

# Estimation au sens du maximum de vraisemblance d'un modèle autorégressif. Application au traitement spatial d'antennes

Maximum Likelihood estimation of an autoregressive model

Application to spatial signal processing



Jean-Pierre LE CADRE

CEPHAG, BP n 46, 38402 SAINT-MARTIN-D'HÈRES et GERDSM, Le  
Brusc, 83140 SIX-FOURS-LES-PLAGES

Jean-Pierre LE CADRE a soutenu une thèse de doctorat de 3<sup>e</sup> cycle en 1982 à l'INPG en traitement du signal. Depuis 1980 son activité au Groupe d'Études et de Recherche de Détection Sous-Marine porte essentiellement sur les traitements d'antennes appliqués à l'acoustique sous-marine.

## RÉSUMÉ

On détermine un algorithme récursif (sur l'ordre du modèle) d'estimation des coefficients d'un modèle autorégressif au sens du maximum de vraisemblance. Les coefficients ainsi estimés servent alors à estimer les densités spectrales des sources par la méthode du maximum d'entropie. Enfin, on présente des résultats de simulation permettant de comparer cet algorithme avec l'algorithme de Durbin-Levinson.

## MOTS CLÉS

Maximum de vraisemblance, modèle AR, Toeplitz, Pisarenko, pôles.

## SUMMARY

*A recursive algorithm (the recursion is on the order of model) related to the estimation of the coefficients of an autoregressive model based on maximization of likelihood is presented. The estimated coefficients are used to obtain the soatial power densities by means of maximum of entropy method. Lastly, simulation results are presented for comparison with Durbin-Levinson's algorithm.*

## KEY WORDS

## TABLE DES MATIÈRES

1. Introduction et motivations
  2. Présentation du problème
  3. Implications des équations de Yule-Walker
  4. Maximisation récursive de la fonction de la log-vraisemblance
  5. Utilisation en traitement spatial
  6. Performances de l'algorithme
  7. Application de la méthode à l'estimation au sens de maximum de vraisemblance Approche d'une matrice de structure imposée
  8. Conclusion
- Bibliographie

## Principales notations utilisées

- $A(n, m)$ , matrice à  $n$  lignes et  $m$  colonnes;  
 $A^t$ , transposée de la matrice  $A$ ;  
 $\begin{pmatrix} A \\ 0 \end{pmatrix}$ , vecteur de dimension  $\dim A + 1$ , sa dernière ligne étant 0;  
 $\tilde{A}$ , défini par :  $\tilde{A}(i) = A(n+1-i)$  pour un vecteur de dimension  $n$ , c'est le vecteur obtenu en inversant l'ordre des éléments puis en les conjuguant;  
 $A^*$ , matrice ligne transposée conjuguée du vecteur  $A$ ;  
 $\delta$ , distribution de Dirac;  
 $x^*$ , conjugué de  $x$ ;  
 $\det$ , déterminant d'une matrice;  
 $\text{Id}$ , matrice élément neutre de la multiplication.

## 1. Introduction et motivations

Le but de cet article est d'explicitier un algorithme d'estimation des coefficients d'un modèle autorégressif (AR) au sens du maximum de vraisemblance pour le traitement de signaux d'antennes.

Cet article n'a pas pour objet d'ajouter une méthode nouvelle aux déjà nombreuses méthodes d'analyse spectrale; il répond au besoin de déterminer un estimateur des paramètres AR possédant une meilleure stabilité statistique que l'estimateur déduit des équations de Yule-Walker.

Cette recherche de stabilité a déjà fait l'objet de nombreux travaux pour l'analyse spectrale d'un signal monodimensionnel parmi lesquels on peut citer l'algo-

rithme de Burg utilisant les prédictions aller et retour [1, 3, 7]. Ce type d'algorithme a été étendu à l'estimation des paramètres d'un modèle AR multidimensionnel [11]. Mais le but poursuivi ici est l'utilisation d'une modélisation AR des capteurs d'une antenne en vue de l'estimation des densités spectrales des sources.

La méthode du maximum d'entropie [7, 3] nécessite le calcul des coefficients déterminés par l'équation de Yule-Walker. Elle possède un pouvoir de résolution angulaire élevé mais présente l'inconvénient de conduire à des estimations très peu stables des densités spectrales des sources. C'est pourquoi on développe ici un algorithme d'estimation des paramètres d'une modélisation AR des sorties de capteurs au sens du maximum de vraisemblance. Pour arriver à des calculs assez simples, on ne considère qu'une méthode récursive de maximisation de la vraisemblance, la récursion portant sur l'ordre du modèle. Ce n'est donc pas une méthode de maximisation globale de la vraisemblance, mais comme on le verra, le volume de calcul nécessite est guère plus important que celui nécessité par l'algorithme de Durbin-Levinson utilisé pour résoudre l'équation de Yule-Walker. Cette méthode a été développée dans le cas monodimensionnel par Kay et Clergeot [9, 4], la formulation et les développements sont cependant assez différents dans ce cas.

L'estimation des paramètres AR par la maximisation de la vraisemblance peut en général conduire à de meilleures estimations (au sens sens statistique) que celles obtenues par une méthode du type moindres carrés (Yule-Walker) et ceci explique l'intérêt de la méthode même si la borne de Cramer-Rao (pour les paramètres d'un modèle AR) ne peut être atteinte qu'asymptotiquement.

Le plan de cet article va être le suivant : après une présentation du problème, on étudie les implications des équations de Yule-Walker, de ceci on dérive le problème de maximisation sous forme récursive, puis la résolution explicite de ce problème. Enfin on étudie sur des signaux simulés les propriétés statistiques des estimateurs obtenus.

## 2. Présentation du problème

On considère une antenne linéaire constituée de  $N$  capteurs équirépartis, celle-ci fournit après transformation de Fourier, un vecteur appelé par la suite vecteur d'observation  $X(f)$  ( $f$ , fréquence considérée, dans la suite, la fréquence considérée étant constante, le terme  $f$  ne sera pas indiqué). On suppose en outre que les divers vecteurs  $X(f)$  sont indépendants (nombre de points de FFT suffisants par exemple) et gaussiens.

Le but est de déterminer un modèle AR décrivant le vecteur  $X$ , i. e. :

$$(1) \quad x_n = \sum_{i=1}^{n-1} a_i x_{n-i} + e$$

$$A'_n = (1, a_1, \dots, a_{n-1}).$$

( $x_n$ , sortie du  $n$ -ième capteur de l'antenne;  $e$ , bruit blanc).



Le calcul développé dans le paragraphe suit le schéma suivant :

1° Grâce à la décomposition de Choleski (6) de  $M_n$ , on obtient une expression explicite de  $\det M_n$  en fonction de  $K_{n-1}$  et  $\sigma_{n-1}$  et du terme quadratique.

2° On maximise la fonction ainsi déterminée en fonction de  $\sigma_{n-1}$ , la log-vraisemblance n'est plus alors qu'une fonction de  $K_{n-1}$ .

3° On maximise cette fonctionnelle relativement à  $K_{n-1}$ ,  $K_{n-1}$  est déterminé par résolution d'une équation du troisième degré à coefficients réels.

