

Détection multiple par les valeurs propres de la matrice spectrale

Multiple detection using the eigenvalues of the spectral matrix



Idir TAS

CEPHAG-ENSIEG, B. P. n° 46, 38402 SAINT-MARTIN-D'HÈRES
CEDEX

Né le 14 mai 1960 à Alger. Diplôme d'études supérieures en Physique Électronique de l'Université de Constantine en 1983, DEA en Automatique et Traitement du Signal de l'INPG en 1984, prépare actuellement une thèse de Doctorat à l'INPG. Ses travaux de recherche intéressent l'analyse spectrale et le traitement d'antenne en écoute passive, en particulier, la détection et la localisation de sources émettrices.



Claudine LATOMBE

CEPHAG, UA 346, B. P. n° 46, 38402 SAINT-MARTIN-D'HÈRES

Claudine Latombe est ancienne élève de l'ENSET, docteur de 3^e cycle en Informatique (1974), docteur d'État en Traitement de Signal à l'INPG (1982). Elle enseigne à l'INPG depuis 1970; son travail de recherche porte sur les problèmes d'analyse spectrale et les traitements multicapteurs.

RÉSUMÉ

Nous traitons ici le problème de la détermination du nombre d'excitations non corrélées en écoute passive, problème qui se pose en acoustique sous-marine, en traitement d'antenne ou en sismique.

Nous utilisons quatre critères de détection : certains connus de type AIC et MDL avec pour mesure d'information l'information directe de Kullback; nous les prolongeons en utilisant l'information de Kullback inverse. Nous adaptons également un nouveau critère qui utilise la loi du logarithme du rapport de vraisemblance (c'est une loi du chi-deux) et nous proposons aussi un critère à seuil avec les valeurs propres de la matrice spectrale.

Dans cette étude, nous étudions et comparons les performances de ces critères sur des simulations réalistes. La première est issue des situations rencontrées en traitement d'antenne, et la deuxième, des situations rencontrées en sismique.

Sur ces simulations nous abordons également l'étude de la robustesse de ces critères lorsque l'hypothèse classique de bruits non corrélés et de même puissance sur tous les capteurs, n'est plus vérifiée. Nous dégageons ainsi les limites d'application de ces critères.

MOTS CLÉS

Analyse spectrale multidimensionnelle, vraisemblance, détection, mesure d'information.

SUMMARY

In this study we treat the problem of detecting from multidimensional data, the number of uncorrelated signals in passive array treatment as it is the case in underwater acoustics, array processing and seismology.

We use four detection criteria. Some of them are known, like AIC and MDL criteria where direct Kullback's divergence is the information measure; we prolong them using the inverse Kullback's divergence. We also adapt a new criterion using the logarithm of the likelihood ratio that has a chi square distribution and we suggest a simplified threshold criterion that uses the eigenvalues of the spectral matrix of the data.

We study and compare the performances of these criteria in realistic simulations. The first one is inspired by the problems of array processing and the second one by seismic problems.

Finally we study the robustness of these criteria when the classical hypothesis of uncorrelated noises having equal variances is not fulfilled. Thus we outline some application limits of these criteria.

KEY WORDS

Multidimensional spectral analysis, likelihood, detection, information measure.

TABLE DES MATIÈRES

- 1. Introduction**
- 2. Formulation du problème**
- 3. Généralités**
 - 3.1. Fonction de vraisemblance
 - 3.2. Rapport de vraisemblance généralisé
- 4. Critères de détection**
 - 4.1. Test du chi-deux
 - 4.2. Critère AIC
 - 4.3. Critère MDL
 - 4.4. Critère simplifié à seuil
- 5. Résultats sur des simulations**
 - 5.1. Traitement d'antennes
 - 5.2. Fronts sismiques
 - 5.3. Résultats et commentaires
- 6. Conclusion**
- Bibliographie**

sources. Dans ce cas, ces critères ont tendance à sous-estimer le nombre de sources.

D'autres critères proposés pour la détermination de l'ordre d'un modèle (11, 12 et 14) ont été appliqués pour la détermination du nombre de sources [9, 15, 17 et 18]. Ces critères (AIC et MDL) formulés théoriquement pour le cas asymptotique ($B_e \cdot T$ très grand) doivent très souvent, pour des signaux réels, être utilisés avec des $B_e \cdot T$ faibles. Leurs résultats, dans ces cas réels ou sur des signaux simulés restent insatisfaisants [15] (nombre de sources détectées trop grand). Afin d'améliorer les résultats des critères AIC et MDL, d'autres mesures d'information autre que l'information directe de Kullback ont été utilisées [9]. Un autre critère peut être construit en utilisant la loi du logarithme du rapport de vraisemblance qui est une loi de chi-deux.

Une formulation générale du problème est donnée au paragraphe 2, les notions théoriques utilisées sont rappelées au paragraphe 3, les critères précis sont développés au paragraphe 4 et des résultats de simulation commentés sont présentés au paragraphe 5.

2. Formulation de problème

Dans les situations d'écoute passive (radar, sonar, sismologie), où on ne dispose que de peu d'information *a priori*, un nombre inconnu P d'excitations non corrélées traversent un milieu de propagation et sont reçues sur une antenne composée de N capteurs séparément accessibles (P inférieur à N).

A la sortie des capteurs les N fonctions du temps $r_i(t)$, $i \in [1, N]$, qui sont généralement des tensions électriques contiennent toute l'information disponible sur le milieu et sur les sources.

Chaque fonction $r_i(t)$ est considérée comme une réalisation d'une fonction aléatoire et est un mélange additif des contributions, supposées centrées, de chaque source et d'un bruit (parasite) dû au milieu de propagation et/ou au système de réception.

1. Introduction

Le problème de la détection du nombre de sources en écoute passive et en large bande reste toujours posé malgré de nombreux travaux effectués récemment dans ce domaine.

L'utilisation de la matrice interspectrale des signaux enregistrés a conduit à l'utilisation de nombreux critères à seuil utilisant les valeurs propres de cette matrice spectrale [3 et 8]. Bien que ces critères restent assez largement utilisés et donnent souvent des résultats acceptables, un problème se pose lorsque l'énergie d'une des sources est grande devant celle des autres

Les hypothèses habituellement faites sur le milieu et les sources sont les suivantes [3, 9, 15, 17 et 18] :

- (i) milieu linéaire, homogène et isotrope, ce qui entraîne que chaque front d'onde est spatialement cohérent;
- (ii) signaux émis non corrélés (ou partiellement) entre eux et non corrélés avec les bruits.

En passant dans le domaine fréquentiel et en large bande, toute l'information sur les signaux et sur les sources est inchangée. Cependant, en ne considérant que la matrice spectrale des signaux $r_i(t)$, $i \in [1, N]$, nous perdons de l'information, en particulier l'information de phase. Ceci n'a pas d'influence sur le problème étudié de la détection du nombre P de sources.

Par suite des décorrélations des signaux émis et des bruits, la matrice spectrale $\gamma(f)$ des signaux $r_i(t)$, $i \in [1, N]$, se décompose comme suit :

$$(1) \quad \gamma(f) = \gamma_s(f) + \gamma_B(f)$$

où $\gamma_s(f)$ est la matrice spectrale des sources seules (elle est de rang P , vu les hypothèses faites) et $\gamma_B(f)$ est la matrice spectrale des bruits.

L'hypothèse classique sur les bruits est de considérer qu'ils ont même puissance sur tous les capteurs et sont non corrélés entre eux. La matrice $\gamma_B(f)$ s'écrit alors :

$$(2) \quad \gamma_B(f) = \beta(f) \cdot \mathbb{I}_N$$

Où \mathbb{I}_N est la matrice identité et $\beta(f)$ est la densité spectrale de puissance des bruits.

