

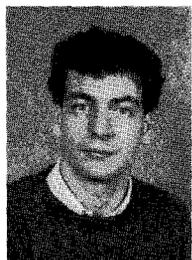
Représentation des sorties de capteurs

par un système linéaire,

applications au traitement d'antenne

Description of the sensor outputs

by means of a linear system in the array processing field



Patrice RAVAZZOLA

GERDSM, LE BRUSC, 83140 SIX-FOURS-LES-PLAGES.

Patrice Ravazzola a obtenu le diplôme d'ingénieur ENSM, Nantes et le DEA d'automatique en 1986. Depuis, il exerce des activités de recherche au GERDSM en traitement d'antenne appliqué à l'acoustique sous-marine dans le cadre de la préparation d'une thèse de doctorat.



Jean-Pierre LE CADRE

IRISA, Campus de Beaulieu, 35042 RENNES CEDEX.

Jean-Pierre Le Cadre a soutenu une thèse de 3^e cycle en juin 1982, puis une thèse de Doctorat d'État en octobre 1987, toutes deux en traitement du signal et à l'INPG. De 1980 à 1988, ses travaux portent essentiellement sur le traitement d'antenne appliqué à l'acoustique sous-marine et sont effectués au GERDSM. Depuis le 1^{er} octobre 1988, J.-P. Le Cadre est affecté à l'IRISA (Rennes) en qualité de Chargé de recherche au CNRS.

RÉSUMÉ

Pour une antenne linéaire à capteurs équidistants, la description des sorties des capteurs par un système linéaire fournit un cadre bien adapté à l'utilisation des hypothèses physiques. La prise en compte de l'hypothèse onde plane et de la corrélation spatiale du bruit le long de l'antenne par un modèle d'état nous conduit à considérer les méthodes de réalisation qui ont été développées pour l'étude des systèmes linéaires. Parmi celles-ci, on s'intéresse plus particulièrement aux méthodes fondées sur l'analyse en composantes principales, et on établit leurs liens avec le cadre habituel du traitement d'antenne.

Les performances comparées de ces méthodes et du goniomètre sont illustrées par quelques résultats de simulation.

MOTS CLÉS

Traitement d'antenne, méthode de réalisation, modèle d'état, onde plane, corrélation spatiale du bruit.

SUMMARY

For a linear array with equispaced sensors, the sensor outputs may be efficiently described by a linear system. This general frame allows us to take into account classical physical hypotheses like plane wave hypothesis or spatial correlation of the noise received along the array; this is achieved by means of state space modelling. This leads us to consider stochastic realization methods which have been developed in the general context of linear systems. Among these, a particular attention is focused on methods based on principal components; links with array processing are presented.

The performances of these methods are compared with the MUSIC method for simulated data.

KEY WORDS

Array processing, realization method, state space model, plane wave, spatial correlation of the noise.

Principales notations utilisées

Une lettre (majuscule ou minuscule) en caractère gras correspond à un vecteur.

Une lettre majuscule maigre (droite ou cursive) correspond à une matrice (sauf L, M, N).

dét : déterminant d'une matrice carrée;

tr : trace d'une matrice;

I : matrice identité;

E(.) : espérance mathématique d'une v. a.;

X|Y : projection orthogonale de la v. a. X sur l'espace engendré par la v. a. Y;

$\bar{\cdot}$: quantité conjuguée;

\cdot^T : quantité transposée;

\cdot^* : quantité transposée-conjuguée;

\cdot^* : pseudo-inverse à gauche d'une matrice;

$|\cdot|$: module d'un nombre complexe;

$j^2 = -1$.

1. Introduction

On s'intéresse à l'estimation des angles d'incidence de signaux sur un réseau de capteurs.

On considère que la propagation d'une onde plane est déterministe alors que le signal porté par cette onde est un processus aléatoire.

Après transformée de Fourier, chaque signal est caractérisé par son amplitude et sa phase; à cause des variations aléatoires des signaux émis, des fluctuations des propriétés du milieu de propagation et des mouvements des sources et de l'antenne, ces paramètres sont considérés comme des réalisations de variables aléatoires. L'amplitude est supposée suivre une loi de Rayleigh et la phase une loi uniforme entre 0 et 2π , ce qui revient à supposer que le signal (d'une source) suit une loi normale centrée (circulaire) de variance la puissance de la source considérée.

Le but du traitement d'antenne (TA) étant de fournir une analyse spatiale du champ de bruit reçu sur l'antenne, on effectue habituellement une analyse fréquentielle suivie d'une analyse spatiale.

Les méthodes dites à haute résolution (HR) ont été l'objet de nombreux développements en traitement d'antenne et en analyse spectrale. Fondées sur une décomposition algébrique (hypothèse de cohérence des sources) de l'espace des observations \mathcal{E} en sous-espace « source » \mathcal{E}_S (signal utile) et sous-espace « bruit » \mathcal{E}_B , ces méthodes ont conduit à une amélioration sensible de la résolution angulaire relativement aux méthodes usuelles non paramétriques [1, 2]. Cette séparation s'effectue par l'analyse des éléments propres de la matrice interspectrale des observations (MIO), que nous notons Γ .

Si on considère le cas de M ondes planes incidentes sur une antenne linéaire composée de N capteurs

équidistants, et si de plus le bruit ambiant est décorrélé de capteur à capteur, la MIO Γ s'écrit :

$$(1) \quad \Gamma = \mathcal{D} \cdot \gamma_S \mathcal{D}^* + b^2 \cdot I = \Gamma_S + \Gamma_B$$

où : \mathcal{D} est la matrice (N, M) des directions des sources

$$\mathcal{D}(\mathbf{D}_{\theta_1}, \mathbf{D}_{\theta_2}, \dots, \mathbf{D}_{\theta_M})$$

\mathbf{D}_{θ_k} est le vecteur direction (N, 1) de la source de gisement θ_k (angle entre la direction de l'onde et l'axe de l'antenne). \mathbf{D}_{θ_k} est la fonction de transfert multidimensionnelle de l'ensemble (milieu + antenne).

$$\mathbf{D}_{\theta_k}^T = (1, \exp[-j \cdot \Phi_k], \dots, \exp[-j(N-1) \Phi_k])$$

Φ_k dépendant de θ_k , de la longueur d'onde et de la géométrie de l'antenne ($j^2 = -1$).

γ_S est la matrice (M, M) de covariance des sources.

b^2 est la densité spectrale de puissance (d. s. p.) du bruit (blanc, gaussien, centré).

Γ a pour système propre (U, Σ^2) :

$$(2) \quad \Gamma = U \cdot \Sigma^2 \cdot U^*$$

où Σ^2 est la matrice diagonale (N, N) des valeurs propres rangées par ordre de module décroissant et U la matrice unitaire (N, N) des vecteurs propres.

Lorsque la MIO correspond exactement au modèle (somme d'ondes planes et de bruit blanc), le sous-espace \mathcal{E}_S , respectivement \mathcal{E}_B , est engendré par les vecteurs propres correspondants aux M plus grandes valeurs propres de Γ , respectivement à la plus petite valeur b^2 de multiplicité (N-M); G. Bienvenu et L. Kopp ont justifié statistiquement cette décomposition [3].

V. F. Pisarenko [4] est l'un des instigateurs des méthodes HR. La méthode qu'il propose en 1973 repose sur le théorème de Caratheodory qui permet de représenter la séquence finie des covariances des observations comme une somme pondérée de sinusoides (sources) et de bruit blanc. Il s'agit alors d'identifier et d'éliminer la contribution du bruit dans la MIO, puis d'en extraire le mineur principal [de rang M et de dimension (M+1, M+1)] pour calculer le vecteur propre \mathbf{V}_0 correspondant à l'unique valeur propre nulle.

Les gisements des sources (les fréquences des sinusoides en analyse spectrale) s'obtiennent en déterminant les zéros du polynôme en z ($z = \exp[j \Phi_k]$) ayant pour coefficients les composantes de \mathbf{V}_0 . Ceci revient à chercher les gisements θ_k qui annulent la projection du vecteur direction \mathbf{D}_{θ_k} sur le sous-espace « bruit ». Connaissant les gisements des sources, Pisarenko détermine les puissances des sources comme solutions d'un système linéaire d'ordre M.

Lorsque la MIO est estimée à partir des observations, l'orthogonalité des vecteurs directions des sources avec un vecteur du sous-espace « bruit » \mathcal{E}_B n'est qu'approximativement vérifiée. G. Bienvenu et L. Kopp (le goniomètre : [5]), Schmidt (MUSIC : [6]) suggèrent alors d'utiliser l'orthogonalité de \mathbf{D}_{θ_k} aux (N-M) vecteurs propres de Γ engendrant \mathcal{E}_B .

Dans le même but, R. Kumaresan et D. W. Tufts [7] proposent une méthode qui consiste à déterminer un vecteur \mathbf{d} , sous les contraintes de norme minimale et de première composante égale à 1, comme combinaison linéaire des $(N-M)$ « plus petits » vecteurs propres de Γ et décrivant entièrement \mathcal{E}_B .

Cependant les méthodes citées ci-dessus, sauf Pisarenko, n'utilisent pas de façon fondamentale l'hypothèse « onde plane » pour la séparation entre sous-espaces sources et bruit. Une plus grande richesse de description est rendue possible par l'introduction de méthodes utilisant un modèle d'état. L'hypothèse « onde plane » est alors incluse au modèle même de l'état du système.