4. 1. EXPRESSION RÉCURSIVE DE  $\det M$

D'après (6) :

$$\det \Delta_n = \det T_n \cdot \det M_n \cdot \det(T_n^*),$$

or  $T_n$  étant triangulaire (diagonale formée de 1),  $\det T_n = 1$ , d'où :

$$\det M_n = \det \Delta_n = \prod_{i=0}^{n-1} \sigma_i,$$

or (5) :

$$\sigma_i = \sigma_{i-1} \cdot (1 - |K_{i-1}|^2),$$

d'où :

$$\begin{aligned} \sigma_{n-1} &= \sigma_{n-1}, \\ \sigma_{n-2} &= \sigma_{n-1} \cdot (1 - |K_{n-1}|^2)^{-1} = \sigma_{n-1} \cdot f_{n-2}, \\ &\vdots \\ \sigma_1 &= \sigma_{n-1} \cdot (1 - |K_{n-1}|^2)^{-1} \times \dots \\ &\quad \times (1 - |K_2|^2)^{-1} = \sigma_{n-1} \cdot f_1, \\ \sigma_0 &= \sigma_{n-1} \cdot (1 - |K_{n-1}|^2)^{-1} \times \dots \\ &\quad \times (1 - |K_1|^2)^{-1} = \sigma_{n-1} \cdot f_0, \end{aligned}$$

on en déduit que :

$$(7) \quad \begin{aligned} \text{Log } \det M_n &= n \text{ Log } \sigma_{n-1} \\ &- (n-1) \cdot \text{Log}(1 - |K_{n-1}|^2) - \dots - \text{Log}(1 - |K_1|^2). \end{aligned}$$

4. 2. MAXIMISATION RELATIVE A  $\sigma_{n-1}$

De la formule (6), on déduit que :

$$\begin{aligned} X^* \cdot M^{-1} \cdot X &= (T_n X)^* \cdot \Delta_n^{-1} \cdot (T_n X) \\ &= \sigma_{n-1}^{-1} (T_n X)^* \delta_n (T_n X), \end{aligned}$$

où  $\delta_n$  est une matrice diagonale ne faisant pas intervenir  $\sigma_{n-1}$  :

$$\delta_n = \begin{bmatrix} f_0^{-1} & & & 0 \\ & f_1^{-1} & & \\ & & \ddots & \\ 0 & & & f_{n-1}^{-1} \end{bmatrix}$$

D'où l'expression suivante de la log-vraisemblance :

$$\begin{aligned} L_n(X_1, \dots, X_m/A, \sigma_e) \\ = -m \cdot n \text{ Log } (\sigma_{n-1}) \\ - (\sigma_{n-1})^{-1} \cdot \sum_{i=1}^m (T_n X_i)^* \cdot \delta_n (T_n X_i) + C_0. \end{aligned}$$

$C_0$  ne dépendant pas de  $\sigma_{n-1}$ .

Par dérivation de l'expression précédente, on obtient l'estimation au sens du maximum de vraisemblance de  $\sigma_{n-1}$ , soit :

$$(8) \quad \sigma_{n-1} = \frac{1}{n \cdot m} \cdot \sum_{i=1}^m (T_n X_i)^* \cdot \delta_n \cdot (T_n X_i).$$

On en déduit une expression de  $L_n$  relative à  $K_{n-1}$  seulement, i. e. :

$$\begin{aligned} (9) \quad L_n(X_1, \dots, X_m/A, \sigma_e) \\ = -m \cdot n \text{ Log} \left( \frac{1}{m \cdot n} \cdot \sum_{i=1}^m (T_n X_i)^* \cdot \delta_n (T_n X_i) \right) \\ + m \cdot (n-1) \text{ Log}(1 - |K_{n-1}|^2) + C_1. \end{aligned}$$

$C_1$  ne dépendant pas de  $K_{n-1}$ .

4. 3. DÉTERMINATION DE  $K_{n-1}$

Il reste à expliquer  $(T_n X_i)^* \cdot \delta_n \cdot (T_n X_i)$  en fonction de  $K_{n-1}$ . Or :

$$\delta_n = (1 - |K_{n-1}|^2) \cdot \gamma_n,$$

avec  $\gamma_n$  matrice diagonale définie par :

$$\gamma_n = \begin{pmatrix} (1 - |K_{n-2}|^2) & \dots & (1 - |K_1|^2) \\ & & 0 \\ & & & & 1 \\ & & & & (1 - |K_{n-1}|^2)^{-1} \end{pmatrix}$$

$\gamma_n$ , matrice diagonale

$$\Rightarrow L_n(\bar{X}_1, \dots, \bar{X}_m/\bar{A}, \sigma_e)$$

$$(10) \quad \begin{aligned} = -mn \text{ Log} \left( \frac{1}{m} \cdot \sum_{i=1}^m (T_n X_i)^* \cdot \gamma_n \cdot T_n X_i \right) \\ - m \text{ Log}(1 - |K_{n-1}|^2) + C_1. \end{aligned}$$

L'avantage de la formulation (10) de  $L_n$  est le suivant : seule la dernière ligne de  $T_n$  (notée  $A_n^i$ ) et l'élément de la dernière ligne et dernière colonne de  $\gamma_n$  dépendent de  $K_{n-1}$ , d'où :

$$\begin{aligned} L_n(X_1, \dots, X_m/A) \\ = -mn \text{ Log} \left( A_n^* \cdot \left( \frac{1}{m} \cdot \sum_{i=1}^m X_i \cdot X_i^* \right) \cdot A_n \right) \\ + m(n-1) \text{ Log}(1 - |K_{n-1}|^2) + C_2. \end{aligned}$$

$C_2$  ne dépendant pas de  $K_{n-1}$ .

De plus,  $A_n$  s'exprime en fonction de  $K_{n-1}$  et  $A_{n-1}$  (supposé connu) par la relation vectorielle (4) :

$$A_n = \begin{pmatrix} A_{n-1} \\ 0 \end{pmatrix} + K_{n-1} \begin{pmatrix} 0 \\ \tilde{A}_{n-1} \end{pmatrix}.$$

Soit  $\Gamma$  la matrice  $1/m \cdot \sum_{i=1}^m X_i \cdot X_i^*$ ,  $\Gamma$  correspond à une estimation de la matrice interspectrale obtenue par la méthode du périodogramme. (On verra au paragraphe 5 que cette matrice doit être légèrement modifiée pour s'adapter aux données du problème.) La partition de  $A_n$  donnée par la formulation (4) impose de découper  $\Gamma(n, n)$  en blocs afin d'effectuer le produit  $A_n^* \cdot \Gamma \cdot A_n$  par multiplication des blocs et d'explicitier ainsi  $K_{n-1}$  dans cette expression, on écrit donc :

$$(11) \quad \Gamma = \begin{bmatrix} \Gamma(1, 1) & L_{n-2}^* & \bar{\Gamma}(n, 1) \\ L_{n-2} & \Gamma_{n-2} & H_{n-2} \\ \Gamma(n, 1) & H_{n-2}^* & \Gamma(n, n) \end{bmatrix} \\ = \begin{bmatrix} & & \bar{\Gamma}(n, 1) \\ & \Gamma_{n-1} & H_{n-2} \\ \Gamma(n, 1) & H_{n-2}^* & \Gamma(n, n) \end{bmatrix}.$$

Où  $\Gamma(1, 1)$  est l'élément de la première ligne et colonne de  $\Gamma$ ,  $\Gamma_{n-2}$  est la matrice  $(n-2, n-2)$  obtenue en extrayant les éléments de  $\Gamma$  par la formulation :

$$\Gamma_{n-2}(i, j) = \Gamma(i+1, j+1), \quad 1 \leq i \leq n-2, \quad 1 \leq j \leq n-2$$

et  $L_{n-2}$  et  $H_{n-2}$  sont les vecteurs colonnes résultats du découpage précédent.