Nous cherchons ici à évaluer dans des situations réalistes (signaux large bande et $B_e \cdot T$ faibles) les performances de certains algorithmes détectant le nombre P de sources. Nous abordons également l'étude de leur robustesse lorsque l'hypothèse de bruits de même puissance et/ou non corrélés n'est plus vérifiée.

3. Généralités

3.1. FONCTION DE VRAISEMBLANCE

Nous rappelons des notions qui, bien que connues, sont nécessaires pour l'établissement des notations.

En écoute passive, nous ne disposons que de peu d'informations *a priori*. Nous supposons cependant que les signaux sortant des capteurs sont gaussiens et centrés.

Les signaux enregistrés sont représentés par le vecteur $\mathbf{r}(t)$, $\mathbf{r}(t) = [r_1(t), \dots, r_N(t)]^T$, où T désigne le transposé d'un vecteur. $\mathbf{r}(t)$ est supposé être une réalisation d'un processus stationnaire, gaussien et centré. $\mathbf{R}(f)$, vecteur dont les composantes sont les transformées de Fourier des $r_i(t)$, est alors approximativement gaussien, complexe et centré.

La matrice spectrale $\gamma(f)$ des observations peut être estimée par :

$$(3) \quad \hat{\gamma}(f) = \frac{1}{K} \sum_{k=1}^K \mathbf{R}_k(f) \cdot \mathbf{R}_k^+(f)$$

Où $+$ désigne le vecteur complexe conjugué.

K représente le nombre de réalisations indépendantes $\mathbf{R}_k(f)$ moyennées pour estimer $\gamma(f)$. Dans la méthode du périodogramme moyenné [6], K représente le nombre de périodogrammes utilisés. Dans la méthode du périodogramme lissé [6], K est le produit « bande équivalente d'analyse » par le « temps d'intégration du traitement ».

En supposant que l'observation $\mathbf{R}(f)$ est une variable aléatoire gaussienne complexe et centrée (cf. [20]) (c'est-à-dire : $E\{\mathbf{R}(f)\mathbf{R}^+(f)\} = 0$), et que les réalisations $\mathbf{R}_k(f)$ sont indépendantes, la log-vraisemblance de $\mathbf{R}(f)$ s'écrit :

$$(4) \quad \log[V(\mathbf{R})] = -KN \log(\pi) - K \log |\gamma(f)| - K \text{Trace} [\hat{\gamma}(f) \cdot \gamma^{-1}(f)]$$

$|\gamma(f)|$ est le déterminant de $\gamma(f)$ et $\gamma^{-1}(f)$ son inverse.

L'estimation $\hat{\gamma}(f)$ de $\gamma(f)$ est supposée sans biais.

Comme les bruits et les signaux sont non corrélés entre eux, nous avons :

$$\gamma(f) = \gamma_s(f) + \gamma_B(f)$$

$\gamma_s(f)$ est la matrice spectrale des signaux qui seraient reçus en l'absence de bruits et $\gamma_B(f)$ celle des bruits.

La matrice $\gamma_s(f)$ est déterminée par sa décomposition spectrale qui, lorsqu'il y a P sources, s'écrit :

$$(5) \quad \begin{cases} \gamma_s(f) = \sum_{i=1}^P \alpha_i(f) \cdot \mathcal{V}_i(f) \cdot \mathcal{V}_i^+(f) \\ \mathcal{V}_i(f) \perp \mathcal{V}_j(f) \quad \text{et} \quad \|\mathcal{V}_i(f)\| = 1 \end{cases}$$

où $(\alpha_i(f), \mathcal{V}_i(f))_{i=1, P}$ sont respectivement les valeurs et les vecteurs propres de $\gamma_s(f)$.

Dans le cas de bruits non corrélés et de même puissance, $\gamma_B(f) = \beta(f) \mathbb{I}_N$, la maximisation de la log-vraisemblance donne les estimations $\hat{\alpha}_i(f)$ et $\hat{\beta}_P(f)$ au sens du maximum vraisemblance :

$$(6) \quad \begin{cases} \hat{\alpha}_i(f) = \hat{\lambda}_i(f) - \hat{\beta}_P(f) \\ \text{et} \\ \mathcal{V}_i(f) = \hat{\mathcal{V}}_i(f), \quad i \in [1, P] \end{cases}$$

$$(7) \quad \hat{\beta}_P(f) = \frac{1}{N-P} \sum_{i=P+1}^N \hat{\lambda}_i(f)$$

$(\hat{\lambda}_i(f), \hat{\mathcal{V}}_i(f))_{i=1, N}$ sont les valeurs et vecteurs propres de $\hat{\gamma}(f)$.

Les $\hat{\lambda}_i(f)$ étant rangées par valeurs décroissantes, $\hat{\beta}_P(f)$ est la moyenne arithmétique des $(N-P)$ plus petites valeurs propres de $\hat{\gamma}(f)$.

Le maximum de la log-vraisemblance devient (cf. [3]) :

$$(8) \quad \text{Max} [\log V(\mathbf{R}, f) | P] \\ = -\log \left[\pi^N \cdot \prod_{i=1}^P \hat{\lambda}_i(f) \cdot [\hat{\beta}_P(f)]^{N-P} \right]^K$$

qui s'écrit aussi :

$$(9) \quad \text{Max} [\log V(\mathbf{R}, f) | P] \\ = -K \cdot \left[(N-P) \log \hat{\beta}_P(f) - \sum_{i=P+1}^N \log(\hat{\lambda}_i(f)) \right] + C_1(f)$$

Où $C_1(f)$ est indépendant de P .
Le nombre positif

$$(N-P) \log \hat{\beta}_P(f) - \sum_{i=P+1}^N \log \hat{\lambda}_i(f)$$

peut s'interpréter comme une mesure de proximité des P plus petites valeurs propres de $\hat{\gamma}(f)$ et est à la base des tests d'Akaike et de Rissanen (cf. §4).

3. 2. RAPPORT DE VRAISEMBLANCE GÉNÉRALISÉ

Nous introduisons ici le rapport de vraisemblance généralisé de l'observation $\mathbf{R}(f)$. Ce rapport servira d'une part à exprimer les critères AIC et MDL et d'autre part à élaborer le critère du test du chi-deux (cf. §4).

L'observation $\mathbf{R}(f)$ dépend d'un ensemble de paramètres ϵ_P .

Afin de tester l'hypothèse H_p , contre l'hypothèse H_q , nous utilisons le rapport de vraisemblance généralisé suivant [3 et 16] :

$$A(P, f) = \frac{\text{MAX}_{H_p} [V(\mathbf{R}(f) | \epsilon_p)]}{\text{MAX}_{H_q} [V(\mathbf{R}(f) | \epsilon_q)]}$$

$V(\mathbf{R}(f) | \epsilon_p)$ est la fonction de vraisemblance conditionnelle de $\mathbf{R}(f)$ et ϵ_p représente le vecteur des paramètres du modèle sous l'hypothèse H_p .

ϵ_p contient dans le cas de P sources, les P plus grandes valeurs propres de $\gamma(f)$, les P vecteurs propres correspondants et la puissance du bruit; il s'écrit sous la forme suivante :

$$(10) \quad \epsilon_p = [\lambda_1, \dots, \lambda_p, \mathcal{V}_1^T, \dots, \mathcal{V}_P^T, \beta_p]^T$$

Partant de $P=0$ et en augmentant les valeurs de P jusqu'à $N-1$, on teste pour chaque valeur de P les hypothèses suivantes :

- Hypothèse H_p : présence de P sources au plus.
- Hypothèse H_q : présence de plus de P sources.