L'information sur la position angulaire (gisement) des sources est contenue dans la matrice de transition du système, et la détermination des gisements est ramené à un problème de réalisation pour lequel de nombreuses solutions ont été proposées en automatique [8, 9]. Si le signal reçu sur l'antenne est une somme d'ondes planes et de bruit blanc, alors la sortie de chaque capteur se prête très bien à la modélisation par un modèle d'état sans bruit sur l'équation d'évolution de l'état :

$$(3) \quad \begin{cases} \mathbf{X}_{i+1} = \mathbf{F} \cdot \mathbf{X}_i \\ y_i = \mathbf{h}^* \cdot \mathbf{X}_i + w_i \end{cases} \quad (\mathcal{M}_B)$$

où :

i est le numéro de capteur (variable spatiale et non temporelle).

y_i est la sortie (scalaire) du i -ième capteur.

\mathbf{X}_i est le vecteur $(M, 1)$ d'état du système.

$\{w_i\}$ est un bruit blanc gaussien centré de d. s. p. b^2 .

\mathbf{F} est la matrice (M, M) de transition du système.

\mathbf{h} est un vecteur $(M, 1)$ de pondération.

S. Y. Kung, K. S. Arun et D. V. Bhaskar Rao [10, 20] ont proposé une méthode utilisant l'analyse en composantes principales de la MIO Γ et permettant de déterminer la matrice de transition \mathbf{F} (dont on déduit aisément les gisements des sources) d'un modèle d'état sans bruit sur l'équation d'état. Cette méthode consiste à exprimer la MIO Γ en fonction des paramètres du système, notamment la matrice d'observabilité \mathbf{O} (§ 2). La matrice de transition \mathbf{F} se calcule alors à partir de la matrice \mathbf{O} par simple considération sur la « forme » même de \mathbf{O} .

Cette approche s'exprime très bien dans le contexte habituel du traitement d'antenne (§ 3).

Par ailleurs, on peut prendre en compte la corrélation du bruit en considérant un modèle d'état de forme innovation (modèle ARMA à phase minimale) pour la sortie des capteurs :

$$(4) \quad \begin{cases} \mathbf{X}_{i+1} = \mathbf{F} \cdot \mathbf{X}_i + \mathbf{T} \cdot v_i \\ y_i = \mathbf{h}^* \cdot \mathbf{X}_i + v_i \end{cases} \quad (\mathcal{M}_C)$$

(i variable spatiale) où \mathbf{T} est un vecteur $(M, 1)$ et $\{v_i\}$ un bruit blanc gaussien centré de d. s. p. b^2 .

Outre les algorithmes usuels de réalisation stochastique, les méthodes fondées sur l'analyse en composantes principales sont d'un intérêt pratique certain.

On peut citer deux approches :

– La première, proposée par K. S. Arun, D. V. Bhaskar Rao et S. Y. Kung [11, 12], consiste à chercher un vecteur d'état \mathbf{X}_i , appartenant à l'espace engendré par le « passé » des observations, qui minimise la variance de l'erreur de prédiction du « futur » des observations. Ceci nous fournit une expression de la matrice d'observabilité du système \mathbf{O} qui comme pour la méthode de réalisation dans le cas bruit blanc (MRB) nous conduit à une estimation de la matrice de transition du système \mathbf{F} , laquelle contient toute l'information qui nous intéresse.

– Une autre approche proposée par U. B. Desai et D. Pal [13] consiste à utiliser les coefficients de corrélation canoniques entre le « futur » et le « passé » des observations [17, 21] que l'on introduit en considérant l'information mutuelle entre le « futur » et le « passé » des observations. La matrice de transition du système \mathbf{F} se déduit alors des dynamiques de l'état \mathbf{X} par projection.

Ces deux approches que l'on regroupe sous l'appellation méthode de réalisation dans le cas d'un bruit additif corrélé (MRC) seront développées dans le paragraphe 4.

L'emploi de méthodes de réalisation (bruit blanc ou bruit corrélé) qui utilisent mieux l'hypothèse « onde plane » laisse supposer une amélioration des performances vis-à-vis des autres méthodes HR, notamment une meilleure détection des sources faibles, et, pour les MRC, une plus grande robustesse à la coloration spatiale du bruit [14, 15].

Après avoir présenté les méthodes de réalisation fondées sur l'analyse en composantes principales, on étudie leur comportement sur des signaux simulés et on compare leurs performances à celles du goniomètre.

2. Introduction d'un modèle d'état

2.1. CARACTÉRISATION DE LA MATRICE DE TRANSITION [10]

On considère la représentation d'état d'un système linéaire ne comportant pas de bruit d'entrée sur l'équation d'état :

$$\begin{cases} \mathbf{X}_{i+1} = \mathbf{F} \cdot \mathbf{X}_i \\ y_i = \mathbf{h}^* \cdot \mathbf{X}_i + w_i \end{cases} \quad (\mathcal{M}_B)$$

La matrice de covariance de l'état $\mathbf{P}_i = E(\mathbf{X}_i \cdot \mathbf{X}_i^*)$ satisfait à l'équation de Lyapunov :

$$(5) \quad \mathbf{P}_{i+1} = \mathbf{F} \cdot \mathbf{P}_i \cdot \mathbf{F}^*$$

La fonction de corrélation Φ du processus $\{y_i\}$ s'exprime en fonction des paramètres du système :

$$(6) \quad \Phi(y_{i+m}, y_i^*) = \mathbf{h}^* \cdot \mathbf{F}^m \cdot \mathbf{P}_i \cdot \mathbf{h} + b^2 \cdot \delta(m)$$

avec

$$\delta(m) = \begin{cases} 1, & \text{si } m=0 \\ 0, & \text{si } m \neq 0. \end{cases}$$

Soit encore, en tenant compte de l'équation de Lyapunov et en notant :

$$(7) \quad \begin{aligned} P &= P_1 = E(X_1 \cdot X_1^*) \\ \Phi(y_{i+m} \cdot y_i^*) &= \mathbf{h}^* \cdot F^{m+i-1} \cdot P \cdot (F^{(i-1)})^* \cdot \mathbf{h} + b^2 \cdot \delta(m) \end{aligned}$$

On vérifie aisément que la matrice de covariance $\Gamma = [\Phi(y_i \cdot y_j^*)]$ s'écrit sous forme factorisée :

$$(8) \quad \Gamma = O \cdot P \cdot O^* + b^2 \cdot I$$

où :

$$O = \begin{bmatrix} \mathbf{h}^* \\ \mathbf{h}^* \cdot F \\ \mathbf{h}^* \cdot F^2 \\ \vdots \end{bmatrix}$$

est la matrice d'observabilité du système [19], et

$$I = \begin{bmatrix} 1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & \ddots \end{bmatrix}$$

Lorsque le processus $\{y_i\}$ est stationnaire, on note : $\Phi(y_{i+m} \cdot y_i^*) = \Gamma_m$ pour tout i . La matrice de covariance du processus est alors une matrice de Toeplitz positive :

$$(9) \quad \Gamma = \begin{bmatrix} \Gamma_0 & \Gamma_{-1} & \Gamma_{-2} & \dots \\ \Gamma_1 & \Gamma_0 & \Gamma_{-1} & \dots \\ \Gamma_2 & \Gamma_1 & \Gamma_0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

avec :

$$(10) \quad \Gamma_m = \mathbf{h}^* \cdot F^m \cdot P \cdot \mathbf{h} + b^2 \cdot \delta(m)$$

Pour déterminer la matrice de transition du système, on considère l'opérateur « up » qui décale les lignes d'une matrice vers le haut d'une ligne en supprimant la première. La matrice O satisfait l'équation :

$$(11) \quad O \cdot F = O^{up}$$

La matrice de transition F se déduit de l'équation précédente :

$$(12) \quad F = O^* \cdot O^{up}$$

où : $O^* = (O^* \cdot O)^{-1} \cdot O^*$, pseudo-inverse à gauche,

$$O^{up} = \begin{bmatrix} \mathbf{h}^* \cdot F \\ \mathbf{h}^* \cdot F^2 \\ \mathbf{h}^* \cdot F^3 \\ \vdots \end{bmatrix}$$

Il faut noter que l'écriture de la matrice de covariance sous la forme $\Gamma = O \cdot P \cdot O^* + b^2 \cdot I$ n'est pas valable pour un modèle ARMA général alors que la matrice

Hankel du système peut se décomposer en $H = O \cdot G$ pour tout modèle ARMA. Déterminer O à partir de Γ au lieu de H revient à chercher un modèle dans un sous-ensemble (qui contient le modèle recherché) de l'ensemble des modèles ARMA. De plus, le problème est mieux conditionné quand on l'exprime en fonction de Γ . En effet, Γ est hermitienne, et en présence de bruit blanc, elle est toujours de rang maximal alors que H est singulière. Il est donc raisonnable de déterminer O à partir de la matrice de covariance et non pas de la matrice de Hankel.