Mais surtout il importe de considérer que malgré son apparence compliquée, ce découpage de  $\Gamma$  s'impose naturellement en considérant la formulation (4) de  $A_n$ . De (11) et (4) on déduit que :

$$(12) \quad A_n^* \cdot \Gamma \cdot A_n = A_{n-1}^* \cdot \Gamma_{n-1} \cdot A_{n-1} \\ + |K_{n-1}|^2 \cdot \tilde{A}_{n-1}^* \cdot \begin{pmatrix} \Gamma_{n-2} & H_{n-2} \\ H_{n-2}^* & \Gamma(n, n) \end{pmatrix} \cdot \tilde{A}_{n-1} \\ + K_{n-1} \cdot A_{n-1}^* \cdot \begin{pmatrix} L_{n-2}^* & \bar{\Gamma}(n, 1) \\ \Gamma_{n-2} & H_{(n-2)} \end{pmatrix} \tilde{A}_{n-1} \\ + \bar{K}_{n-1} \tilde{A}_{n-1}^* \cdot \begin{pmatrix} L_{n-2} & \Gamma_{n-2} \\ \Gamma(n, 1) & H_{n-2}^* \end{pmatrix} \cdot n^{-1}.$$

Ce qu'on écrit sous la forme suivante :

$$A_n^* \cdot \Gamma \cdot A_n = \mathcal{E}_n = \mathcal{E}_{n-1} \\ + 2 \operatorname{Re}(K_{n-1} \cdot C_n) + |K_{n-1}|^2 \cdot D_n,$$

avec :

$$D_n = \tilde{A}_{n-1}^* \cdot \begin{pmatrix} \Gamma_{n-2} & H_{n-2} \\ H_{n-2} & \Gamma(n, n) \end{pmatrix} \cdot \tilde{A}_{n-1}, \\ C_n = A_{n-1}^* \cdot \begin{pmatrix} L_{n-2}^* & \bar{\Gamma}(n, 1) \\ \Gamma_{n-2} & H_{n-2} \end{pmatrix} \cdot \tilde{A}_{n-1}.$$

Alors  $L_n(X_1, \dots, X_m)$  s'exprime comme une fonction de  $K_{n-1}$  (en négligeant les termes ne faisant pas intervenir  $K_{n-1}$ ), soit :

$$(13) \quad L_n(X_1, \dots, X_m / K_{n-1}) \\ = -m \cdot n \operatorname{Log}(\mathcal{E}_{n-1} + 2 \operatorname{Re}(K_{n-1} \cdot C_n) + |K_{n-1}|^2 \cdot D_n) \\ + m \cdot (n-1) \operatorname{Log}(1 - |K_{n-1}|^2).$$

Maintenant,  $L_n$  s'exprime comme une fonction explicite de  $K_{n-1}$  et il s'agit donc de déterminer le coefficient de réflexion  $K_{n-1}$  maximisant (13).

#### 4.4. MAXIMISATION DE LA LOG-VRAISEMBLANCE RELATIVEMENT A $K_{n-1}$

On peut remarquer que nécessairement  $K_{n-1}$  se trouve dans le disque unité, de plus, si  $|K_{n-1}| \rightarrow 1$  alors la log-vraisemblance tend vers  $-\infty$ ,  $L_n$  possède donc au moins un maximum dans le disque unité puisqu'elle est continue dans le disque (théorème des valeurs intermédiaires).

Posons :

$$K_{n-1} = x + iy, \\ C_n = a + ib.$$

En dérivant  $L_n$  par rapport aux parties réelles et imaginaires de  $K_{n-1}$ , on obtient que nécessairement en un extrémum :

$$(14) \quad y = -\frac{b}{a} x.$$

De (14), on déduit l'expression de  $L_n$  en fonction du seul paramètre  $x$ , d'où l'équation nécessairement vérifiée par  $x$  en un extrémum :

$$(15) \quad x^3 \cdot \left( \frac{a^2 + b^2}{a^2} \right) D_n \cdot (\alpha_n - 1) \\ + x^2 \cdot \left( \frac{a^2 + b^2}{a} \right) \cdot (2\alpha_n - 1) \\ + x \cdot (\alpha_n \cdot \mathcal{E}_{n-1} + D_n) + a = 0,$$

où :

$$\alpha_n = \frac{n-1}{n}.$$

Cette équation du troisième degré à coefficients réels a une ou trois solutions réelles dans le disque (au moins une!). Sa résolution est obtenue par les formu-



Données des simulations des figures 1 et 2 :

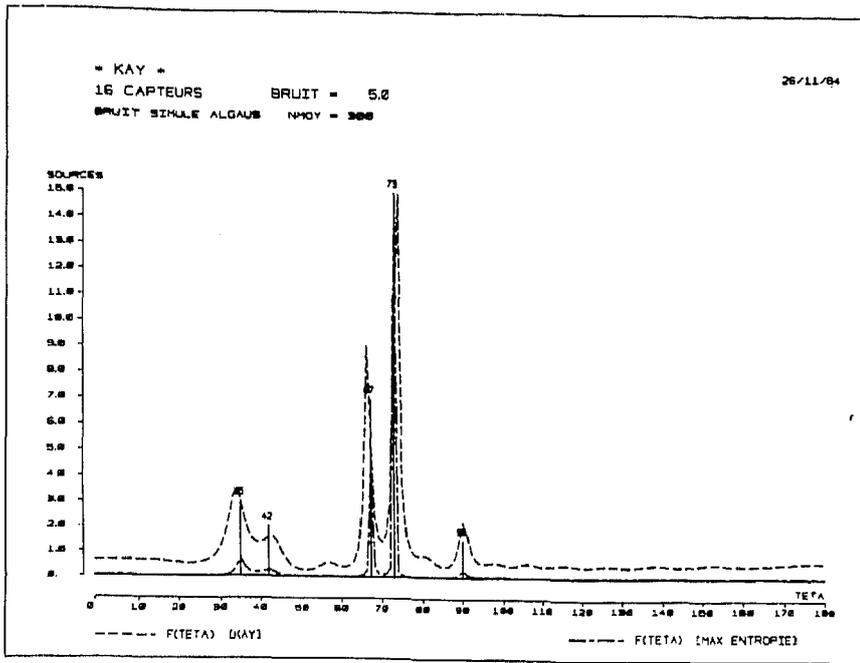


Fig. 1

Figure 1  
 NMOY = 300  
 b = 5  
 ns = 5

$\sigma_1 = 3, \theta_1 = 35$   
 $\sigma_2 = 2, \theta_2 = 42$   
 $\sigma_3 = 7, \theta_3 = 67$   
 $\sigma_4 = 15, \theta_4 = 73$   
 $\sigma_5 = 1, \theta_5 = 90$

