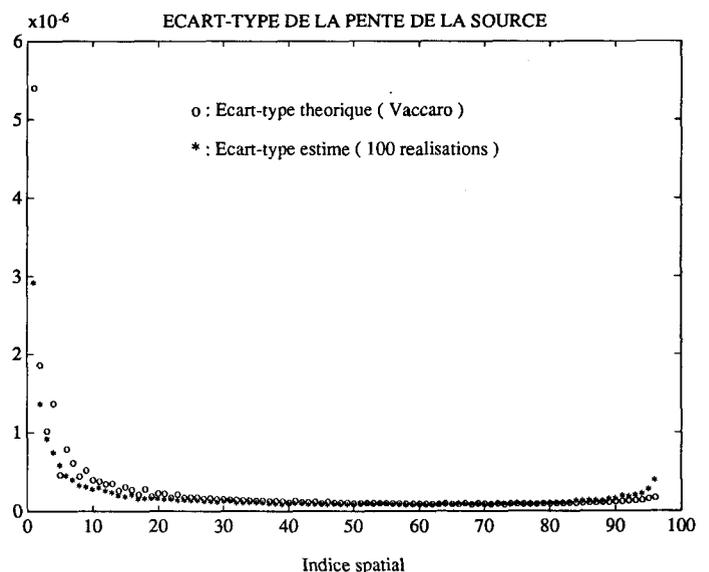


Figure 3. — Comparaison entre l'écart-type théorique de  $\hat{k}$  donné par (28) et son écart type expérimental ( $\sigma_1^2 = -10$  dB,  $\sigma_2^2 = -20$  dB,  $p = 17$ ,  $N = 100$ ) moyenné sur 50 réalisations.

On peut aussi remarquer que cette analyse peut être étendue sans difficulté à d'autres méthodes d'estimation de  $\hat{k}$  (modélisation AR, analyse de Fourier...). Dans tous les cas, l'analyse des performances suit les mêmes étapes. Le choix des fils optimaux ne sera donc guère différent.

Nous avons détaillé, dans ce paragraphe, un exemple d'utilisation de la formule fondamentale de cet article dans le cadre du traitement d'antenne. Il nous semble que ce simple exemple met en évidence l'importance du choix des observations pour les traitements sous-optimaux. Un choix convenable peut améliorer très sensiblement les performances des traitements.

Bien que l'application, présentée ici, soit limitée à l'estimation des paramètres  $\hat{k}$ , le cadre général de cette étude

pourrait s'appliquer à de nombreux problèmes d'une grande importance pratique, tels que :

- problèmes liés à l'utilisation de grandes antennes,
- faibles BT,
- problèmes multiplateformes (détermination de la distance optimale interplateformes),
- analyse multi-échelles,
- etc.

## 6. Conclusion

Nous avons considéré la variance d'estimation des corrélations spatiales estimées après rectification. Une formule simple a été prouvée à la section 2 (10). L'influence des moyennes spatiales pour des situations simples mais réalistes a été étudiée à la section 4.

Un exemple d'application en traitement d'antenne des calculs précédents est présenté dans la section 5. Tous les problèmes posés ne sont pas résolus mais nous pensons avoir posé des questions importantes que sont :

1. Faut-il considérer avec la même « importance » toutes les corrélations spatiales estimées (rectifiées) ?
2. Comment déterminer le compromis optimal entre le facteur de dilatation d'échelle (le choix des indices spatiaux par exemple) et la variance du bruit d'estimation des corrélations spatiales ?
3. Peut-on, à partir des observations et d'hypothèses physiques élémentaires, déterminer une configuration spatio-temporelle optimale ?
4. Comment optimiser les méthodes actuelles de traitement d'antenne (par exemple en tenant compte de la corrélation spatiale des covariances estimées rectifiées, etc.) ?

Quoiqu'il en soit, le but de cette correspondance n'est surtout pas de proposer une formule « magique » de traitement d'antenne mais de proposer des voies soit pour améliorer les traitements existants ou en définir de nouveaux.

## Remerciements

Les auteurs remercient les experts de la revue TS qui ont permis d'améliorer cet article.

## A. Annexe 1

Cette annexe a pour objet le calcul des termes entre crochets de la formule (6), i.e. :

$$\mathbb{E} \{ (X_k^* \cdot Z^i \cdot X_k) (X_k^* \cdot Z^{j'} \cdot X_k) \} . \quad (33)$$

Pour cela, l'ingrédient de base est constitué par l'expression de la forme quadratique  $X^* \cdot A \cdot X$  définie relativement au vecteur  $X$  supposé gaussien complexe, circulaire, la matrice  $A$  étant hermitienne. Cette fonction caractéristique (notée  $\Phi(t)$ ) est donnée par la formule (7), qui est une formule classique [10]. Toutefois, il importe de remarquer que les matrices  $Z^i$  ne sont pas hermitiennes. Il convient donc de se ramener à ce cas. Dans ce but, les formules suivantes s'avèrent fondamentales, et sont déduites (par dérivation) de l'expression  $\Phi(t)$  :

$$\mathbb{E} [(X^* \cdot A_1 \cdot X)^2] = \text{tr}^2 (A_1 R) + \text{tr} (A_1 R A_1 R)$$

$$\mathbb{E} [(X^* \cdot A_2 \cdot X)^2] = \text{tr}^2 (A_2 R) + \text{tr} (A_2 R A_2 R)$$

$$\mathbb{E} [X^* A_1 X \cdot X^* A_2 X] = \text{tr} [A_1 R A_2 R] + \text{tr} (A_1 R) \cdot \text{tr} (A_2 R)$$

avec :  $A_1$  et  $A_2$  hermitiennes, et  $X$  gaussien complexe circulaire ( $A_2$ ).