En dimension finie ($\dim = N$), l'expression de O^{up} nécessite l'introduction d'une dernière ligne de zéros :

$$(13) \quad O = \begin{bmatrix} \mathbf{h}^* \\ \mathbf{h}^* \cdot F \\ \vdots \\ \mathbf{h}^* \cdot F^{(N-1)} \end{bmatrix}, \quad O^{up} = \begin{bmatrix} \mathbf{h}^* \cdot F \\ \mathbf{h}^* \cdot F^2 \\ \vdots \\ \mathbf{h}^* \cdot F^{(N-1)} \\ \mathbf{0}^* \end{bmatrix}$$

($\mathbf{0}$: vecteur nul).

Le calcul de F par pseudo-inverse nous conduit alors à l'approximation :

$$(14) \quad F \simeq O^* \cdot O^{up}$$

Une solution exacte s'obtient en définissant les opérateurs « \uparrow » et « \downarrow » qui sélectionnent respectivement les $(N-1)$ premières et les $(N-1)$ dernières lignes d'une matrice :

$$(15) \quad O^\uparrow = \begin{bmatrix} \mathbf{h}^* \\ \mathbf{h}^* \cdot F \\ \vdots \\ \mathbf{h}^* \cdot F^{(N-2)} \end{bmatrix}, \quad O^\downarrow = \begin{bmatrix} \mathbf{h}^* \cdot F \\ \mathbf{h}^* \cdot F^2 \\ \vdots \\ \mathbf{h}^* \cdot F^{(N-1)} \end{bmatrix}$$

[matrices de dimension $(N-1, M)$].

D'où :

$$(16) \quad F = (O^\downarrow)^* \cdot O^\uparrow$$

2. 2. EXPRESSION DE LA MATRICE DE TRANSITION EN FONCTION DES ÉLÉMENTS PROPRES DE LA MATRICE DE COVARIANCE

Considérons maintenant la décomposition de la matrice de covariance en éléments propres :

$$(17) \quad \Gamma = U \cdot \Sigma^2 \cdot U^* = [U_1 \ U_2] \begin{bmatrix} \Sigma_1^2 & 0 \\ 0 & \Sigma_2^2 \end{bmatrix} \begin{bmatrix} U_1^* \\ U_2^* \end{bmatrix}$$

où Σ_1^2 est la matrice (M, M) diagonale contenant les M plus grandes valeurs propres, et Σ_2^2 les $(N-M)$ autres, U_1 et U_2 étant les matrices des vecteurs propres associées respectivement à Σ_1^2 et Σ_2^2 .

La meilleure approximation de rang M de Γ au sens de la norme matricielle (norme de Frobenius, ou norme spectrale) s'obtient en ne retenant que les M principales composantes :

$$(18) \quad U_1 \cdot \Sigma_1^2 \cdot U_1^*$$

En présence de *bruit blanc*, seules les valeurs propres étant affectées, on peut éliminer l'effet du bruit en soustrayant aux valeurs propres l'estimation au sens du maximum de vraisemblance de la densité spectrale du bruit \hat{b}^2 (moyenne arithmétique des $(N-M)$ plus petites valeurs propres de Γ) :

$$(19) \quad U_1 (\Sigma_1^2 - \hat{b}^2) U_1^* = U_1 \cdot \hat{\Sigma}_1^2 \cdot U_1^*$$

Remarque. — Cette approximation de la partie « signal » de la matrice de covariance ne permet pas de conserver le caractère Toeplitz qu'elle possède si le processus est stationnaire.

La comparaison de cette expression de la partie « signal » de la matrice de covariance estimée, à la forme asymptotique « bruitée » (8) nous permet d'établir la relation suivante :

$$(20) \quad O \cdot P \cdot OO^* = U_1 \cdot \hat{\Sigma}_1^2 \cdot U_1^*$$

Par utilisation de la SVD (décomposition en valeurs singulières) des deux produits matriciels ($O \cdot P^{1/2}$) et ($U_1 \cdot \Sigma_1$), on montre qu'il existe une matrice unitaire W de dimension (M, M) telle que :

$$(21) \quad OO \cdot P^{1/2} = U_1 \cdot \Sigma_1 \cdot W^*$$

Il s'agit donc maintenant de déterminer la matrice de transition F à partir de l'expression précédente. Ceci sera obtenu grâce au lemme ci-dessous de démonstration évidente.

Lemme 1 : *Propriété des opérateurs « \uparrow » et « \downarrow ».*

Soit A et B deux matrices quelconques de dimensions (l, m) et (m, n) , on a :

$$(22) \quad (A \cdot B)^\uparrow = A^\uparrow \cdot B \quad \text{et} \quad (A \cdot B)^\downarrow = A^\downarrow \cdot B$$

Appliquons cette propriété à l'équation (21) :

$$(23) \quad \begin{cases} O^\uparrow \cdot P^{1/2} = U_1^\uparrow \Sigma_1 \cdot W^* \\ O^\downarrow \cdot P^{1/2} = U_1^\downarrow \Sigma_1 \cdot W^* \end{cases}$$

En utilisant la propriété $O^\uparrow \cdot F = O^\downarrow$, on en déduit une expression de F :

$$(24) \quad F = P^{1/2} \cdot W \cdot \hat{\Sigma}_1^{-1} \cdot [(U_1^\uparrow)^* \cdot U_1^\downarrow] \cdot \hat{\Sigma}_1 \cdot W^* \cdot P^{-1/2}$$

Nous ne donnons pas ici les équations permettant de déterminer les autres paramètres du système linéaire car ils sont pour nous de moindre intérêt.

Bien que ne connaissant pas les matrices P et W , on montre, en utilisant la propriété classique du déterminant d'un produit matriciel ($\det(A \cdot B) = \det(B \cdot A)$) que cette expression suffit pour déterminer le spectre de F (ensemble des valeurs propres de F). En effet, on obtient alors :

$$\det(F - \lambda I) = \det([(U_1^\uparrow)^* \cdot U_1^\downarrow] - \lambda I)$$

Ce qui signifie encore que les spectres des matrices F et $[(U_1^\uparrow)^* \cdot U_1^\downarrow]$ sont identiques.

Propriété 1 : Les valeurs propres de F sont identiques aux valeurs propres de $[(U_1^\uparrow)^* \cdot U_1^\downarrow]$.

Il suffit donc de déterminer les M vecteurs propres « principaux » de Γ pour calculer les valeurs propres de la matrice de transition du système F .

Remarques. On peut noter que ce calcul a nécessité quelques hypothèses :

- bruit blanc;
- matrice de covariance de l'état du système de rang maximal (rang $P = M$).

Par ailleurs, l'estimation de la partie signal de la matrice de covariance à l'aide des M composantes principales détruit le caractère de Toeplitz qu'elle possède dans le cas stationnaire.

3. Lien avec le traitement d'antenne

3.1. MÉTHODE DE RÉALISATION SPATIALE, CAS DU BRUIT ADDITIF BLANC (MRB)

Le champ de bruiteurs est supposé composé de M sources $\{S_k\}_{k=1,2,\dots,M}$ de gisements $\{\theta_k\}_{k=1,2,\dots,M}$ et de niveaux $\{\sigma_k^2\}_{k=1,2,\dots,M}$. L'antenne est linéaire et comporte N capteurs équidistants (d = distance intercapteurs). Soit c la célérité des ondes dans le milieu de propagation. Le bruit ambiant est supposé blanc, gaussien, centré, de densité spectrale de puissance b^2 .

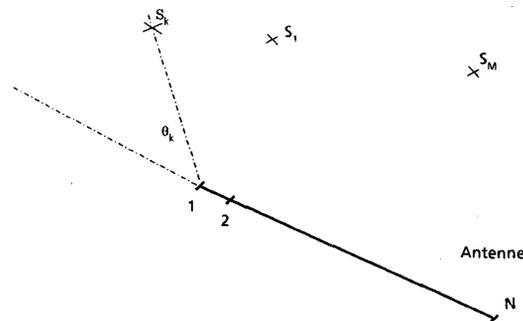


Fig. 1

A la fréquence f considérée, la MIO s'écrit :

$$\Gamma = \mathcal{D} \cdot \gamma_S \cdot \mathcal{D}^* + b^2 \cdot I$$

où :

\mathcal{D} est la matrice (N, M) des directions des sources, $\mathcal{D}(\mathbf{D}_{\theta_1}, \mathbf{D}_{\theta_2}, \dots, \mathbf{D}_{\theta_M})$. \mathbf{D}_{θ_k} est le vecteur direction $(N, 1)$ de la source de gisement θ_k

$$\mathbf{D}_{\theta_k} = (1, \exp[-j\Phi_k], \dots, \exp[-j(N-1)\Phi_k])^T$$

$\Phi_k = 2\pi f \tau_k$, avec $\tau_k = d/c \cos \theta_k$ (retard intercapteurs). γ_S est la matrice (M, M) de covariance des sources.

Lorsque les sources ne sont pas corrélées, γ_S est la matrice diagonale des puissances des sources ($\gamma_S = \text{diag}(\sigma_k^2)$).