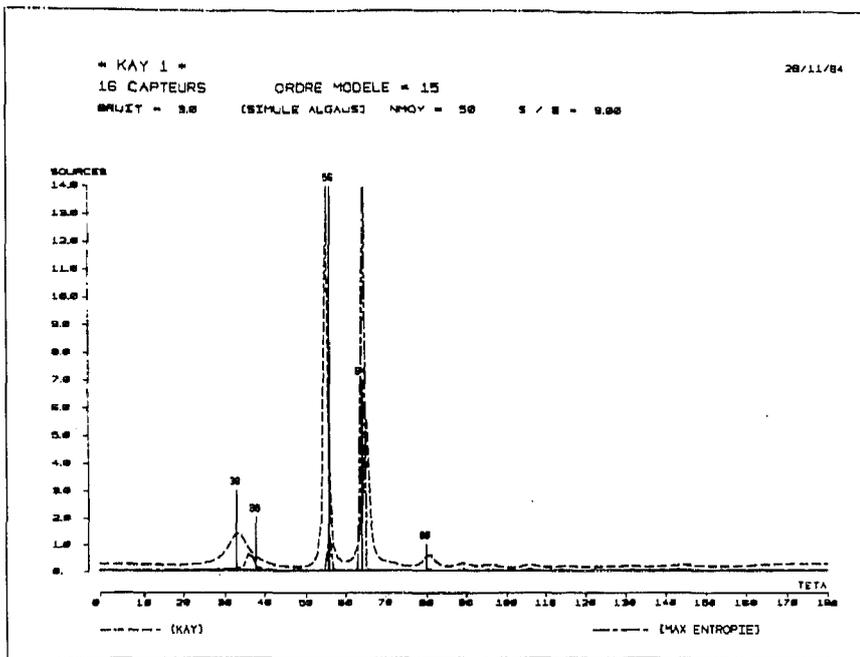


Fig. 2

Figure 2  
 NMOY = 50  
 b = 3  
 ns = 5

$\sigma_1 = 3, \theta_1 = 33$   
 $\sigma_2 = 2, \theta_2 = 38$   
 $\sigma_3 = 14, \theta_3 = 56$   
 $\sigma_4 = 7, \theta_4 = 64$   
 $\sigma_5 = 1, \theta_5 = 80$

On peut remarquer que les calculs nécessaires ne sont pas beaucoup plus importants que ceux nécessités par l'algorithme de Durbin-Levinson.

## 6. Performances de l'algorithme

On utilise pour cela une simulation définie de la façon suivante :

X est un vecteur gaussien dont la matrice de covariance s'écrit :

$$(18) \quad C = \sum_{i=1}^{ns} \sigma_i D_{\theta_i} D_{\theta_i}^* + b \text{Id.}$$

On fait l'hypothèse que X est gaussien centré, cette hypothèse semblant justifiée dans le cadre de l'écoute

passive (en acoustique sous-marine) par les résultats d'expérimentations, où :  $ns$  est le nombre de sources simulées;  $\{(\sigma_b, \sigma_i)\}_{i=1}^{ns}$ , les densités spectrales et les gisements des sources;  $b$ , la densité spectrale du bruit. La simulation du vecteur  $X$  correspondant est obtenue par décomposition de Choleski de la matrice  $C$ . On indiquera sur les résultats de simulation  $ns$ , les couples :  $\{\sigma_b, \theta_i\}_{i=1}^{ns}$  et  $b$ , ainsi que le nombre de moyennages utilisés pour estimer  $\Gamma'$  (ce qui détermine le facteur BT correspondant et donc la variance d'estimation des éléments de  $\Gamma'$ ).

On a cherché à comparer les propriétés de deux estimateurs des paramètres d'un modèle AR des sorties des capteurs de l'antenne. L'un est celui développé au cours de cet article (maximisation récursive de la vraisemblance), l'autre est celui obtenu par la résolution des équations de Yule-Walker par l'algorithme de Durbin-Levinson appliqué à la matrice de Toeplitz obtenue par moyennage des diagonales de la matrice du périodogramme [7].

Sur les figures 1 et 2, on présente les résultats obtenus par les deux méthodes,  $P(\theta)$  est représentée avec une échelle linéaire et les sources sont représentées avec leurs gisements et leurs puissances par un trait vertical dont la longueur représente la puissance de la source. Les gisements sont représentés de 0 à 180°. La courbe en tirets courts (notée Kay) correspond à la courbe

obtenue par la maximisation de la vraisemblance et celle en tirets longs correspond à l'application de l'algorithme de Durbin-Levinson.

Sur la figure 1, il y a cinq sources dont deux fortes, la puissance des trois sources faibles est mal estimée par Durbin-Levinson, par contre, elle est bien estimée (relativement aux sources fortes) par la maximisation de la vraisemblance.

Ce résultat est confirmé sur la figure 2 où de plus la source forte dans le gisement 56 est très mal estimée par Durbin-Levinson, quant à la source dans le gisement 80, elle n'est pas vue par cette méthode.

On présente sur les figures 3 et 4 les résultats obtenus pour dix réalisations d'une simulation dont les paramètres (puissances et gisements des sources, niveau du bruit ambiant, facteur BT) sont indiqués en haut des figures.

La mention « Max entropie » correspond à l'application de l'algorithme de Durbin-Levinson, alors que la mention « Max vraisemblance » correspond à l'application de la méthode développée dans le présent article. On peut noter la stabilité de la méthode du maximum de vraisemblance pour l'estimation des puissances des sources qui contraste avec l'instabilité de la solution directe des équations de Yule-Walker. Ceci étant particulièrement visible pour les deux sources fortes (dans les gisements 75 et 82).

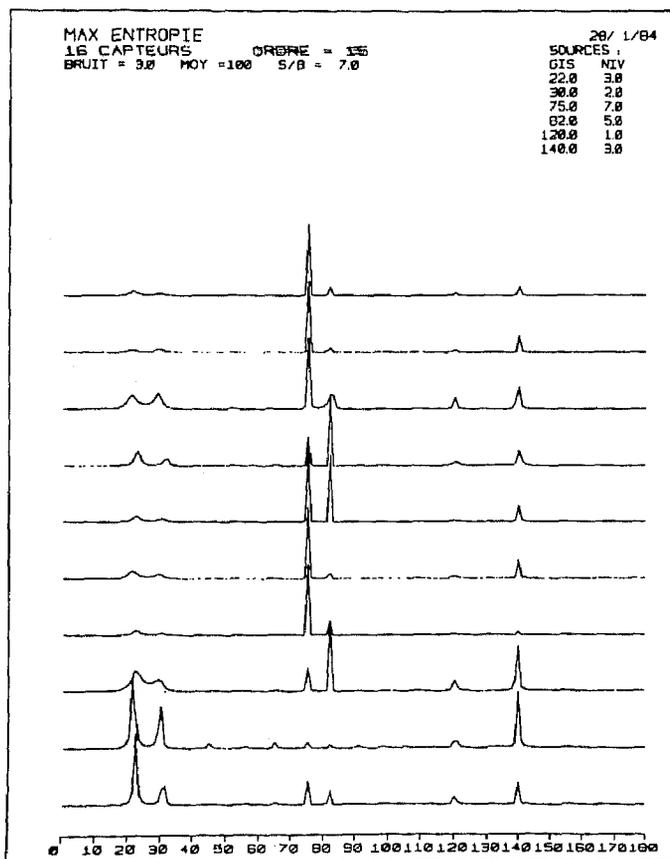


Fig. 3. — Comparaison des deux méthodes.