D'après [3], la fonction $\text{MAX} [V(\mathbf{R}(f) | \epsilon_q)]$ est monotone croissante en q , nous avons alors :

$$\text{MAX}_{q \leq P} [V(\mathbf{R}(f) | \epsilon_q)] = V(\mathbf{R}(f) | \epsilon_P)$$

et

$$\text{MAX}_{P < q \leq N-1} [V(\mathbf{R}(f) | \epsilon_q)] = V(\mathbf{R}(f) | \epsilon_{N-1})$$

En utilisant l'expression explicite de la log vraisemblance [formule (8)] en fonction des valeurs propres de $\hat{\gamma}(f)$ rangées par valeurs décroissantes, $A(P, f)$ s'écrit :

$$(11) \quad A(P, f) = \left[\frac{\prod_{i=P+1}^N \hat{\lambda}_i(f)}{(\hat{\beta}_P(f))^{N-P}} \right]^K$$

où

$$\hat{\beta}_P(f) = \frac{1}{N-P} \sum_{i=P+1}^N \hat{\lambda}_i(f)$$

4. Critères de détection

Avant de caractériser le sous-espace source, les paramètres des sources et le bruit ambiant, il faut estimer le nombre de sources. Nous présentons ici quatre critères de détection qui utilisent soit la vraisemblance de $\mathbf{R}(f)$ soit le rapport de vraisemblance généralisé, soit comparant les valeurs propres de la matrice de covariance $\hat{\gamma}(f)$ de $\mathbf{R}(f)$ à un seuil.

4. 1. TEST DU CHI-DEUX

En utilisant le rapport de vraisemblance généralisé $A(P, f)$ [expression (11) du paragraphe 3.2], un critère de détection serait alors :

$$A(P, f) \underset{H_p}{\overset{H_q}{\geq}} \alpha_p$$

α_p est un seuil à déterminer défini par :

$$\text{Proba} [A(P, f) > \alpha_p | H_p] = f_a$$

f_a est la probabilité de fausse alarme.

La loi de probabilité exacte du rapport $A(P, f)$ n'est pas connue et le problème de la détermination du seuil n'est pas résolu. En pratique, il est possible de fixer les seuils α_p par simulation mais ce travail manque de souplesse et de simplicité.

Asymptotiquement, lorsque K tend vers l'infini, d'après la théorie générale des rapports de vraisemblance [5], la statistique $-2 \log [A(P, f)]$ suit une loi de chi-deux à $(N-P)^2 - 1$ degrés de liberté.

En pratique K est petit, il est donc préférable [7] d'utiliser une statistique modifiée $X(P, f)$ qui s'écrit :

$$(12) \quad X(P, f) = C(P) [-2 \log A(P, f)]$$

où la fonction $C(P)$ vaut :

$$(13) \quad C(P) = \left[1 - \frac{P}{K} - \frac{2(N-P)^2 + 1}{6K(N-1)} \right]$$

Lorsque K tend vers l'infini (cas asymptotique) $C(P) = 1$, on retrouve ainsi la statistique $-2 \log A(P, f)$ issue de la théorie générale.

D'après [5] si la statistique $-2 \log A(P, f)$ suit une loi du chi-deux à $(N-P)^2 - 1$ degrés de liberté, la statistique $X(P, f)$ suit alors approximativement la même loi du chi-deux.

Un critère de détection est alors le suivant :

$$(14) \quad X(P, f) \underset{H_p}{\overset{H_q}{\geq}} \chi_p^2$$

χ_p^2 est un seuil à déterminer défini par :

$$\text{Proba} [X(P, f) > \chi_p^2 | H_p] = \eta_p$$

η_p est le niveau de « signification » du test, c'est le risque de rejeter H_p à tort et il s'écrit [2] :

$$(15) \quad \eta_p = \frac{1}{2^{r/2} \cdot \Gamma(r/2)} \int_{\chi_p^2}^{\infty} (x)^{-1+r/2} \exp\left(-\frac{x}{2}\right) dx$$

r est le nombre de degrés de liberté, $r = (N-P)^2 - 1$ et Γ la fonction gamma.

On définit la probabilité complémentaire $P_d = 1 - \eta_p$ qui représente la probabilité d'acceptation de H_p sous l'hypothèse H_p . Cette probabilité sera appelée par la suite « probabilité de détection ».

A une fréquence donnée, en fixant les niveaux de signification η_p du test ou les probabilités de détection P_d , on calcule les seuils χ_p^2 correspondants et on les compare à la statistique $X(P, f)$.

Si $X(P, f)$ est supérieure à χ_p^2 , l'hypothèse H_p , présence de P sources au plus, est rejetée.

En suivant la démarche présentée dans [19], les hypothèses H_p ($p \in [0, N-2]$) sont testées séquentiellement pour un niveau de signification donné et la valeur de P pour laquelle H_p est la première acceptée est prise comme une estimée du nombre de sources à cette fréquence.

Nous avons éliminé la valeur de P qui vaut $(N-1)$ car pour cette valeur le nombre de degrés de liberté de la statistique $X(P, f)$ est nul ($r = (N-P)^2 - 1$). Le test ne peut être utilisé dans ce cas-là.

Ce traitement est ainsi appliqué à toutes les fréquences. Il peut arriver que, pour un niveau de signification donné, toutes les hypothèses H_p ($p \in [0, N-2]$)

soient rejetées. Ce cas doit être interprété soit comme un mauvais choix du modèle, soit qu'il y ait $(N-1)$ sources.

4.2. CRITÈRE AIC (AKAIKE INFORMATION CRITERION)

Ce critère a été proposé par H. Akaike [1] en 1974 pour la détermination de l'ordre d'un modèle. On peut l'appliquer pour la détection du nombre de sources (cf. [18]) en minimisant par rapport à P , la quantité

$$(16) \quad \text{AIC}(P) = -\text{MAX} [\log\text{-vraisemblance}] + \text{NPL}$$

NPL est le nombre de paramètres libres du modèle représenté ici par le vecteur ε_p (§ 3.2). Il vaut $P(2N-P) + 1$. En négligeant le terme indépendant de P : $\text{NPL} = P(2N-P)$.

En utilisant l'expression explicite du maximum de la log-vraisemblance, nous estimons le nombre de sources, en minimisant par rapport à P , l'expression suivante :

$$(17) \quad \text{AIC1}(P, f) = -\log \left[\frac{\prod_{i=P+1}^N \hat{\lambda}_i(f)}{\left(\frac{1}{N-P} \sum_{i=P+1}^N \hat{\lambda}_i(f) \right)^{N-P}} \right]^K + P(2N-P)$$

On peut constater que ce critère peut aussi s'écrire en faisant intervenir le rapport de vraisemblance généralisé $A(P, f)$:

$$\text{AIC1}(P, f) = -\log [A(P, f)] + P(2N-P)$$

L'information de Kullback [4, 1], pour deux vecteurs aléatoires X et Y de lois de probabilité $f(X)$ et $g(Y)$ est définie par :

$$I(f, g) = \int \log \left(\frac{dg}{df} \right) dg$$

Cette information est positive ou nulle. Elle est nulle si et seulement si $f(X)$ et $g(Y)$ sont égales presque partout. Elle n'est pas symétrique car $I(f, g) \neq I(g, f)$. Soient V_{ε_p} et V_{ε_q} les lois de probabilité des observations $R(f/\varepsilon_p)$ et $R(f/\varepsilon_q)$. Ces lois sont supposées gaussiennes complexes, centrées et de matrice de covariance $\gamma(f)$ et $\hat{\gamma}(f)$.