On a alors recours à l'écriture ci-dessous de la forme bilinéaire  $X^* \cdot Z \cdot X$  :

$$X^* \cdot Z \cdot X \triangleq f(X) = \text{Re} f(X) + i \text{Im} f(X)$$

avec :

$$\begin{aligned} \text{Re} f(X) &= \frac{1}{2} X^* (Z + Z') \cdot X \\ \text{Im} f(X) &= \frac{1}{2} X^* (-i \cdot (Z - Z')) \cdot X . \end{aligned} \quad (34)$$

Considérons la matrice  $-i(Z - Z')$ , on a :

$$(-i(Z - Z'))^* = i(Z' - Z) = -i(Z - Z') . \quad (35)$$

Les deux matrices  $(Z + Z')$  et  $-i(Z - Z')$  sont des matrices hermitiennes, il est alors possible d'appliquer les formules données ci-dessus. Si on considère l'exemple du calcul de  $\mathbb{E}(X^* \cdot Z \cdot X)$ , le formalisme précédent conduit aux calculs ci-dessous :

$$\begin{aligned} \mathbb{E} (|f(X)|^2) &= \frac{1}{4} [\mathbb{E}(X^* A_1 X^2) - \mathbb{E}(X^* A_2 X^2) - \\ &\quad - 2i \mathbb{E}(X^* A_1 X X^* A_2 X)] \end{aligned}$$

avec :

$$A_1 = Z + Z'$$

et

$$A_2 = -i(Z - Z') .$$

Soit en développant :

$$\begin{aligned} \mathbb{E} (|X^* \cdot Z \cdot X|^2) &= \\ &= \text{tr} (Z R) \text{tr} (Z' R) + \text{tr} (Z R Z' R) . \end{aligned} \quad (36)$$

Ce calcul s'étend directement au cas général. On notera cependant que la décomposition de  $X^* \cdot Z^i \cdot X$  en parties réelles et imaginaires semble être un intermédiaire de calcul indispensable.

## BIBLIOGRAPHIE

- [1] J. A. CADZOW, « Signal enhancement using canonical projection operators ». *ICASSP 87*, Dallas, pp. 1651-1654.
- [2] J. P. LE CADRE et P. LOPEZ, « Estimation d'une matrice interspectrale de structure imposée ». *Traitement du signal*, vol. TS-1, pp. 3-17, décembre 1984.
- [3] P. FORSTER, Méthodes de traitement d'antenne après filtrage spatial. *Thèse de doctorat de l'université de Rennes I*, novembre 1988.
- [4] B. G. LUCAS and J. P. LE CADRE, « Experimentation of spatial processing methods ». *Nato advanced study institute on Adaptive Methods in Underwater Acoustics*, H. G. Urban Ed., D. Reidel, 1984.
- [5] G. CYBENKO, « On the eigenstructure of Toeplitz matrices ». *IEEE Trans. on Acoust., Speech and Signal Processing*, vol. ASSP-32, n° 4, August 1984.
- [6] S. V. PARTER, « On the distribution of the singular values of Toeplitz matrices ». University of Wisconsin-Madison, *Computer Sciences Technical Report 609*, August 1985.
- [7] B. PORAT and B. FRIEDLANDER, « Parametric techniques for adaptive detection of Gaussian signals ». *IEEE Trans. on Acoust., Speech and Signal Processing*, vol. ASSP-32, n° 4, August 1984.
- [8] R. GODEMENT, *Cours d'algèbre*, Hermann, Paris 1966.
- [9] T. W. ANDERSON, *An Introduction to Multivariate Statistical Analysis*, 2nd Edition, 1984, Wiley.
- [10] K. S. MILLER, *Complex Stochastic Processes. An Introduction to Theory and Applications*. Addison-Wesley, 1974.
- [11] KOTZ, JOHNSON and READ, *Encyclopedia of Statistical Sciences*. Wiley, vol. 2, 1985 (9 vol.).
- [12] D. THUBERT, L. KOPP et J. P. LE CADRE, Précision des méthodes haute-résolution et bornes de Cramer-Rao. *Actes du GRETSI*, pp. 369-374, juin 1985.
- [13] R. J. VACCARO, D. W. TUFTS and G. FAYE BOUDREAU-BARTELS, *Advances in principal component signal processing, SVD and signal processing*, E. F. Deprettere (Editor), Elsevier Science Publishers B.V. (North-Holland), 1988.
- [14] O. ZUGMEYER et J. P. LE CADRE, « Méthodes d'intégration temporelle en traitement d'antenne ». *Rapport IRISA*, n° 566, décembre 1990.
- [15] J. P. LE CADRE et O. ZUGMEYER, Intégration temporelle en traitement d'antenne. *Actes du colloque GRETSI*, Juan-les-Pins, septembre 1991, pp. 689-692.
- [16] S. Y. KUNG, D. V. BHASKAR RAO and S. Y. KUNG, State space and singular value decomposition based approximation methods for the harmonic retrieval problem. *Journal of the optical society of America*, vol. 73, n° 12, décembre 1983, pp. 1789-1811.
- [17] O. ZUGMEYER et J. P. LE CADRE, Méthodes paramétriques pour la détection de sources en mouvement. *Rapport IRISA*, n° 567, décembre 1990.

Manuscrit reçu le 3 juin 1991.