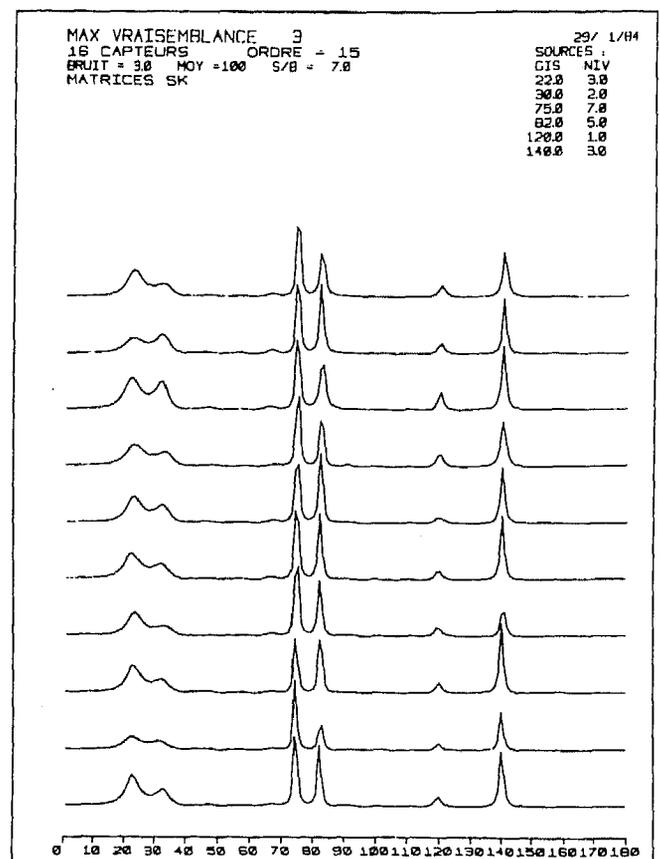


Fig. 4. — Comparaison des deux méthodes.

L'ordre 15 est l'ordre maximal (il y a 16 capteurs); suivant une idée répandue en analyse spectrale, on a donc réduit l'ordre du modèle jusqu'à l'ordre 8 (ce qui correspond au nombre de capteurs divisés par 2).

On peut constater sur la figure 5 la dégradation du pouvoir séparateur, alors que par contre, les pics parasites semblent moins importants (ceci est surtout visible à faible rapport signal sur bruit).

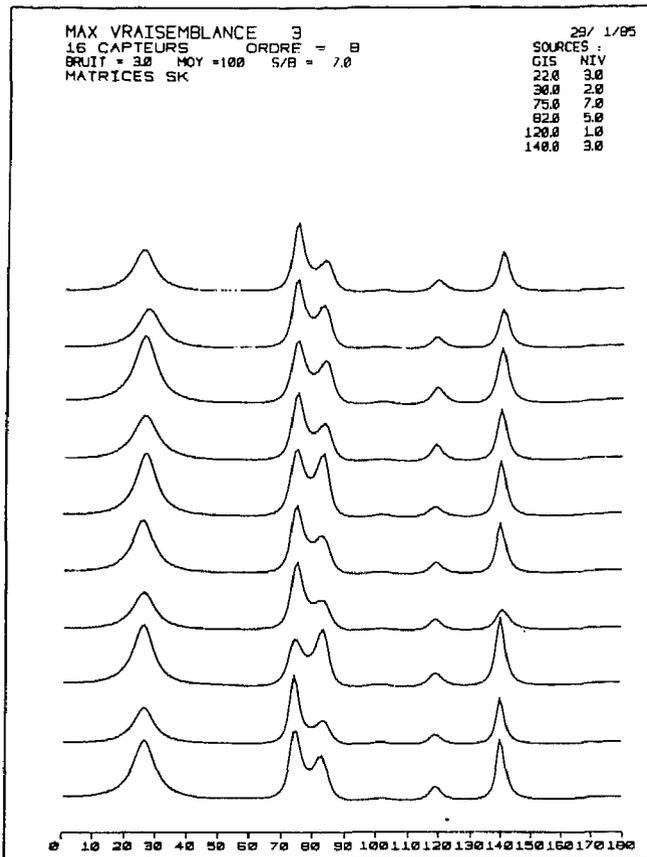


Fig. 5. — Influence de l'ordre du modèle.

Enfin sur les figures 6 et 7, on présente les résultats obtenus par les deux méthodes avec des sources de faible niveau (par rapport au bruit ambiant). Dans ce cas, les résultats apparaissent semblables, en particulier les pics parasites sont semblables. La solution directe des équations de Yule-Walker présente une bonne stabilité (cette stabilité est-elle due au niveau élevé du bruit ambiant?). Dans le cas d'un bruit (ambiant) élevé, la modélisation AR ne paraît pas appropriée, ce qui a été démontré par de nombreux auteurs; et le fait d'estimer les coefficients du modèle AR par maximisation approchée de la vraisemblance ne change pas le fond du problème qui est l'adéquation de la modélisation.

La figure 8 utilise un autre type de présentation, précisément on présente les pôles de la fonction de

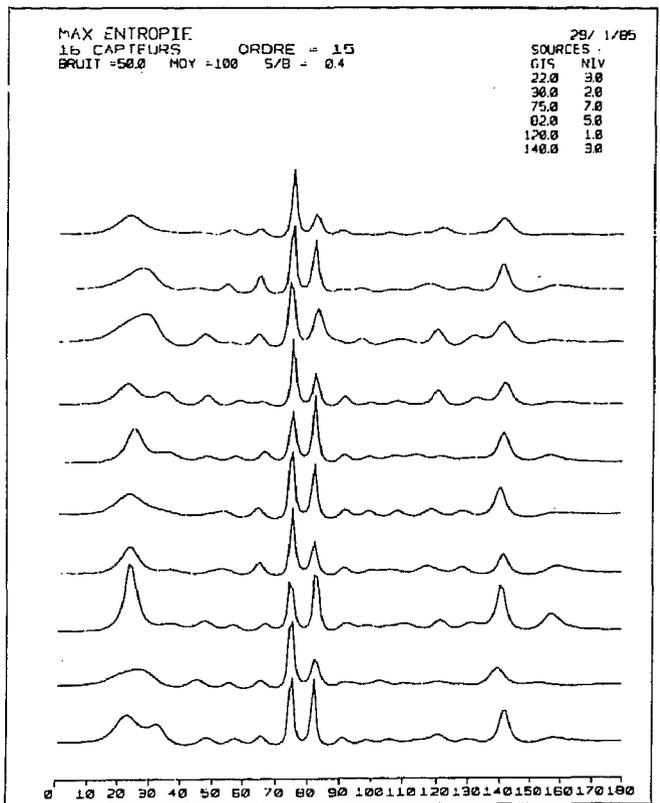


Fig. 6. — Influence du rapport signal sur bruit.