En utilisant l'information de Kullback $I(V_{\varepsilon_p}, V_{\varepsilon_q})$, Clergeot [21] et Lecadre [9] ont montré que le critère AIC s'écrit sous la forme :

$$(18) \quad \text{AIC}(P) = \text{Min} [I(V_{\varepsilon_p}, V_{\varepsilon_q})] + \text{NPL}$$

Comme la mesure d'information est un moyen d'évaluer la proximité entre deux lois de probabilité, on peut utiliser (du moins formellement) une autre mesure d'information pour évaluer la proximité entre les deux lois V_{ε_p} et V_{ε_q} .

En utilisant l'information de Kullback inverse $I(V \varepsilon_p, V \varepsilon_p)$, on obtient alors (cf. [9]) un autre critère, noté AIC2(P, f) :

$$(19) \text{ AIC2}(P, f) = \log \left[\frac{\prod_{i=P+1}^N \hat{\lambda}_i(f)}{\left((N-P) \left(\sum_{i=P+1}^N \hat{\lambda}_i^{-1}(f) \right)^{-1} \right)^{N-P}} \right]^K + P(2N-P)$$

L'information de Kullback n'étant pas symétrique, les deux expressions du critère AIC (AIC1 et AIC2) sont différentes. Outre l'opposition de signe du premier terme, nous constatons que :

— Pour AIC1 l'estimation de la puissance de bruit est la moyenne arithmétique des (N-P) plus petites valeurs propres de $\hat{\gamma}(f)$:

$$\hat{\beta}_p(f) = \frac{1}{N-P} \sum_{i=P+1}^N \hat{\lambda}_i(f)$$

— Pour AIC2 l'estimation de la puissance du bruit est la moyenne harmonique des (N-P) plus petites valeurs propres de $\hat{\gamma}(f)$.

4.3. CRITÈRE MDL (MINIMUM DESCRIPTION LENGTH)

Inspirés par les travaux de H. Akaike, J. Rissanen et G. Schwartz [11, 12 et 14] ont approché le problème de la détermination de l'ordre d'un modèle de deux points de vue différents mais aboutissant à une même forme du critère.

Ce critère peut être appliqué à la détection du nombre de sources et il s'écrit :

$$(20) \text{ MDL}(P) = -\text{MAX} [\text{Log-vraisemblance}] + \frac{1}{2} \text{NPL}(\log K)$$

Par extension, en suivant [9], on peut le mettre sous la forme suivante :

$$\text{MDL} = \text{Mesure d'information} + \frac{1}{2} \text{NPL}(\log K)$$

Le critère MDL, expression (20), diffère du critère AIC, expression (16), par le deuxième terme qui est introduit. Le critère MDL est un critère consistant, c'est-à-dire qui converge vers le rang exact P du modèle quand K devient grand.

Ce critère utilisé avec l'information de Kullback directe $I(V \varepsilon_p, V \varepsilon_p)$ conduit à l'expression suivante, symétrique de l'expression (17) :

$$(21) \text{ MDL1}(P, f) = -\log \left[\frac{\prod_{i=P+1}^N \hat{\lambda}_i(f)}{\left(\frac{1}{N-P} \sum_{i=P+1}^N \hat{\lambda}_i(f) \right)^{N-P}} \right]^K + \frac{1}{2} P(2N-P) \log K$$

D'autre part, utilisé avec l'information de Kullback inverse $I(V \varepsilon_p, V \varepsilon_p)$, ce critère a l'expression suivante, symétrique de l'expression (19) :

$$(22) \text{ MDL2}(P, f) = +\log \left[\frac{\prod_{i=P+1}^N \hat{\lambda}_i(f)}{\left((N-P) \left(\sum_{i=P+1}^N \hat{\lambda}_i^{-1}(f) \right)^{-1} \right)^{N-P}} \right]^K + \frac{1}{2} P(2N-P) \log K$$

A chaque fréquence, le nombre de sources est donné par la valeur de P, $P \in [0, N-1]$, pour laquelle les critères AIC et MDL [expressions (18), (19), (21) et (22)] sont minimaux.

4.4. CRITÈRE SIMPLIFIÉ A SEUIL (NOTÉ CSS)

Dans ce critère simplifié, à une fréquence donnée, le nombre de sources est égal au nombre de valeurs propres de la matrice $\hat{\gamma}(f)$ supérieures à un certain seuil S qui devrait estimer le niveau de bruit.

L'estimation de la puissance du bruit au sens du maximum de vraisemblance, lorsque P est connu, est :

$$\hat{\beta}_p(f) = \frac{1}{N-P} \sum_{i=P+1}^N \hat{\lambda}_i(f)$$

où $\hat{\lambda}_i(f)$ sont les valeurs propres de $\hat{\gamma}(f)$ rangées par valeurs décroissantes.

En l'absence de sources ($P=0$), $\beta_0(f)$ représente la puissance du bruit à la fréquence f. C'est pourquoi nous proposons de choisir pour seuil S :

$$(23) S = \frac{1}{N} \text{Max}_{f_j} \left[\frac{1}{N} \sum_{i=1}^N \hat{\lambda}_i(f_j) \right] = \frac{1}{N} \text{Max}_{f_j} [\beta_0(f_j)]$$

S représente la puissance maximale de bruit par capteur et en absence de sources.

Des seuils plus généraux et plus difficiles à mettre en œuvre ont été proposés [3 et 8].

5. Résultats sur des simulations

Nous présentons ici les résultats obtenus avec les six critères annoncés (AIC1, AIC2, MDL1, MDL2, CSS et test du chi-deux) sur des simulations réalistes, proches des situations réelles rencontrées soit en traitement d'antennes passives, soit en sismique.

Pour chaque simulation, les signaux sont échantillonnés à la période de 4 ms. La matrice spectrale est estimée par la méthode du périodogramme lissé avec une fonction de pondération de Tuckey caractérisée par un produit bande équivalente d'analyse et temps d'intégration du traitement de 12,56.

Si nous utilisons plusieurs réalisations de bruit, nous calculons une matrice spectrale moyenne qui s'écrit :

$$\hat{\gamma}(f) = \frac{1}{M} \sum_{k=1}^M \hat{\gamma}^k(f)$$

M est le nombre de moyennes et $\hat{\gamma}^k(f)$ la matrice spectrale estimée pour la k -ième réalisation de bruit.

Nous calculons les valeurs et les vecteurs propres de la matrice $\hat{\gamma}(f)$ et nous appliquons les critères de détection ci-dessus (§4) pour estimer le nombre de sources.

5. 1. TRAITEMENT D'ANTENNE

Nous avons simulé une situation pouvant se rencontrer en traitement d'antenne avec des sources bande étroite.

L'hypothèse d'ondes planes n'est pas nécessaire pour les critères de détection que nous utilisons, nous l'avons faite afin de simplifier la simulation.

Les signaux émis par les trois sources s'écrivent :

$$S_1(t) = A_1 \cos [2 \pi f_1 (1 - m_1 t) t + \varphi_1]$$

$$S_2(t) = \sum_{i=2}^3 A_i \cos [2 \pi f_{i-1} (1 - m_i t) t + \varphi_i]$$

$$S_3(t) = \sum_{i=4}^6 A_i \cos [2 \pi f_{i-3} (1 - m_i t) t + \varphi_i]$$

Les A_i , $i \in [1, 6]$, sont des amplitudes certaines.