Le modèle d'état correspondant à cette configuration est le modèle d'état (\mathcal{M}_B) présenté au paragraphe

précédent :

$$\begin{cases} \mathbf{X}_{i+1} = \mathbf{F} \cdot \mathbf{X}_i \\ y_i = \mathbf{h}^* \cdot \mathbf{X}_i + w_i \end{cases} \quad (\mathcal{M}_B)$$

où y_i est la sortie du i -ième capteur (à la fréquence f). y_1 est la sortie du 1^{er} capteur, pris comme référence. L'état correspondant \mathbf{X}_1 contient les contributions de chaque source au capteur numéro 1, à un instant donné. Chaque contribution dépend de la réalisation de la variable aléatoire dont est fonction l'amplitude du signal. L'état \mathbf{X}_i peut donc être défini comme étant l'ensemble des contributions des sources au capteur numéro i pour les mêmes réalisations des variables aléatoires. Par conséquent, les liens entre les vecteurs d'état sont déterministes et s'expriment à l'aide de la matrice de transition du système qui est diagonale dans une certaine base (base qui rend indépendantes les contributions des sources).

$$\mathbf{F} = \text{diag} (\exp [-j \Phi_k])$$

L'information sur la puissance des sources peut se répartir entre le vecteur de pondération \mathbf{h} et la matrice de variance de l'état.

Comme précédemment, l'analyse en composantes principales de la MIO Γ nous permet d'écrire :

$$(25) \quad \mathcal{D} \cdot \gamma_s \cdot \mathcal{D}^* \mathbf{U}_1 \cdot \hat{\Sigma}_1^2 \cdot \mathbf{U}_1^*$$

avec (\mathbf{U}, Σ^2) système propre de Γ .

On montre également par utilisation de la SVD qu'il existe une matrice (M, M) unitaire \mathbf{W} telle que :

$$(26) \quad \mathcal{D} \cdot \gamma_s^{1/2} = \mathbf{U}_1 \cdot \Sigma_1 \cdot \mathbf{W}^*$$

La matrice des vecteurs directions de sources s'écrit :

$$\mathcal{D} = \begin{pmatrix} 1 & \dots & 1 \\ \exp [-j \Phi_1] & & \exp [-j \Phi_M] \\ \vdots & \vdots & \\ \exp [-j(N-1) \Phi_1] & \dots & \exp [-j(N-1) \Phi_M] \end{pmatrix}$$

l'utilisation des opérateurs « \uparrow » et « \downarrow » pour cette matrice nous permet de prendre en compte l'hypothèse « onde plane » :

$$(27) \quad \mathcal{D}^\uparrow \cdot \mathbf{F} = \mathcal{D}^\downarrow$$

où \mathbf{F} est la matrice diagonale relative aux déphasages entre capteurs pour chaque source qui traduit la propagation des ondes planes le long de l'antenne.

On peut alors calculer, comme au paragraphe précédent, la matrice de transition du système \mathbf{F} :

$$(28) \quad \mathbf{F} = \gamma_s^{1/2} \cdot \mathbf{W} \cdot \hat{\Sigma}_1^{-1} \cdot [(\mathbf{U}_1)^* \cdot \mathbf{U}_1] \cdot \hat{\Sigma}_1 \cdot \mathbf{W}^* \cdot \gamma_s^{-1/2}$$

On montre aisément (de manière analogue au calcul proposé au paragraphe précédent) que les matrices \mathbf{F} et $[(\mathbf{U}_1)^* \cdot \mathbf{U}_1]$ ont mêmes valeurs propres. On a donc une estimation des valeurs propres de la matrice de transition du système correspondant au problème

posé :

$$(29) \quad \hat{\mathbf{F}} = \text{diag} (\hat{\rho}_k \cdot \exp [-j \hat{\Phi}_k])$$

(ρ_k : module de la k -ième valeur propre).

On en déduit la position angulaire des sources :

$$(30) \quad \hat{\theta}_k = \text{Arc Cosinus} \left(\frac{c}{2 \cdot \pi \cdot f \cdot d} \hat{\Phi}_k \right) \\ k = 1, \dots, M$$

Propriété 2 : Les gisements des sources sont estimés à partir des valeurs propres de la matrice $[(\mathbf{U}_1)^* \cdot \mathbf{U}_1]$.

3.2. SENSIBILITÉ A LA CORRÉLATION DES SOURCES

3.2.1. Corrélation des sources. Lissage temporel et spatial

Le problème de l'estimation de la position angulaire de sources corrélées se pose dans certains domaines d'application du traitement d'antenne.

On s'intéresse ici plus particulièrement à l'acoustique sous-marine et on peut citer deux exemples d'application.

Le premier concerne la localisation d'un bruiteur à partir des divers rayons issus de celui-ci. En effet, il est fréquent qu'un signal émis par une source parvienne au récepteur par plusieurs trajets du fait des réflexions sur la surface et sur le fond marin. Ces signaux reçus sur l'antenne peuvent être corrélés et de sites (angles d'arrivée dans le plan vertical) proches. Il s'agit alors de séparer les divers trajets et, moyennant des connaissances complémentaires (propagation), utiliser ces informations pour localiser le bruiteur (télémétrie).

L'analyse des bruits rayonnés par un bâtiment constitue le deuxième exemple d'application de la séparation de sources corrélées. Dans ce cas, des sources ponctuelles situées sur le bâtiment peuvent émettre des signaux corrélés (fonction de transfert du type couplage mécanique, etc.) qu'il est nécessaire de séparer.

Le problème de la séparation de trajets corrélés est donc d'un intérêt certain. Les moyens dont on dispose pour le résoudre sont les suivants.

D'une part, les différentes méthodes HR qui ont été développées fournissent des estimations sans biais, même dans le cas de sources partiellement corrélées (γ_s de rang maximal). En pratique, leurs performances peuvent se dégrader. Dans le cas de sources parfaitement corrélées, ces méthodes ne s'appliquent plus (γ_s n'est plus de rang maximal).

En effet, ces méthodes rencontrent, en théorie, des difficultés uniquement dans le cas de sources parfaitement corrélées (la matrice γ_s n'est pas de rang maximal) [16].

Mais pour des données réelles, le cas de sources parfaitement corrélées est peu probable, car si l'on considère deux trajets issus d'une même source, la propagation des deux ondes dans le milieu sous-marin dépend de nombreux paramètres et les causes de dégradation

de la corrélation sont multiples (réflexion sur la surface ou sur le fond, présence d'ondes dans le milieu, mouvement des bruiteurs et du récepteur...).

D'autre part, on a la possibilité d'employer des techniques de décorrélation pour éviter une dégradation des performances des méthodes pouvant se produire lorsque la corrélation entre les sources est importante. Le lissage temporel et le lissage spatial sont deux de ces techniques.

Le lissage temporel s'effectue de façon naturelle lorsqu'on emploie une méthode qui nécessite l'estimation de la MIO Γ .

En effet, on estime Γ en « moyennant » les transformées de Fourier des échantillons temporels des signaux reçus sur l'antenne (périodogramme), ce qui implique une intégration sur une certaine durée des divers phénomènes qui sont la cause de la décorrélation des sources. Il apparaît donc que la décorrélation est d'autant meilleure que la durée d'intégration est grande.

Autre technique de décorrélation, le lissage spatial [16, 22, 23] consiste à diviser l'antenne en $(N-n+1)$ sous-antennes de n capteurs avec $(n-1)$ capteurs communs entre deux sous-antennes adjacentes, puis à calculer la MIO $\Gamma_k^{(n)}$ qui correspond à chaque sous-antenne.

Soient δ_{θ_i} le vecteur $(n, 1)$ extrait du vecteur D_{θ_i} en ne considérant que les n premières lignes :

$$\delta_{\theta_i} = \begin{pmatrix} 1 \\ \exp[-j\Phi_i] \\ \vdots \\ \exp[-j(n-1)\Phi_i] \end{pmatrix}$$

Δ la matrice (n, M) des vecteurs δ_{θ_i} :

$$\Delta = (\delta_{\theta_1}, \delta_{\theta_2}, \dots, \delta_{\theta_M})$$

et F la matrice diagonale (M, M) :

$$F = \text{diag}(\exp[-j\Phi_i])$$

La MIO correspondant à la k -ième sous-antenne s'écrit :

$$(31) \quad \Gamma_k^{(n)} = \Delta \cdot F^{k-1} \cdot \gamma_S \cdot (F^{k-1})^* \cdot \Delta^* + b^2 \cdot I$$

La MIO Γ lissée spatialement se définit alors simplement comme étant la moyenne des MIO des sous-antennes :

$$(32) \quad \Gamma = \frac{1}{N-n+1} \sum_{k=1}^{N-n+1} \Gamma_k^{(n)}$$

Remarque : On peut noter que, sans utiliser la méthode ci-dessus, le fait de projeter orthogonalement la MIO estimée pour l'antenne entière sur le sous-espace des matrices de Toeplitz [15] (ce qui se réalise aisément par moyennage sur chaque diagonale et que nous appelons technique de « rectification ») revient également à effectuer un lissage spatial.

3.2.2. Sensibilité à la corrélation de la MRB

Comme toutes les méthodes HR, la MRB s'applique encore lorsque les sources sont partiellement corrélées. En effet, on peut vérifier que « γ_S de rang maximal » est la seule hypothèse à poser sur γ_S pour déterminer la matrice de transition du système F [donnée par l'équation (28)].

Nous avons pu constater sur de nombreux exemples de simulation que la robustesse de la MRB vis-à-vis de la corrélation des sources est semblable à celle de la méthode du goniomètre.