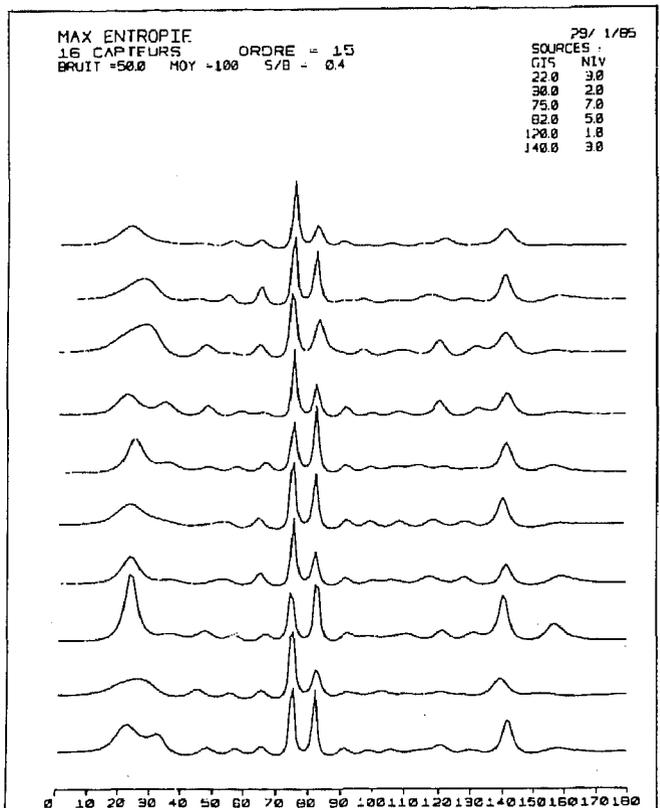
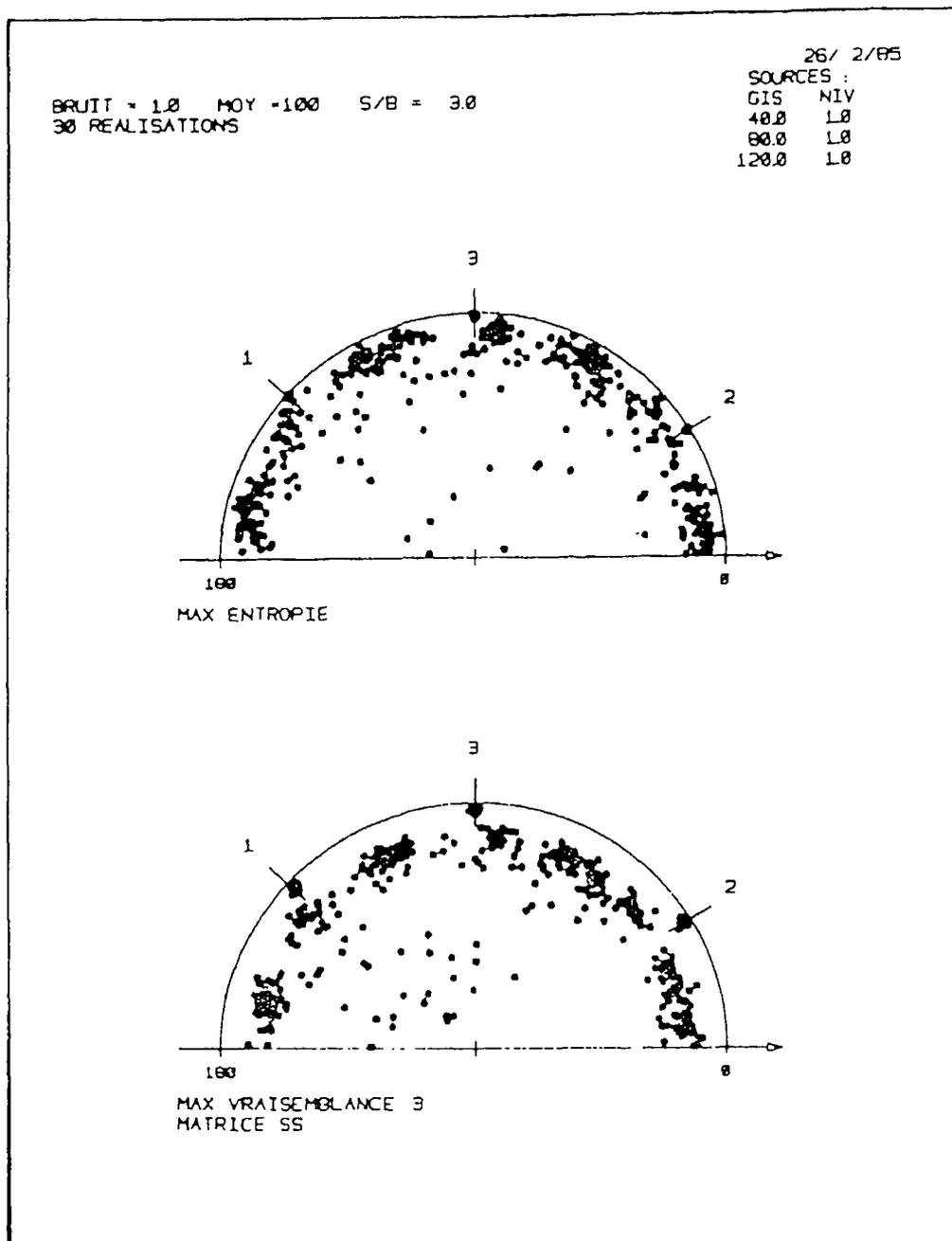


Fig. 7. — Influence du rapport signal sur bruit.



8. — Influence de la méthode d'estimation des coefficients du modèle AR sur la position des pôles du modèle estimé.

transfert du modèle AR. Ces pôles sont déterminés à partir des coefficients AR par le calcul des racines du polynôme en  $Z$  associée au modèle. Les racines étant déterminées par un programme utilisant la méthode de Laguerre.

Dans cette simulation, il y a 3 sources, l'ordre étant maximal, il y a donc 16 pôles. Les traits 1, 2, 3 correspondent aux positions exactes des bruiteurs simulés qui sont définis par leurs gisements et leurs puissances respectives. De plus, on considère 30 réalisations de cette simulation.

On observe que les pôles s'accroissent autour des pôles correspondants aux bruiteurs pour les deux méthodes. Mais la méthode du maximum de vraisemblance ramène les pôles non liés aux sources vers l'intérieur du cercle unité (sur une couronne de rayon assez nettement inférieur à 1). La méthode du maximum de vraisemblance apparaît donc comme stabilisante, ce qui n'est guère étonnant au vue de l'expression du déterminant de la matrice  $M_n$ .

Il faut toutefois noter que la méthode du maximum de vraisemblance a tendance à ramener tous les pôles

\* KAYTOP \* MAX VRAISEMBLANCE  
16 CAPT ORDRE = 15 MOY = 150

Sources : 1 2 3  
Gisement : 48.0 55.0 62.0  
Niveau : 1.5 0.5 5.0  
Niv Bruit : 1.5