Les φ_i , $i \in [1, 6]$, sont des phases aléatoires équiparties sur $[0, 2 \pi]$.

Les f_i , $i \in [1, 3]$, sont des fréquences pures fixées aux valeurs suivantes : $f_1 = 32$ Hz, $f_2 = 64$ Hz et $f_3 = 96$ Hz.

Les m_i , $i \in [1, 6]$, sont des modulations linéaires fixes, choisies pour simuler des sources à bande étroite et non corrélées. Les signaux simulés sont de fait non stationnaires, mais vu les valeurs de modulation utilisées ($m_i < 10^{-1}$) et la durée $T = 1$ s des réalisations, ces signaux peuvent être considérés comme quasi stationnaires sur la durée T .

La disposition des sources et des capteurs est représentée sur la figure 1.

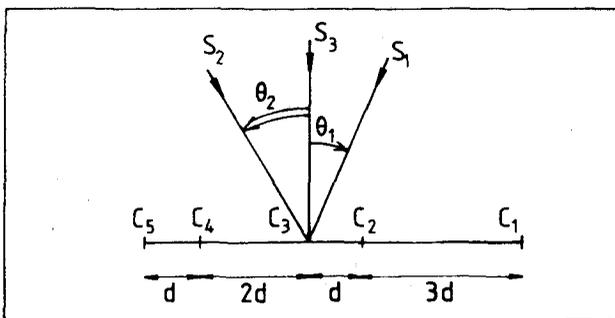


Fig. 1

Les signaux reçus sont des ondes planes au niveau des capteurs, additivement bruitées qui s'écrivent :

$$r_i(t) = \sum_{j=1}^3 S_j(t - \tau_j^i) + a_i \cdot N_i(t), \quad i \in [1, 5]$$

Les τ_j^i sont des retards de propagation et a_i des constantes.

Les $N_i(t)$ sont des réalisations indépendantes d'un bruit blanc gaussien réduit.

Pour avoir des bruits corrélés sur les voies k et j , nous remplaçons $B_k(t) = a_k \cdot N_k(t)$ par la combinaison linéaire suivante :

$$B_k(t) = a_k \cdot N_k(t) + C \cdot a_j \cdot N_j(t)$$

C représente le coefficient de corrélation des voies k et j . L'élément (k, j) de la matrice spectrale des bruits ne sera pas nul. Pour C donné, les bruits seront de même puissance sur tous les capteurs si le coefficient a_k vaut :

$$a_k = (1 - C^2)^{1/2} \cdot a_j$$

L'expérience simulée devrait donc pouvoir conduire aux conclusions suivantes : trois sources autour de 32 Hz, deux sources autour de 64 Hz et une source autour de 96 Hz.

5. 2. FRONTS SISMIQUES

Nous avons cherché à simuler une situation pouvant se rencontrer en sismique où il s'agit de séparer deux sources dont les fronts d'ondes non nécessairement plans arrivent sur un réseau de capteurs avec des retards très proches. C'est le cas par exemple de séparation de multiples et d'ondes réfléchies ou réfractées traversant une discontinuité du sous-sol caractérisée par une différence d'impédance (densité et vitesse de propagation).

Pour cela nous avons simulé deux sources émettant des signaux de la forme :

$$S_i(t) = \sum_{j=1}^2 A_j^i \cos [2 \pi f_j (1 - m_j t) t + \varphi_j^i], \quad i \in [1, 2]$$

Pour la définition des paramètres A_j^i , m_j^i et φ_j^i , cf. §5. 1. Nous avons pris :

$$f_1 = 32 \text{ Hz} \quad \text{et} \quad f_2 = 86 \text{ Hz}$$

$S_1(t)$ est une onde plane dans la direction de l'antenne.

$S_2(t)$ est une onde hyperbolique.

La disposition des fronts d'ondes et des capteurs est représentée figure 2.

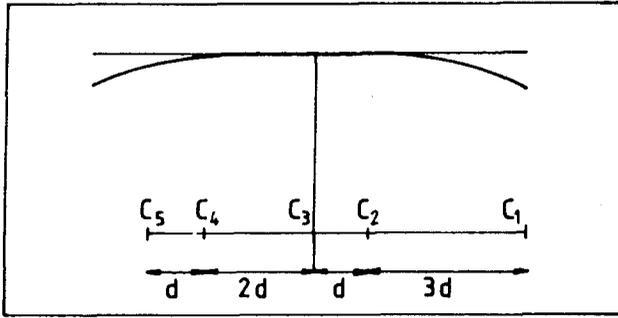


Fig. 2

Les signaux reçus peuvent se mettre sous la forme :

$$r_i(t) = S_1(t) + S_2(t + \delta k_i T_E) + a_i N_i(t),$$

$$i \in [1, 5], \quad \delta = \text{constante} (\delta \leq 1)$$

L'onde $S_2(t)$ arrive sur le réseau de capteurs avec des retards de propagation inférieurs à la période d'échantillonnage T_E ($T_E = 4$ ms).

Nous avons pris : $k_1 = 1, k_2 = 1/4, k_3 = 0, k_4 = 1/2$ et $k_5 = 3/4$.

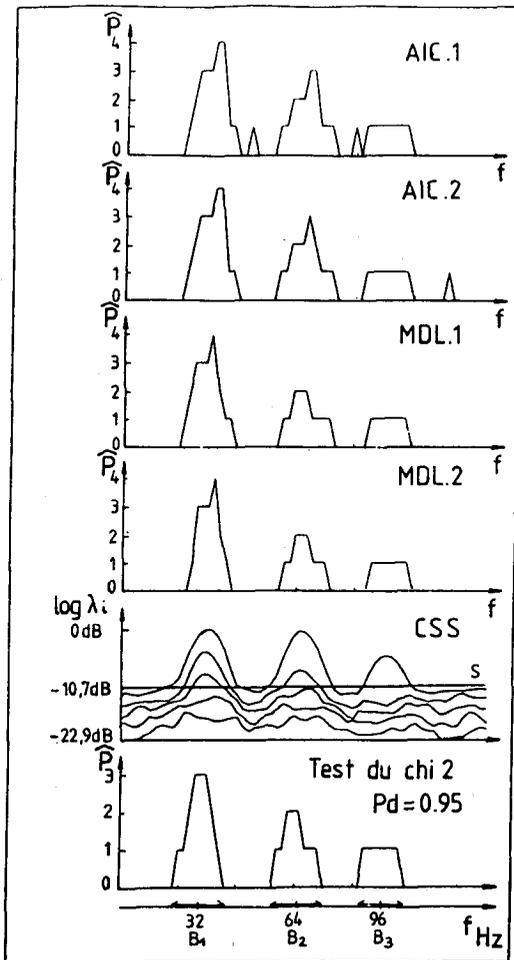


Fig. 3. — Première simulation. Bruits non corrélés et de même puissance sur les capteurs. $M=1$ et $K=12,56$.

5. 3. RÉSULTATS ET COMMENTAIRES

Notre but n'est pas ici de comparer les différents critères sur un plan statistique, mais plutôt d'étudier leur comportement sur un petit nombre d'épreuves puisqu'en pratique, il est souvent demandé de conclure à partir d'un très petit nombre de réalisations. Nous avons cependant, dans la première simulation, effectué un traitement statistique portant sur 110 réalisations.

Lorsque le test du chi-deux est appliqué, la quantité $C(P)$ varie, dans les situations décrites, entre 0,82 et 0,73.