3.3. ESTIMATION DE LA MATRICE DE CORRÉLATION DES SOURCES

3.3.1. Sources non corrélées

Première méthode

Connaissant une estimation de la position angulaire des sources, on peut reconstruire une matrice \mathcal{D} et étant donnée l'égalité $\mathcal{D} \cdot \gamma_S \cdot \mathcal{D}^* \cdot U_1 \cdot \hat{\Sigma}_1^2 \cdot U_1^*$, calculer une matrice $\hat{\gamma}_S$ (matrice diagonale des puissances des sources).

Deuxième méthode (fondée sur les moments du second ordre)

Lorsque le champ de bruiteurs est celui décrit au paragraphe 3.1, le signal reçu sur le i -ième capteur s'écrit :

$$(33) \quad y_i(t) = \sum_{k=1}^M \alpha_k(t) \times \exp[-j\{(i-1)\Phi_k - 2\pi ft\}] + w_i(t)$$

où α_k est un bruit blanc gaussien centré, de variance σ_k^2 .

Ce qui correspond à la sortie du modèle d'état (3) :

$$y_i = \mathbf{h}^* \cdot \mathbf{X}_i + w_i$$

Par identification des deux équations précédentes, on déduit une expression du vecteur déterministe constant \mathbf{h} et du vecteur d'état \mathbf{X}_i dans la base qui diagonalise la matrice de transition :

$$\mathbf{h} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \mathbf{X}_i = \begin{pmatrix} \alpha_1 \cdot \exp[-j\{(i-1)\Phi_1 - 2\pi ft\}] \\ \alpha_2 \cdot \exp[-j\{(i-1)\Phi_2 - 2\pi ft\}] \\ \vdots \\ \alpha_M \cdot \exp[-j\{(i-1)\Phi_M - 2\pi ft\}] \end{pmatrix}$$

Le choix de cette base permet de calculer aisément les covariances de la sortie du modèle d'état, celles-ci étant indépendantes du choix de la base pour exprimer le modèle.

Les $\{\alpha_k\}_{k=1, \dots, M}$ étant indépendants deux à deux (*hyp. sources non corrélées*), la matrice de variance de

l'état s'écrit :

$$P = E(\mathbf{X}_i \cdot \mathbf{X}_i^*) = \text{diag}(\sigma_k^2).$$

Si de plus on considère l'estimation de la matrice de transition du système :

$$F = \text{diag}(\hat{\rho}_k \cdot \exp[-j\hat{\phi}_k]).$$

On obtient l'expression de la covariance de la sortie du modèle d'état :

$$(34) \quad \Gamma_{-m} = \begin{cases} \sum_{k=1}^M \sigma_k^2 + b^2, & \text{si } m=0 \\ \sum_{k=1}^M \hat{\rho}_k^m \cdot \sigma_k^2 \cdot \exp[j \cdot m \cdot \hat{\phi}_k], & \text{si } m \neq 0. \end{cases}$$

Les puissances des sources, $\{\sigma_k^2\}$, sont donc solutions du système linéaire obtenu ci-dessus.

Si on choisit de prendre tous les coefficients ρ_k égaux à 1, on retrouve la formule d'estimation de la puissance des sources de Pisarenko.

3.3.2. Sources corrélées

On suppose que γ_s est non dégénérée (pas de sources totalement corrélées).

On a vu que l'estimation de la MIO par le périodogramme (lissage temporel) suivi d'une projection sur le sous-espace des matrices de Toeplitz (lissage spatial) permet de nous affranchir des problèmes dus à la corrélation des sources. Mais cette opération a pour effet de transformer la matrice de corrélation des sources γ_s qui est non diagonale et plus ou moins proche de la dégénérescence dans le cas de sources partiellement corrélées, en une matrice de rang maximal qui tend vers une matrice diagonale.

L'estimation de la matrice de corrélation ainsi déformée n'a alors que peu d'intérêt. La dimension de l'espace source ainsi que les gisements des sources étant estimés à partir de la MIO lissée spatialement, il s'agit maintenant d'estimer la matrice de corrélation des sources à partir de la MIO avant lissage spatial.

Soit (V, Δ) le système propre de la MIO non lissée spatialement (ou non rendue de Toeplitz). Δ est la matrice diagonale des valeurs propres rangées par ordre décroissant et V est la matrice des vecteurs propres qui leurs sont associés. Le sous-espace « sources » est engendré par les M vecteurs propres principaux : $V_1 = (V_1, V_2, \dots, V_M)$. On élimine l'effet du bruit, supposé blanc, en soustrayant aux M plus grandes valeurs propres la moyenne des $(N-M)$ plus petites valeurs propres : Δ_1 matrice (M, M) diagonale.

La partie signal contenue dans la MIO s'écrit alors :

$$\mathcal{D} \cdot \gamma_s \cdot \mathcal{D}^* V_1 \cdot \hat{\Delta}_1 \cdot V_1^*.$$

Première méthode

Connaissant les gisements des sources, on construit une matrice \mathcal{D} et on en déduit la matrice de corrélation des sources par pseudo-inverse à partir de l'équation précédente.

Deuxième méthode

On a vu que dans le cas de sources non corrélées, on retrouve la formulation de Pisarenko. En présence de sources corrélées, le problème est différent. En effet, le signal reçu sur l'antenne (la sortie du modèle d'état) n'est plus un processus stationnaire. Sa fonction de corrélation (7) s'exprime en fonction des éléments de la matrice de corrélation des sources et de leurs gisements :

$$\begin{aligned} \Phi(y_{i+m} \cdot y_i^*) &= \sum_{l=1}^M \sum_{k=1}^M \sigma_{lk} \cdot \exp[-j\{m\Phi_l + (i-1)(\Phi_l - \Phi_k)\}]. \end{aligned}$$

En identifiant $\Phi(y_{i+m} \cdot y_i^*)$ aux éléments de la MIO « non bruitée » estimée par (V_1, Δ_1, V_1^*) , on détermine les éléments $\sigma_{l,k}$ de la matrice de corrélation des sources (par résolution du système qui a pour équation générale l'équation précédente).

En fait, cette méthode d'estimation de la matrice de corrélation est équivalente à la précédente.

3.4. STABILITÉ DE LA SOLUTION

Pour assurer la stabilité du modèle, les valeurs propres de la matrice de transition doivent se trouver à l'intérieur du cercle unité [19].

Toutefois, les pôles d'un processus harmonique étant situés sur le cercle unité, on parlera de stabilité marginale [19]; c'est-à-dire, pour qu'un processus soit stable il suffira qu'il ait ses pôles à l'intérieur ou sur le cercle unité. Dans le cadre de l'identification d'un processus harmonique bruité, il est donc normal de trouver des pôles de module inférieur ou égal à 1, mais pas de pôles de module supérieur à 1. Bien sûr, les problèmes de stabilité en traitement d'antenne n'ont pas la même importance (quant à leurs conséquences!) qu'en automatique, mais leur étude permet une justification de la procédure d'identification. En effet, quelle confiance peut-on accorder à une méthode d'identification qui trouve les pôles d'un processus stable à l'extérieur du cercle unité?

L'étude de la stabilité ne paraît donc pas négligeable.

Aussi, nous allons vérifier si les différentes expressions obtenues pour la matrice de transition fournissent des modèles stables (au sens où nous l'avons indiqué ci-dessus, c'est-à-dire, si leurs pôles sont de modules inférieur ou égal à 1).

Propriété 3 : Soit la matrice $[O^* \cdot O^{up}]$ [cf. équation (14)] avec $O = U_1 \cdot T$, où U_1 est une matrice (N, M) constituée des M vecteurs propres principaux de la MIO et T une matrice (M, M) inversible [cf. équation (21)].

Les valeurs propres de $[O^* \cdot O^{up}]$ sont toutes à l'intérieur du cercle unité.

On donne en annexe A.1 les éléments de la démonstration.

Pour une matrice de type $[(O^t)^* \cdot O^t]$ [cf. équation (16), (24), ou encore (28)] on ne peut pas conclure en général quant à la stabilité de la solution obtenue (Annexe B). Toutefois, on remarque que dans deux cas la stabilité est garantie :

1^{er} cas (cas asymptotique, c'est-à-dire lorsque la MIO correspond exactement au modèle) :

Les vecteurs propres de la MIO (vecteurs colonnes de la matrice U_1) et les vecteurs sources $\{D_{\theta_k}\}_{k=1, \dots, M}$ engendrent asymptotiquement le même sous-espace : le sous-espace « source » \mathcal{E}_s . On vérifie alors que les valeurs propres de la matrice de transition F qui sont aussi les valeurs propres de la matrice $[(U_1)^* \cdot U_1]$ sont bien égales aux valeurs propres théoriques, donc de module égal à 1.

2^e cas :

On a pu vérifier sur de nombreux exemples que lorsque la MIO est rendue sous forme Toeplitz la stabilité est assurée. En fait, les cas asymptotiques et Toeplitz sont équivalents pour l'étude de la stabilité.

Remarque : On peut toujours se ramener au cas Toeplitz (par « rectification » par exemple) et donc garantir la stabilité.