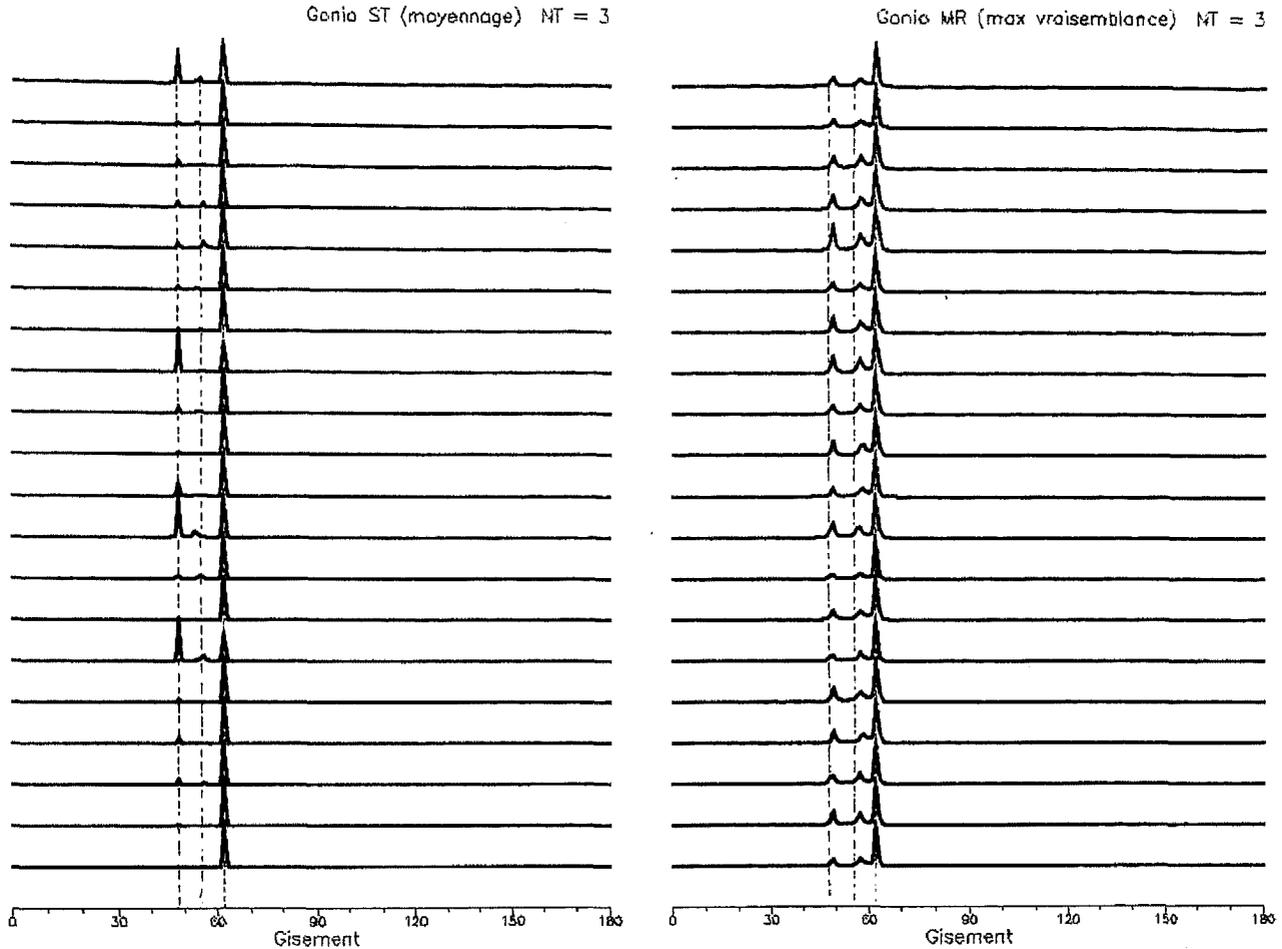


Fig. 9. — Localisation de sources par la méthode de Pisarenko en utilisant la matrice interspectrale approchée par moyennage des diagonales (Gonio ST) ou par la méthode du maximum de vraisemblance (Gonio MR).

à l'intérieur du cercle unité; ainsi quand la puissance du bruit augmente, les pôles dus aux bruiteurs sont également ramenés vers l'intérieur du cercle unité.

### 7. Application de la méthode a l'estimation au sens du maximum de vraisemblance approché d'une matrice de structure imposée

Une application de la méthode qui vient d'être décrite peut être l'estimation d'une matrice de covariance (ou interspectrale) de structure imposée (Toeplitz) au sens du maximum de vraisemblance approché.

Pour cela, on considère la matrice  $\Gamma'_{n+1}$  définie par la formule (16), par application de la méthode on obtient une suite de coefficients  $(1, a_1, \dots, a_n)$  décrivant la modélisation AR. Le problème est maintenant d'estimer la matrice de covariance (ou interspectrale) de Toeplitz, correspondant à ce modèle AR.

Le problème est de calculer les covariances correspondantes définies par :

$$(19) \quad \Gamma_k = \frac{1}{2i\pi} \oint H(Z) \cdot H(Z^{-1}) \cdot Z^k \frac{dZ}{Z}$$

où  $H(Z)$  est la fonction de transfert du modèle, soit :

$$H(Z) = (a_0 + a_1 Z + \dots + a_n Z^n)^{-1}, \\ Z = \exp\left(2i\pi f \frac{d \cos \theta}{C}\right), \quad a_0 = 1.$$

La suite des réponses impulsionnelles correspondant à  $H(Z)$  est la suite :  $\{h_k\}_{-\infty}^{+\infty}$  définie par :

$$(20) \quad h_k = \begin{cases} 0 & \text{pour } k < 0, \\ \delta(k) - \sum_{i=1}^n a_i h_{i+k} & \text{pour } k \geq 0. \end{cases}$$

D'autre part :

$$(21) \quad r_k = \sum_{i=0}^{+\infty} h_i h_{i+k} \quad k \geq 0,$$

$$r_k = r_{-k} \quad (\text{dans le cas réel}),$$

d'où :

$$r_k = \sum_{j=0}^{+\infty} h_j h_{j+k}$$

$$= \sum_{j=0}^{+\infty} h_j \left( \delta(j+k) - \sum_{i=1}^n a_i h_{j+k-i} \right)$$

$$= \sum_{j=0}^{+\infty} h_j \cdot \delta(j+k) - \sum_{i=1}^n a_i \left( \sum_{j=0}^{+\infty} h_j h_{j+k-i} \right).$$

Soit encore :

$$(22) \quad \sum_{i=0}^n a_i r_{k-i} = \sum_{j=0}^{+\infty} h_j \cdot \delta(j+k), \quad k \geq 0.$$

Cette expression peut encore s'écrire sous forme d'un système linéaire :

$$(23) \quad \mathbf{A} \cdot \mathbf{R} = \mathbf{D},$$

avec :

$$\mathbf{R}^t = (r_0, r_1, \dots, r_n),$$

$$\mathbf{D}^t = (\mathbf{E}, 0, \dots, 0).$$

A étant définie par :

$$(24) \quad A(i, j) = \begin{cases} a_{i-1} & \text{pour } j=1, \\ a_{i+j-2} + a_{i-j} & \text{pour } j>1. \end{cases}$$

Cette expression a été obtenue par Dugré, Beex et Scharf [6] pour le calcul des covariances correspondant à un modèle ARMA.

Dans le cas de covariances complexes (ce qui est le cas du traitement spatial) l'expression (24) de A(i, j) ne s'applique plus, l'approche qui nous a semblée la plus simple a été de considérer (22) dans le cas complexe ( $r_{-k} = r_k^*$ ) puis de décomplexifier le problème, on obtient alors le vecteur  $\mathbf{R}$  comme solution d'un système linéaire du même type que celui défini par (23), (24) mais de dimension double.

De plus, il a été prouvé que pour un modèle AR stable la matrice A définie par (24) est inversible [6]. Dans toutes les simulations effectuées, la matrice de Toeplitz associée au vecteur  $\mathbf{R}$ , i. e. :

$$T(i, j) = r_{i-j}$$

s'est avérée être positive.