5. 3. 1. Première simulation

Les paramètres du signal sont fixés aux valeurs suivantes : $A_1 = 20, A_2 = 25, A_3 = 20, A_4 = 30, A_5 = 25, A_6 = 20$; $m_1 = 0,06, m_2 = 0,09, m_3 = 0,03, m_4 = 0,12, m_5 = 0,04$ et $m_6 = 0,05$.

Les caractéristiques bruit varient par l'intermédiaire des coefficients a_i .

(a) Un cas de bruits non corrélés et de même puissance sur tous les capteurs ($a_i = 34$ pour $i = 1, 5$) est présenté figure 3. On utilise une seule réalisation de bruit

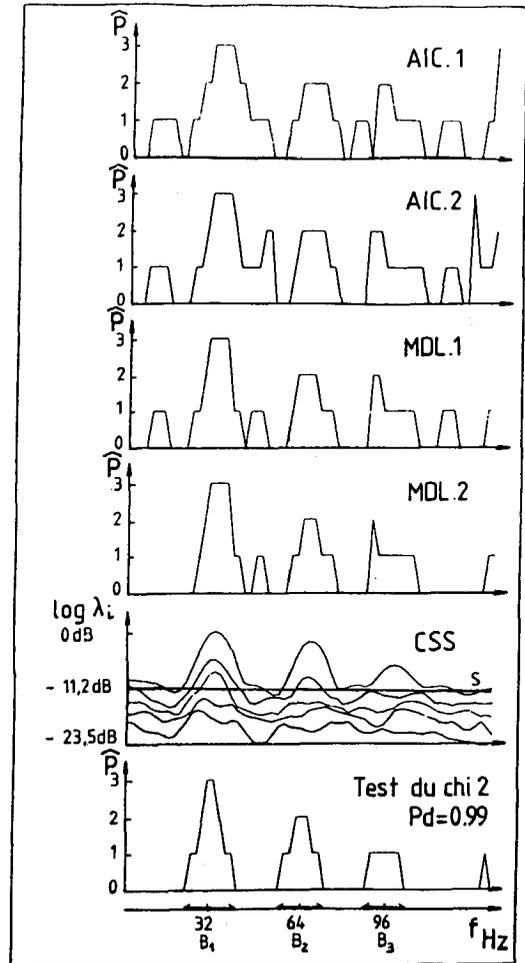


Fig. 4. — Première simulation. Bruits non corrélés et de puissance différentes sur les capteurs (trois voies modifiées). $M=1$ et $K=12,56$.

($K=12,56$ et $M=1$). Le rapport signal à bruit varie de 5 dB à -1 dB selon les voies considérées.

Les critères MDL1 et MDL2 donnent des sources parasites mais cependant moins que AIC1 et AIC2.

Le CSS et le test du chi-deux avec une probabilité de détection de 0,95 donnent les résultats escomptés.

(b) Figure 4, les bruits sont non corrélés et de puissance différentes sur les capteurs (trois voies modifiées). $K=12,56$ et $M=1$. Les pondérations des bruits sont : $a_1=30$, $a_3=40$, $a_5=45$ et $a_2=a_4=34$.

Pour une probabilité de détection de 0,95, le test du chi-deux donne des résultats semblables à MDL2. Si $P_d=0,99$ le test du chi-deux est meilleur. Ceci montre l'influence de la probabilité de détection pour le test du chi-deux. Si P_d est faible, ce critère donne des sources parasites. Les autres critères donnent plusieurs sources parasites à certaines fréquences

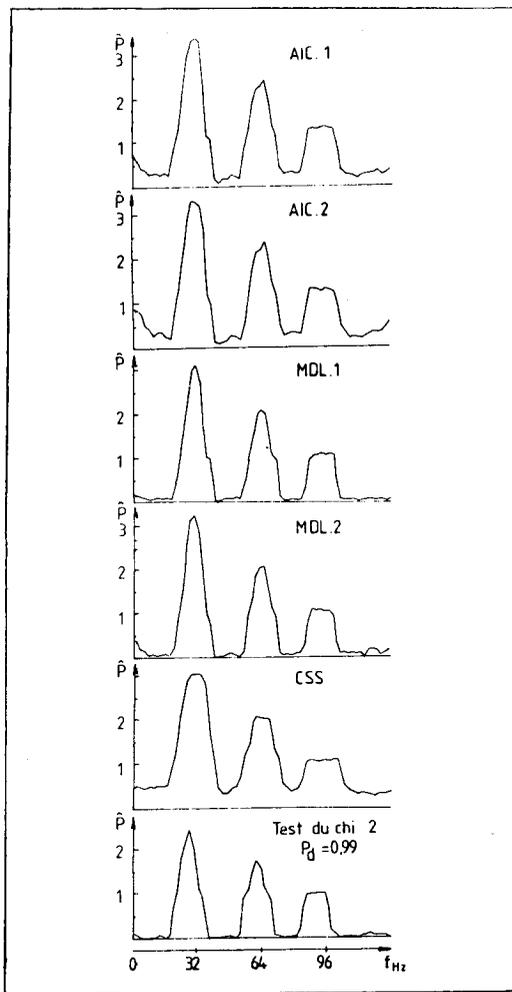


Fig. 5. - Première simulation. Bruits non corrélés et de puissances différentes sur les capteurs (trois voies modifiées). Nombre moyen de sources pour $M=110$ réalisations.

(c) Figure 5, la situation est identique au cas de (b), mais on a utilisé 110 réalisations ($M=110$). La figure représente le nombre moyen de sources obtenu à

partir de ces 110 épreuves. On constate que les critères MDL et surtout AIC surestime le nombre de sources. Le CSS donne aussi des sources parasites aux fréquences où aucune source n'a été simulée, mais donne assez exactement le nombre désiré de sources autour des fréquences 32, 64 et 96 Hz.

(d) Figure 6, les bruits sont corrélés et ont même puissance sur les capteurs. $K=12,56$ et $M=1$. Les bruits des voies 2 et 3 ont un coefficient de corrélation $C=0,7$. Le moins mauvais des critères est le CSS.

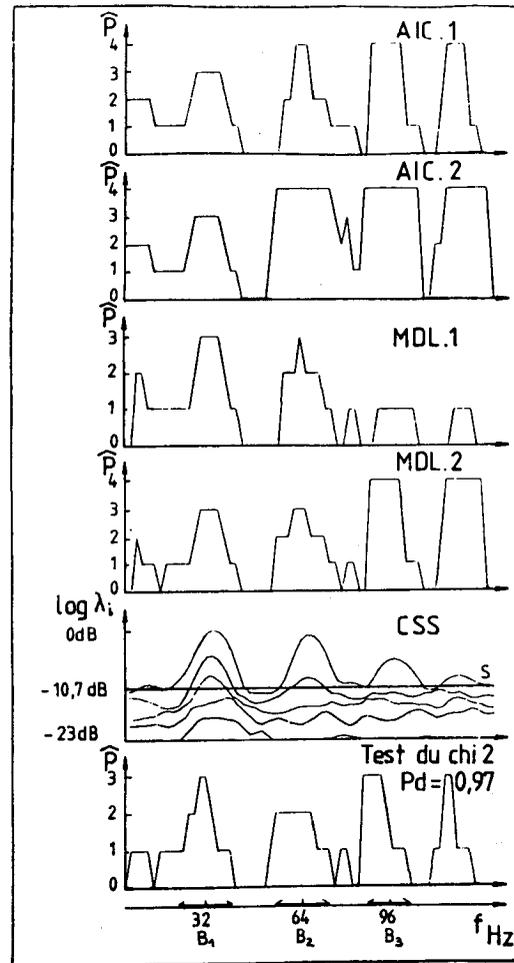


Fig. 6 - Première simulation. Bruits corrélés et de même puissance sur les capteurs. Coefficient de corrélation des voies 2 et 3 : $C=0,7$. $M=1$ et $K=12,56$.