3.5. ESTIMATION DE LA MATRICE DE TRANSITION AVEC CONTRAINTE

Dans le but d'assurer la stabilité de la solution sans pour autant imposer à la MIO d'être une matrice de Toeplitz, on peut imposer une contrainte à la matrice de transition lors de la résolution du problème $O^t = O^t \cdot F$.

Soit alors le problème de minimisation avec contrainte :

$$(35) \quad \begin{cases} \min \| O^t - O^t \cdot F \| \\ \text{contrainte : } F \text{ matrice unitaire.} \end{cases}$$

Cette contrainte implique que toutes les valeurs propres de F sont de module égal à 1 (ce qui correspond bien à un processus harmonique) et assure ainsi la stabilité (au sens que nous avons indiqué).

La résolution de ce problème (voir annexe C) conduit à effectuer la SVD de $[(O^t)^* \cdot O^t]$:

$$(36) \quad (O^t)^* \cdot O^t = V \cdot \Delta \cdot Q^*$$

où Δ est la matrice diagonale des valeurs singulières. V et Q sont deux matrices (M, M) unitaires.

On en déduit la matrice de transition du système qui est solution du problème posé :

$$(37) \quad F = V \cdot Q^*$$

Remarque : Toutes les valeurs propres ont un module proche de 1, même si l'on a surestimé le nombre de sources (distinction vrais pôles/faux pôles moins aisée).

On pourrait trouver d'autres contraintes plus « douces » mais celle-ci a l'avantage de rendre la minimisation facilement réalisable et peu coûteuse en temps de calcul.

3.6. MRB : RÉSULTATS DE SIMULATIONS

On se propose de comparer les performances de la MRB à celles du goniomètre.

Pour cela, on considère la fonction inverse de la projection d'un vecteur type « source » D_θ sur le sous-espace bruit :

$$P(\theta) = \frac{1}{\sum_{i=M+1}^N |U_i^* \cdot D_\theta|^2}$$

N étant le nombre de capteurs, M le nombre supposé de sources et $\{U_i\}_{i=M+1, N}$ engendrant le sous-espace bruit.

Les estimés des gisements des sources correspondent alors aux angles θ pour lesquels la courbe $P(\theta)$ présente des pics.

On peut utiliser plus finement l'information contenue dans les vecteurs propres en considérant la fonction rationnelle en z ($z = \exp(-j2\pi fd/c \cos \theta)$) associée à $P(\theta)$:

$$P(z) = \frac{1}{z^{-N+1} \cdot D(z)}$$

avec :

$$D(z) = d_0 + d_1 \cdot z + \dots + d_{N-2} \cdot z^{N-2} + d_{N-1} z^{N-1} + \bar{d}_{N-2} z^N + \dots + \bar{d}_1 z^{2N-3} + \bar{d}_0 z^{2(N-1)}$$

$D(z)$ est un polynôme de degré $2(N-1)$; si z_k est un zéro de $D(z)$ alors \bar{z}_k^{-1} l'est aussi. On appelle pôles du goniomètre les $(N-1)$ zéros de $D(z)$ situés à l'intérieur du cercle unité. On déduit alors une estimation des gisements des sources à partir des pôles du goniomètre les plus proches du cercle unité, les autres pôles correspondant au bruit.

L'étude des pôles du goniomètre est plus coûteuse en temps de calcul que l'étude de la fonction de projection mais améliore la résolution du goniomètre, et rend plus aisée la comparaison avec la MRB car on peut alors adopter la même représentation pour les deux méthodes. En effet, on représente les pôles du goniomètre (respectivement : les pôles du système pour la MRB sur un demi-cercle unité avec pour coordonnées polaires le module de chaque pôle et le gisement estimé correspondant) :

$$z_k \left\{ \begin{array}{l} \text{module}(z_k) \\ \text{Arc. cos} \left[-\frac{c}{2\pi fd} \cdot \text{Argument}(z_k) \right] \end{array} \right.$$

Pour les simulations, on considère une antenne linéaire composée de 16 capteurs équirépartis à $\lambda/2$; le bruit est supposé blanc, gaussien, centré, de densité spectrale de puissance égale à 1. On agira donc sur le rapport signal/bruit en faisant varier la puissance des sources.

La MIO est estimée par le périodogramme moyenné (BT=100) suivi d'une projection orthogonale sur le sous-espace des matrices de Toeplitz (rectification). Ce dernier traitement constitue un lissage spatial (peu

coûteux en temps de calcul) éliminant ainsi la corrélation éventuelle entre sources.

Simulation : On s'intéresse à la détection d'une source faible au milieu de deux sources fortes.

Pour cela, on simule trois sources non corrélées de gisements 56, 60, 65° et de niveaux respectifs -1, -11, -1 dB.

On applique les méthodes en supposant le nombre de sources égal à 4. Sur 100 réalisations la MRB a détecté 40 fois la source faible, et la méthode du goniomètre 15 fois seulement (cf. tableau).

Détection de la source faible	Pôles gonio. 15/100	Pôles MRB 40/100
-------------------------------	---------------------	------------------

Pour une réalisation on présente :

Figure 2 : La fonction de projection du goniomètre.

Les deux pics importants correspondent aux deux sources fortes, la source faible n'est pas vue.

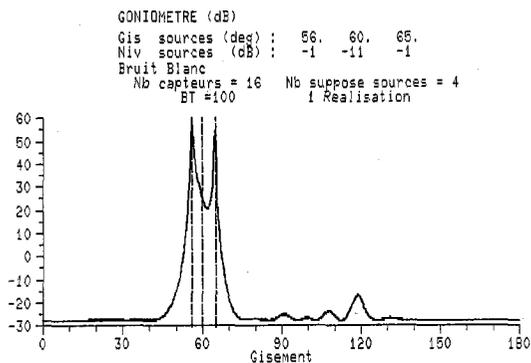


Fig. 2

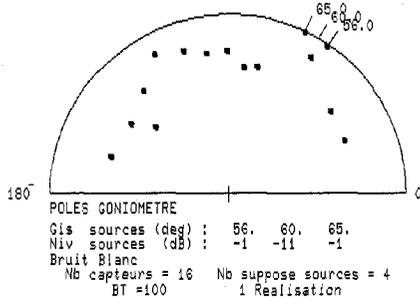


Fig. 3

Figure 3 : Les pôles du goniomètre.

Les pôles correspondants aux sources fortes sont pratiquement sur le cercle unité, le pôle de la source faible est plus à l'intérieur du cercle.

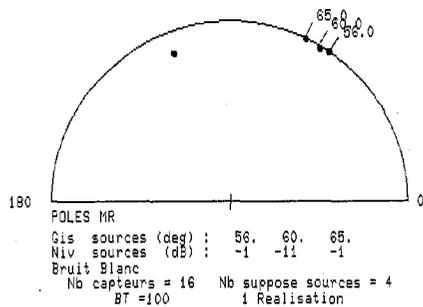


Fig. 4

Figure 4 : Les pôles de la MRB.

Les pôles des trois sources sont quasiment sur le cercle unité.

On constate donc que la représentation polaire des résultats du goniomètre est plus fine que sa fonction de projection mais reste inférieure pour la détection à la MRB.

On peut reprendre cette simulation en augmentant le niveau de la source faible jusqu'à ce que le pourcentage de détection de cette source par le goniomètre soit égal à celui qu'avait le MRB lorsque le niveau de la source faible était -11 dB. Ceci est atteint lorsque la source faible a un niveau de -8 dB.

Dans ce cas de figure, la MRB a donc un gain de 3 dB par rapport au goniomètre.

4. Méthode de réalisation spatiale, cas du bruit additif corrélé (MRC)

4.1. RÉSULTATS GÉNÉRAUX (MODÈLE D'ÉTAT A PHASE MINIMALE). POSITION DU PROBLÈME

4.1.1. Résultats généraux (modèle d'état à phase minimale)

L'approche précédente (MRB) [10] supposait le bruit spatialement blanc. De plus, cette approche ne peut être étendue au cas d'un bruit non blanc même de covariance connue.

On se propose donc de décrire les sorties des capteurs par un modèle ARMA à phase minimale dont le modèle des innovations est de la forme :

$$\begin{cases} \mathbf{X}_{i+1} = \mathbf{F} \cdot \mathbf{X}_i + \mathbf{T} \cdot v_i \\ y_i = \mathbf{h}^* \cdot \mathbf{X}_i + v_i \end{cases} \quad (M)$$

où $\{v_i\}$ est un bruit blanc gaussien centré de d. s. p. b^2 soit :

$$\begin{aligned} \mathbf{Y}_i^+ &= [y_i, y_{i+1}, \dots]^T \\ (38) \quad &\text{le futur des observations au } i\text{-ième capteur, et} \\ \mathbf{Y}_i^- &= [y_{i-1}, y_{i-2}, \dots]^T \\ &\text{le passé de celles-ci.} \end{aligned}$$

L'approche stochastique consiste à déterminer un vecteur d'état \mathbf{X}_i résumant toute l'information passée pour prédire de façon optimale (au sens d'un critère à définir) l'espace des observations futures.