Ainsi sur la figure 9, on présente les résultats d'une simulation définie de la façon suivante :

1° Les sources et le bruit sont définis de façon identique à celle employée au paragraphe 6.

2° On applique la méthode de Pisarenko (projection sur le sous-espace vectoriel engendré par les vecteurs propres associés aux dix valeurs propres minimales) concurremment à deux matrices de Toeplitz. D'une part à la matrice projection orthogonale de la matrice du périodogramme sur le sous-espace vectoriel des matrices de Toeplitz. [Courbes de gauche notées ST (moyennage).] D'autre part, à la matrice de Toeplitz obtenue à partir de l'estimation des coefficients AR au sens du maximum de vraisemblance approché à l'aide des formules (23) et (24). [Courbes de droite notées MR (max de vraisemblance).]

On peut constater que les courbes de droite (MR) semblent présenter une meilleure stabilité statistique, de ce fait les deux sources faibles apparaissent plus nettement (surtout celle dans le gisement 55). Dans ce cas, la méthode exposée dans le paragraphe apporte une amélioration sensible de la méthode de Pisarenko, au prix de calculs assez modestes vis-à-vis de ceux nécessités par le calcul exact de la vraisemblance. Toutefois, il faut remarquer que l'avantage n'est pas toujours aussi marqué.

On présente sur le tableau, les résultats statistiques obtenus sur un nombre donné de réalisations par les deux estimateurs. On peut remarquer que la variance d'estimation de la puissance des sources faibles est, en général, bien plus petite pour la méthode du maximum de vraisemblance que pour la méthode de Levinson, le biais peut aussi être bien plus petit. Ceci est en accord avec les résultats trouvés par Kay [9] en analyse spectrale.

## 8. Conclusion

L'algorithme présenté dans cet article conduit à une maximisation récursive de la vraisemblance au prix de calculs assez réduits. On peut rapprocher cette méthode de celles développées en [10], puisqu'elle peut conduire (cf. § 7) à l'estimation de la matrice du modèle AR estimé (qui est de Toeplitz) au sens du maximum de vraisemblance.

On peut remarquer la supériorité de cet algorithme sur celui de Durbin-Levinson lorsque des sources fortes sont présentes, cela s'explique par l'influence stabilisante du déterminant. Cette conclusion n'est plus vérifiée lorsque toutes les sources sont de faibles puissances relativement au bruit, mais là, le problème réside dans l'inadéquation du modèle AR.

L'auteur tient à remercier les experts de la revue pour leurs critiques qui ont permis d'éclaircir (l'espère-t-il) et de compléter de nombreux points.

Manuscrit reçu le 4 mars 1985, version révisée le 30 septembre 1985.

TABLEAU

J KAY 1 \*

NBRE CAPTEURS	= 16		
DIST. INTER CAPTEURS	= LAMBDA/2		
NBRE DIAG. BRUIT	= 1		
ORDRE DU MODELE	= 15		
NBRE RÉALISATIONS	= 100		
SOURCE 1	GIS=33.0	NIV= 3.0	
SOURCE 2	GIS=41.0	NIV= 2.0	
SOURCE 3	GIS=67.0	NIV=17.0	
SOURCE 4	GIS=76.0	NIV= 5.0	
SOURCE 5	GIS=90.0	NIV= 1.0	SIG=ECART TYPE
BRUIT	=3.00		
S/B	9.33		
BRUIT SIMULE ALGAUS	NMOY = 100		

MOYENNES ET ECART TYPE SUR : 100 REALISATIONS

SOURCE 1	GIS INIT =33.0		NIV INIT = 3.0	
KAY	MOY GIS =31.1	SIG GIS=0.24	MOY NIV = 3.1	SIG NIV=0.74
MAXENT	MOY GIS =33.0	SIG GIS=0.69	MOY NIV = 4.9	SIG NIV=6.07
SOURCE 2	GIS INIT =41.0		NIV INIT = 2.0	
KAY	MOY GIS =41.5	SIG GIS=0.64	MOY NIV = 1.8	SIG NIV=0.43
MAXENT	MOY GIS =41.3	SIG GIS=0.89	MOY NIV = 2.5	SIG NIV=4.15
SOURCE 3	GIS INIT =67.0		NIV INIT =17.0	
KAY	MOY GIS =66.5	SIG GIS=0.50	MOY NIV =17.0	SIG NIV=0.00
MAXENT	MOY GIS =67.0	SIG GIS=0.00	MOY NIV =14.4	SIG NIV=5.41
SOURCE 4	GIS INIT =76.0		NIV INIT = 5.0	
KAY	MOY GIS =77.0	SIG GIS=0.20	MOY NIV 7.9	SIG NIV=1.87
MAXENT	MOY GIS =76.0	SIG GIS=0.22	MOY NIV 2.3	SIG NIV=3.26
SOURCE 5	GIS INIT =90.0		NIV INIT = 1.0	
KAY	MOY GIS =90.9	SIG GIS=0.30	MOY NIV = 1.7	SIG NIV=0.37
MAXENT	MOY GIS =89.8	SIG GIS=0.88	MOY NIV = 1.0	SIG NIV=2.36

### BIBLIOGRAPHIE

- [1] A. ARCESE, On the method of Maximum Entropy Spectrum Estimation, *IEEE Trans. Inform. Theory*, IT-29, n° 1, janvier 1983, p. 161-164.
- [2] T. E. BARNARD, Two Maximum Entropy Beamforming Algorithms for Equally Spaced Line Arrays, *IEEE Trans. on acoustics speech, signal processing*, ASSP-30, n° 2, avril 1982, p. 175-189.
- [3] J. P. BURG, Maximum Entropy Spectral Analysis, *Ph. D. dissertation*, Dep. Geophys., Stanford University, Stanford, CA, May, 1975.
- [4] H. CLERGEOT, Estimation Spectraoe par le critère du Maximum de Vraisemblance pour un modèle AR avec pôles voisins du cercle unité, *Actes du neuvième colloque sur le traitement du signal et ses applications*, Nice, mai 1983.
- [5] H. CLERGEOT, *Thèse de Doctorat d'État*, octobre 1982, Orsay.
- [6] J. P. DUGRE, A. L. BEEX et L. L. SCHARF, Generating covariance sequences and the calculation of quantization and rounding error variances in digitals filters, *IEEE Trans. on ASSP*, ASSP-28, n° 1, février 1980, p. 102-104.
- [7] S. HAYKIN éd., *Non linear Methods of Spectral Analysis*, Springer-Verlag, 1983.
- [8] U. GRENANDER et SZEGO, *Toeplitz forms and their applications*, University of California Press, San Francisco, 1958.
- [9] S. M. KAY, Recursive Maximum Likelihood Estimation of Autoregressive Processes, *IEEE Trans on ASSP*, ASSP-31, n° 1, février 1983, p. 56-65.
- [10] J. P. LE CADRE et P. LOPEZ, Estimation d'une matrice interspectrale de structure imposée, *Traitement du Signal*, 1, n° 1, décembre 1984, p. 3-17.
- [11] O. N. STRAND, Multichannel Complex Maximum Entropy Spectral Analysis, *IEEE Trans. Autom. Control*, AC-22, n° 4, août 1977, p. - .