(e) Figure 7, la situation est identique à celle de d, mais $M=3$. Alors que les critères AIC et MDL donnent toujours des sources parasites, le CSS et le test du chi-deux avec $P_d=0,95$ donnent les résultats attendus. Ceci montre qu'en moyennant sur un nombre relativement faible de réalisations, on élimine les sources parasites.

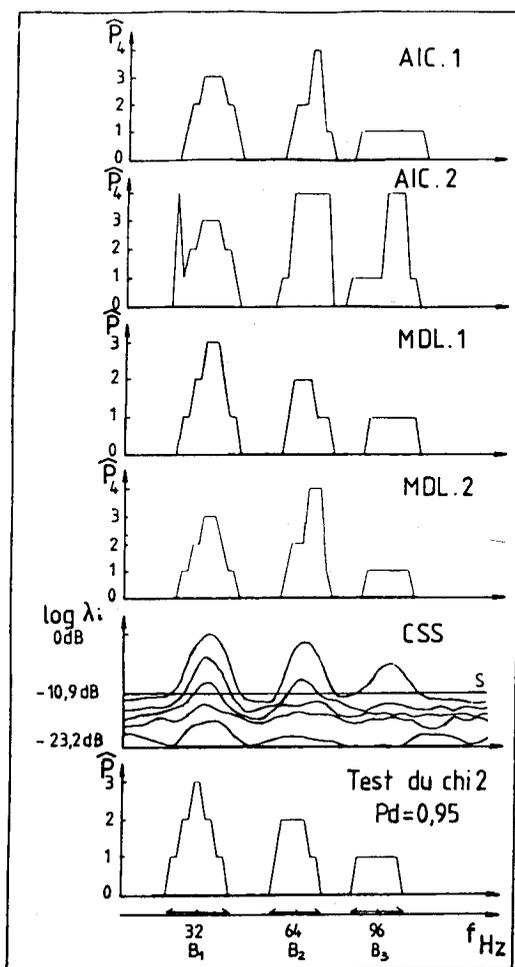


Fig. 7. — Première simulation. Même situation que figure 6. Moyenne sur $M=3$ réalisations.

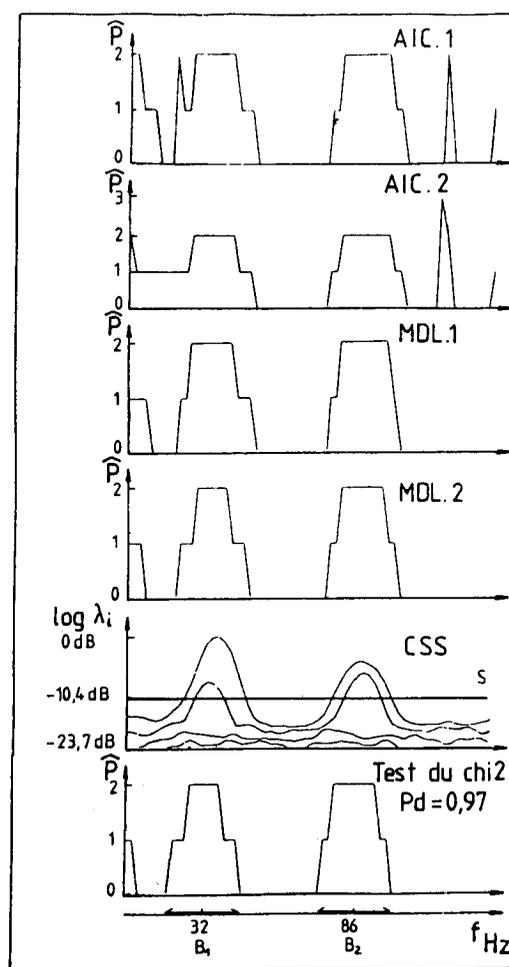


Fig. 8. — Deuxième simulation. Fronts d'ondes proches ($\delta=1$). $M=1$ et $K=12,56$.

5.3.2. Deuxième simulation

Les caractéristiques signal et bruit sont fixées aux valeurs suivantes : $A_1^1=A_1^2=30$ et $A_2^1=A_2^2=25$; $m_1^1=0,08$, $m_1^2=0,10$, $m_2^1=0,05$ et $m_2^2=0,07$.

$a_i=20$ pour $i=1,5$ (S/B moyen et bruits non corrélés et de même puissance sur les capteurs).

Les retards de propagation de l'onde hyperbolique peuvent varier par l'intermédiaire du coefficient δ .

(a) Figure 8, on utilise une seule réalisation de bruit ($K=12,56$ et $M=1$). Pour des fronts d'onde proches ($\delta=1$), le critère AIC donne des sources parasites à de nombreuses fréquences. Les critères MDL, CSS et le test du chi-deux donnent des résultats acceptables.

(b) Figure 9, pour des fronts d'onde très proches ($\delta=0,6$), $K=12,56$ et $M=1$, le test du chi-deux est le

meilleur. Les critères AIC et MDL donnent des sources parasites et le CSS perd une source en basse fréquence.

(c) Figure 10, la situation est identique à celle de (b), mais $M=3$. On améliore tous les résultats des critères sauf ceux du CSS. Le test du chi-deux conclue correctement avec des probabilités de détection plus basses. Pour conclure sur ces simulations, le test du chi-deux que nous proposons, utilisé avec des probabilités de détection de 0,95 à 0,99 donne les résultats attendus dans des situations proches de la réalité (fig. 7 et 9) alors que, comme d'autres auteurs l'ont remarqué, les critères AIC et MDL ont tendance à surestimer le nombre de sources. Le CSS donne également, dans ces cas de simulations d'assez bons résultats mais ne peut être complètement justifié.

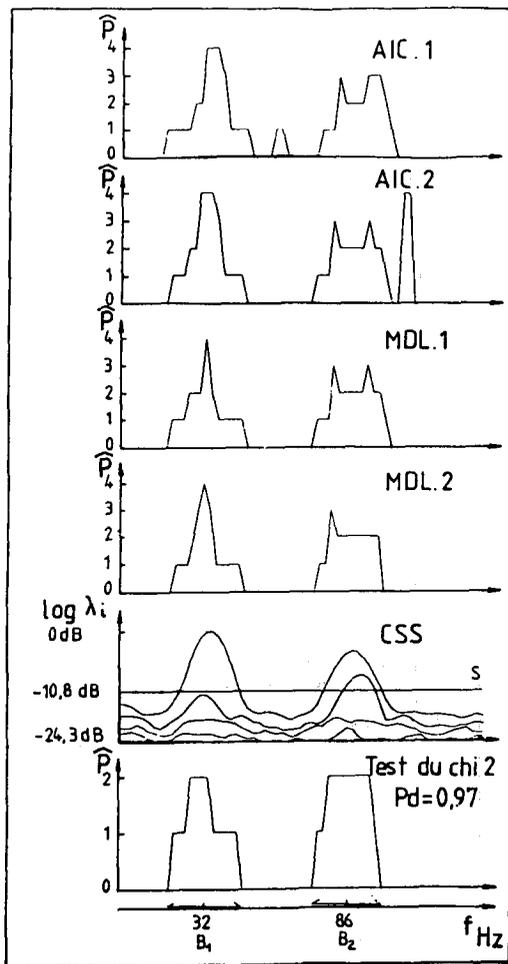


Fig. 9. — Deuxième simulation. Fronts d'ondes très proches ($\delta=0,6$). $M=1$ et $K=12,56$.