L'état est relié à l'entrée et à la sortie du système par les matrices de commandabilité \mathbf{C} et d'observabilité \mathbf{O} . D'une part, l'état est combinaison linéaire du passé des entrées \mathbf{V}_i^- :

$$(39) \quad \mathbf{X}_i = \mathbf{C} \cdot \mathbf{V}_i^-$$

où :

$$\mathbf{C} = (\mathbf{T}, \mathbf{F} \cdot \mathbf{T}, \mathbf{F}^2 \cdot \mathbf{T}, \dots),$$

$$\mathbf{V}_i^- = [v_{i-1}, v_{i-2}, v_{i-3}, v_{i-3}, \dots]^T \quad (\text{passé des entrées})$$

et

$$\mathbf{V}_i^+ = [v_i, v_{i+1}, v_{i+2}, \dots]^T \quad (\text{futur des entrées}).$$

D'autre part, le futur des observations s'exprime en fonction de l'état et du futur des entrées :

$$(40) \quad Y_i^+ = O \cdot X_i + L \cdot V_i^+$$

avec

$$O = \begin{bmatrix} h^* \\ h^* F \\ h^* \cdot F^2 \\ \vdots \end{bmatrix},$$

$$L = \begin{bmatrix} 1 & 0 & 0 & \dots \\ h^* \cdot T & 1 & 0 & \dots \\ h^* F \cdot T & h^* \cdot T & 1 & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}.$$

L'innovation étant un bruit blanc, le futur des entrées V_i^+ n'est pas corrélé avec l'état X_i . La projection orthogonale du futur des observations sur l'état s'écrit alors simplement :

$$(41) \quad Y_i^+ | X_i = O \cdot X_i.$$

Par ailleurs, le modèle étant supposé à phase minimale, son filtre inverse est stable :

$$(42) \quad \begin{cases} X_{i+1} = (F - T \cdot h^*) X_i + T \cdot y_i \\ v_i = -h^* \cdot X_i + y_i \end{cases}$$

Ce qui se traduit par :

$$(43) \quad X_i = A \cdot Y_i^-$$

où

$$A = (T, (F - T \cdot h^*) \cdot T, (F - T \cdot h^*)^2 \cdot T, \dots).$$

Des équations précédentes on déduit alors également :

$$(44) \quad \begin{cases} Y_i^+ = O \cdot A \cdot Y_i^- + L \cdot V_i^+ \\ Y_i^+ | Y_i^- = O \cdot A \cdot Y_i^- = O \cdot X_i. \end{cases}$$

Propriété 4 : L'état d'un modèle d'état à phase minimale résume le passé des observations ($X_i = A \cdot Y_i^-$) et permet de prédire le futur de celles-ci ($Y_i^+ | Y_i^- \equiv O \cdot X_i$).

Ces quelques rappels relatifs au modèle d'état à phase minimale nous permettent d'introduire deux méthodes de réalisation stochastique fondées sur deux critères distincts.

4.1.2. Position du problème

Les méthodes proposées ici reposent sur une décomposition de l'antenne en deux sous-antennes adjacentes dont les sorties des capteurs après transformée de Fourier (vecteurs d'observations) sont notés Y^+ et Y^- .

Il s'agit alors de déterminer un système linéaire de dimension fixée qui modélise les sorties des capteurs. Pour cela, diverses méthodes peuvent être proposées; toutes sont fondées sur l'estimation de l'état du système qui résume la connaissance d'une sous-antenne (le passé des observations). On voit alors que ces

méthodes diffèrent par leur façon de résumer cette connaissance (critère).

La première méthode (§ 4.2) consiste à minimiser la variance de l'erreur de prédiction d'un vecteur d'observations (première sous-antenne) à partir de l'autre vecteur (deuxième sous-antenne).

La deuxième méthode (§ 4.3) est basée sur la maximisation de l'information mutuelle entre les deux vecteurs d'observations (aléatoires), qui est, par définition [24], la quantité d'information sur un vecteur contenue dans l'autre.

4.2. MÉTHODE DU CRITÈRE D'EFFICACITÉ DE LA PRÉDICTION [11, 12]

4.2.1. Présentation de la méthode

Le critère d'efficacité de la prédiction défini par K. S. Arun, D. V. Bhaskar Rao et S. Y. Kung consiste à chercher le vecteur d'état X_i appartenant à l'espace engendré par le passé des observations Y_i^- (i. e. $X_i = A Y_i^-$ qui minimise la variance de l'erreur de prédiction du futur des observations, i. e. :

$$(45) \quad \text{Min}_{X_i} \{ \text{tr} [\text{Cov}(Y_i^+ - Y_i^+ | X_i)] \}.$$

En tenant compte de l'égalité $X_i = A \cdot Y_i^-$, le problème précédent, revient à déterminer la matrice A qui minimise le critère :

$$(46) \quad \text{Min}_A \{ \text{tr} [\text{Cov}(Y_i^+ - Y_i^+ | A \cdot Y_i^-)] \}.$$

Soit, en utilisant le lemme suivant :

Lemme 2 : La projection orthogonale d'un vecteur a sur un vecteur b, a et b vecteurs aléatoires centrés, s'écrit :

$$a | b = E(a \cdot b^*) \cdot [E(b \cdot b^*)]^{-1} \cdot b,$$

le critère s'écrit :

$$(47) \quad \text{Min}_A \{ \text{tr} [\Gamma^+ - H \cdot A^* \cdot (A \cdot \Gamma^- \cdot A^*)^{-1} \cdot A \cdot H^*] \}$$

avec :

$$\Gamma^+ = E(Y_i^+ \cdot Y_i^{+*}), \quad \Gamma^- = E(Y_i^- \cdot Y_i^{-*}),$$

$$H = E(Y_i^+ \cdot Y_i^{-*})$$

Γ^+ et Γ^- sont les MIO estimées à partir du futur et du passé des observations, H est la matrice de Hankel (des covariances) du système.

Lorsque le système est supposé d'ordre M, le vecteur d'état correspondant a M composantes. Le problème (45) consiste alors à chercher la matrice A de dimension (M, L), où L est la dimension des vecteurs Y_i^+ et Y_i^- qui minimise le critère (47).

Propriété 5 (annexe D) : La solution du problème (45) est obtenue lorsque les lignes de la matrice $(\Sigma_1 \cdot A \cdot \Gamma^- \cdot H^{-1})$ forment une base des M « plus grands » vecteurs propres de la matrice $[H \cdot (\Gamma^-)^{-1} \cdot H^*]$. La matrice Σ_1^2 est la matrice diagonale des M plus grandes valeurs propres qui leur sont associées.

On notera donc :

$$(48) \quad H. (\Gamma^-)^{-1}. H^* = U. \Sigma^2. U^*$$

avec :

$$U = [U_1 | U_2]$$

la matrice des vecteurs propres, et

$$\Sigma^2 = \begin{bmatrix} \Sigma_1^2 & 0 \\ 0 & \Sigma_2^2 \end{bmatrix}$$

la matrice diagonale des valeurs propres rangées par ordre de module décroissant.

On a alors :

$$(49) \quad \Sigma_1. A. \Gamma^- . H^{-1} = U_1^*$$

On en déduit la matrice qui minimise le critère (47) :

$$(50a) \quad A = \Sigma_1^{-1}. U_1^*. H. (\Gamma^-)^{-1}.$$

Remarque : On vérifie aisément que le critère d'efficacité de la prédiction est invariant par changement de base. En effet, par la matrice de changement de base B, la réalisation (F, T, h*) devient (BFB⁻¹, BT, h*B⁻¹).

La matrice maximisant le critère :

$$A = (T, (F - Th^*)T, (F - Th^*)^2T, \dots)$$

est alors transformée en :

$$(BT, B(F - Th^*)T, B(F - Th^*)^2T, \dots) = BA.$$

B étant de rang maximal, il est évident que la matrice (BA) minimise également le critère d'efficacité de la prédiction :

$$\begin{aligned} \text{Min}_{(BA)} \{ \text{tr} [\Gamma^+ - H(BA)^*] \\ \times [(BA) \Gamma^- (BA)^*]^{-1} (BA) H^* \} \\ = \text{Min}_{(BA)} \{ \text{tr} [\Gamma^+ - HA^* . (A \Gamma^- A^*)^{-1} AH^*] \}. \end{aligned}$$

On voit donc que l'on peut minimiser le critère dans une base quelconque et en déduire les valeurs propres (invariantes) de la matrice de transition du système.

La solution générale du problème de minimisation s'écrit donc :

$$(50b) \quad A = B. \Sigma_1^{-1}. U_1^*. H. (\Gamma^-)^{-1}$$

avec B matrice (M, M), de rang M, *arbitraire*.

A partir de cette matrice, on détermine la matrice de transition du système par diverses méthodes.

4.2.2. Détermination de la matrice de transition du système

Méthode utilisant la matrice d'observabilité

Considérons maintenant la projection orthogonale du futur des observations sur le passé de celles-ci :

$$(51) \quad Y_i^+ | Y_i^- = H. (\Gamma^-)^{-1} Y_i^-.$$

Or, on a vu au paragraphe 4.1 que l'on avait :

$$(44) \quad Y_i^+ | Y_i^- \equiv O. A. Y_i^-.$$

On en déduit alors des deux équations précédentes :

$$(52) \quad O. A = H. (\Gamma^-)^{-1}$$

$$(53) \quad \Rightarrow O = U_1. \Sigma_1. B^{-1}.$$

Ce qui nous permet de déterminer la matrice de transition du système :

$$F = (O^\dagger)^* . O^\dagger$$

$$(54) \quad F = B. \Sigma_1^{-1} (U_1^\dagger)^* . U_1^\dagger . \Sigma_1 . B^{-1}$$

On a vu, dans le cas de la MRB, que ce type d'estimateur de F ne fournissait pas nécessairement un modèle stable. Aussi, on s'intéresse à d'autres méthodes d'estimation de la matrice de transition.