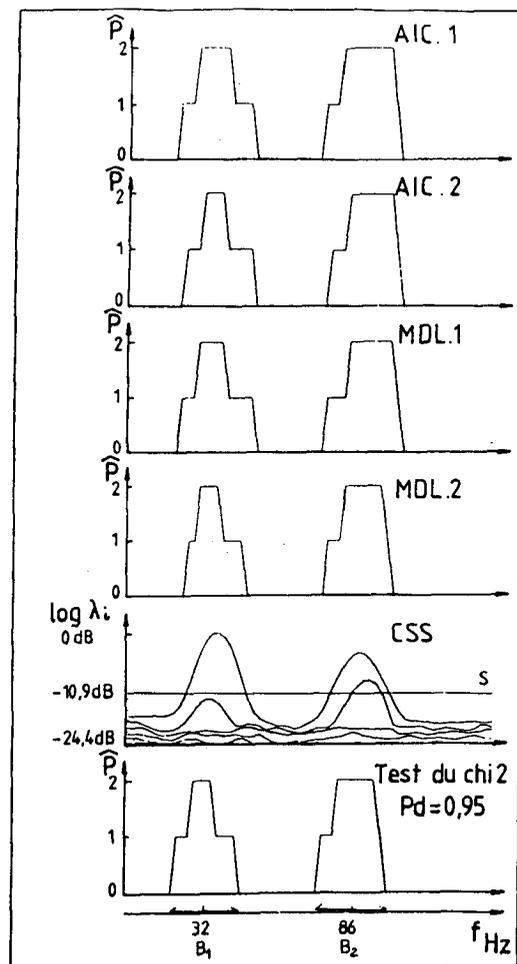


Fig. 10. — Deuxième simulation. Même situation que figure 9. Moyenne sur $M=3$ réalisations.

6. Conclusion

Nous avons utilisé les critères de détection connus de type AIC et MDL avec comme mesure d'information, l'information directe de Kullback. Nous les avons prolongé avec l'information inverse de Kullback; d'autres mesures d'information [4] peuvent être aussi utilisées [9]. Nous avons également adapté un nouveau critère avec la loi du logarithme du rapport de vraisemblance et nous avons proposé un critère simplifié à seuil portant sur les valeurs propres de la matrice spectrale estimée des signaux enregistrés.

Nous avons comparé les performances de ces critères sur des simulations réalistes. La comparaison statistique effectuée dans un cas, comme les comparaisons effectuées sur quelques réalisations aboutissent à des conclusions similaires; le test du chi-deux ou MDL pourront donc être utilisés sur des signaux réels pour lesquels le nombre de moyenne possible est souvent limité du fait de la non-stationnarité à long terme des signaux.

De notre étude, il ressort que le test du chi-deux utilisé avec des probabilités de détection de 0,95 à 0,99 est le plus performant et que le CSS est également souvent acceptable. Les critères MDL donnent, de manière générale des résultats meilleurs que les critères AIC. Dans le cas de bruits non corrélés et lorsqu'on est loin de la situation asymptotique ($B_e \cdot T$ faibles), le critère MDL2 est meilleur que MDL1. Cependant, lorsque les bruits n'ont pas même puissance sur tous les capteurs ou que le rapport signal à bruit est faible (situations fréquentes avec des signaux réels), ces critères MDL surestiment aussi le nombre de sources, ce qui n'est pas le cas avec les deux critères à seuil présentés (CSS et le test du chi-deux).

Nous proposons dans cette étude et appliquons pour la première fois en traitement du signal, un test utilisant la loi du logarithme du rapport de vraisemblance. Ce test apparaît meilleur que les autres dans la mesure où il ne surestime pas le nombre de sources et permet de conclure même sur un nombre restreint de réalisations.

Manuscrit reçu le 17 septembre 1985

BIBLIOGRAPHIE

- [1] H. AKAIKE, A new look at the statistical model identification, *IEEE, Trans. on Aut. Control*, AC-19, n° 6, Dec. 1974.
- [2] J. S. BENDAT et A. G. PIERSON, *Random data: analysis and measurement procedures*, Wiley, Interscience, 1971.
- [3] G. BIENVENU et L. KOPP, Optimality of high resolution array processing using the eigensystem approach, *IEEE, Trans. on ASSP*, ASSP-31, n° 5, Oct. 83, p. 1235-1247.
- [4] B. BOUCHON, Gains d'information : mesure de proximité en lois de probabilités, *Revue du CETHEDEC, Ondes et signal*, n° 70, 1983.
- [5] A. T. JAMES, Tests of equality of the latent roots of the covariance matrix, *Multivariate Analysis*, II, P. R. KRISHNAIAH ed., Academic Press, New York, 1969.
- [6] JENKINS et WATTS, *Spectral analysis and its application*, Holden Day, 1968.
- [7] L. KOPP et G. BIENVENU, Multiple detection using eigenvalues when the noise spatial coherence is partially unknown, *NATO, Advanced Study in Underwater Acoustics*, Lunenburg, 1984.
- [8] C. LATOMBE, Non conventional array using the eigensystem of the spectral matrix, *Proc. EUSIPCO*, 1983, p. 499-502.
- [9] J. P. LECADRE, Applications of some statistical measures to spatial signal processing, *NATO, Advanced Study in Underwater Acoustics*, Lunenburg, 1984.
- [10] W. S. LIGGET, Passive sonar: fitting models to multiple time series, *Proc. NATO, ASI on Signal Processing*, Loughborough, 1972, Academic Press, 1973.
- [11] J. RISSANEN, Modeling by shortest data description length, *Automatica*, 14, 1978, p. 417-431.
- [12] J. RISSANEN, A universal prior for integers and estimation by minimum description length, *Ann. of Stat.*, 11, n° 2, 1983.
- [13] G. SAPORTA, *Théories et méthodes de la statistique*, Technip, 1978.
- [14] G. SCHWARTZ, Estimating the dimension of a model, *Ann. of Stat.*, 6, 1978, p. 461-464.
- [15] I. TAS et C. LATOMBE, Performances de divers algorithmes pour la détection, dans les problèmes d'imagerie, du nombre d'excitations non corrélées, *Actes du GRETSI*, mai 1985, p. 363-367.
- [16] H. L. VAN TREES, *Detection, estimation and modulation theory*, Part. I, New York, Wiley, 1968.
- [17] G. VEZZOSI et P. NICOLAS, Séparation de fronts d'ondes corrélés, *Actes du GRETSI*, mai 1983, p. 277-282.
- [18] M. WAX et T. KAILATH, Determining the number of signals by information theoretic criteria, *IEEE, ASSP, Workshop on spectral estimation II*, Floride, Nov. 1983, p. 192-196.
- [19] M. WAX, TIE-JUN. SHAN et T. KAILATH, Spatio-temporel spectral analysis by eigenstructure methods, *IEEE, Trans. on ASSP*, 32, n° 4, August 1984, p. 817-827.
- [20] N. R. GOODMAN, Statistical analysis based on a certain multivariate complex gaussian distribution (an introduction), *Ann. math. Stat.*, 34, 1963, p. 152-177.
- [21] H. CLERGEOT, Estimation du spectre d'un signal aléatoire gaussien par le critère du maximum de vraisemblance ou du maximum de probabilité à postériori: *Thèse de doctorat d'état*, Paris-Sud, 1982.