Méthode de pénalité

On cherche F tel que : $O^\dagger . F = O^\dagger$.

Nous ne revenons pas sur l'approximation : $O . F \simeq O^{up}$ (introduction d'une ligne de zéros), qui, comme dans le cas de la MRB, fournit un modèle stable (cf. annexe A) mais moins précis.

On détermine ici F comme étant la solution d'un problème de minimisation avec pénalité sur sa norme :

$$(55) \quad \text{Min}_F \{ \| O^\dagger - O^\dagger F \|^2 + \lambda. \| F \|^2 \}$$

λ donné

La différentielle en F de $\{ \| O^\dagger - O^\dagger F \|^2 + \lambda. \| F \|^2 \}$ appliquée au point K s'annule lorsque :

$$(56) \quad \begin{aligned} -\text{tr}(K^* . O^{\dagger*} . O^\dagger) + \text{tr}(K^* . O^{\dagger*} . O^\dagger F) \\ + \lambda. \text{tr}(K^* . F) = 0 \\ \forall K \in \mathcal{M}_C(M, M) \end{aligned}$$

Soit

$$(57) \quad \text{tr}[K^* . (-O^{\dagger*} . O^\dagger + O^{\dagger*} . O^\dagger F + \lambda. F)] = 0,$$

∀ K

L'application :

$$K \rightarrow \text{tr}[K^* . (-O^{\dagger*} . O^\dagger + O^{\dagger*} . O^\dagger F + \lambda. F)]$$

est une forme linéaire de $\mathcal{M}_C(M, M)$; on a en particulier pour $K = (-O^{\dagger*} . O^\dagger + O^{\dagger*} . O^\dagger F + \lambda. F)$:

$$(58) \quad \text{tr}(K^* . K) = 0.$$

Or, l'application : $K \rightarrow \text{tr}(K^* . K)$ est une forme hermitienne définie positive sur $\mathcal{M}_C(M, M)$. Ce qui implique :

$$(59) \quad -O^{\dagger*} . O^\dagger + O^{\dagger*} . O^\dagger F + \lambda. F = O_p$$

$$(60) \quad \Leftrightarrow F = [(O^{\dagger*} . O^\dagger + \lambda. I_p)^{-1} . (O^{\dagger*} . O^\dagger)$$

Lors du calcul de la matrice de transition, on peut donc choisir λ tel que les valeurs propres de cette matrice soient à l'intérieur du cercle unité. La valeur du paramètre λ qui place le pôle de plus grand module à proximité du cercle unité varie suivant la configuration du problème.

Il est important de noter que l'on repousse tous les pôles à l'intérieur du cercle unité sans faire la distinction vrais pôles/faux pôles.

Méthode de projection

On peut déterminer la matrice de transition du système à partir des dynamiques de l'état du système :

$$(61) \quad X_{i+1} | X_i = F \cdot X_i$$

En utilisant le théorème de la projection orthogonale, il vient :

$$(62) \quad F = E(X_{i+1}, X_i^*) \cdot [E(X_i, X_i^*)]^{-1}$$

$$X_j = A \cdot Y_j^-; \quad A = \Sigma_1^{-1} \cdot U_1^* \cdot H \cdot (\Gamma^-)^{-1}$$

En remplaçant dans l'équation précédente on obtient :

$$(63) \quad F = \Sigma_1^{-1} \cdot U_1^* \cdot H \cdot (\Gamma^-)^{-1} \cdot \overleftarrow{\Gamma}^- \cdot (\Gamma^-)^{-1} \cdot H^* \cdot U_1 \cdot \Sigma_1^{-1}$$

avec

$$\overleftarrow{\Gamma}^- = E(Y_{i+1}^-, Y_i^-)$$

ou encore, en utilisant la décomposition en éléments propres de $H \cdot (\Gamma^-)^{-1} \cdot H^*$:

$$F = (\Sigma_1 \cdot U_1^* \cdot H^{-*} \cdot \overleftarrow{\Gamma}^- \cdot H^{-1} \cdot U_1 \cdot \Sigma_1)$$

Pratiquement, $\overleftarrow{\Gamma}^-$ se déduit de Γ^- en décalant les colonnes de Γ^- d'une colonne vers la gauche.

Lorsque l'on introduit une colonne de zéros pour première colonne à droite de $\overleftarrow{\Gamma}^-$, on montre (cf. annexe A.2) que la solution ainsi obtenue est stable. Mais cette solution n'a que peu d'intérêt car ici on connaît la suite des covariances permettant de construire $\overleftarrow{\Gamma}^-$.

On obtient alors une solution dont on ne peut pas garantir la stabilité. Toutefois, cette technique de détermination de la matrice de transition par projection a un effet stabilisant, et parmi les nombreuses simulations que nous avons effectuées nous avons pu constater que le modèle ainsi obtenu était toujours stable.

4.2.3. Remarques

1. Le grammien d'observabilité est diagonal : pour $B=I$,

$$O^* \cdot O = \Sigma_1^2$$

2. Si on calcule la matrice de covariance de $Y_i^+ | Y_i^-$ pour chacune des expressions (44) et (51), on obtient :

$$\text{Var}[H \cdot (\Gamma^-)^{-1} \cdot Y_i^-] = H \cdot (\Gamma^-)^{-1} \cdot H^*$$

$$\text{Var}[O \cdot X_i] = O \cdot P_i \cdot O^*$$

avec $P_i = \text{Var}(X_i)$.

Si la sortie du système linéaire (4) est un processus stationnaire $P_i = P$ et

$$\text{Var}[O \cdot X_i] = O \cdot P \cdot O^*$$

Soit :

$$O \cdot P \cdot O^* = H \cdot (\Gamma^-)^{-1} \cdot H^*$$

Cette expression laisse apparaître le lien existant entre la MRC et MRB.

4.3. UTILISATION DES COEFFICIENTS DE CORRÉLATION CANONIQUES [13]

La méthode proposée par U. B. Desai et D. Pal consiste à utiliser les coefficients de corrélation canoniques entre le futur et le passé des observations.

4.3.1. Variables canoniques [17]

On peut introduire les coefficients de corrélation canoniques de la façon suivante : soient Y_j^+ et Y_j^- les vecteurs du futur et du passé des observations, et Γ^+ , Γ^- leurs matrices de covariance. On définit :

$$\varepsilon^+ = (\Gamma^+)^{-1/2} \cdot Y_j^+, \quad \varepsilon^- = (\Gamma^-)^{-1/2} \cdot Y_j^-$$

$$E[\varepsilon^+ \cdot (\varepsilon^+)^*] = E[\varepsilon^- \cdot (\varepsilon^-)^*] = I_L$$

Soit la décomposition en valeurs singulières de la matrice :

$$\Omega = E(\varepsilon^+ \cdot (\varepsilon^-)^*) = (\Gamma^+)^{-1/2} \cdot H [(\Gamma^-)^{-1/2}]^*$$

$$(64) \quad (\Gamma^+)^{-1/2} \cdot H [(\Gamma^-)^{-1/2}]^* = U \cdot \Sigma^2 \cdot V^*$$

avec U, V matrices (L, L) unitaires et Σ^2 matrice diagonale des valeurs singulières $\Sigma^2 = \text{diag}(\sigma_i^2)$.

Ou encore, en ne retenant que les l valeurs singulières non nulles :

$$(\Gamma^+)^{-1/2} \cdot H [(\Gamma^-)^{-1/2}]^* = \tilde{U} \cdot \tilde{\Sigma}^2 \cdot \tilde{V}^*$$

avec

$$\tilde{U}^* \cdot \tilde{U} = \tilde{V}^* \cdot \tilde{V} = I_L$$

et

$$\tilde{\sigma}_1^2 = \sigma_1^2 \geq \dots \geq \tilde{\sigma}_l^2 = \sigma_l^2 > 0 = \sigma_{l+1}^2 = \dots = \sigma_L^2$$

(les valeurs singulières sont positives par définition de la SVD).

Les variables canoniques α et β sont alors définies par :

$$\alpha = \tilde{U}^* \cdot \varepsilon^+, \quad \beta = \tilde{V}^* \cdot \varepsilon^-$$

et on a :

$$E(\alpha \cdot \alpha^*) = E(\beta \cdot \beta^*) = I_l$$

$$E(\alpha \cdot \beta^*) = \tilde{\Sigma}^2$$

Les éléments de $\tilde{\Sigma}^2$ sont appelés les coefficients de corrélation canoniques.

Le vecteur de variables aléatoires orthonormales α (respectivement : β) contient toute l'information sur

Y_i^- (resp. : Y_i^+) qui est présente dans Y_i^+ (resp. : Y_i^-).

Propriété 6 : L'information mutuelle entre Y_i^+ et Y_i^- s'exprime en fonction des coefficients de corrélation canoniques :

$$(65) \quad I(Y_i^+, Y_i^-) = I(\alpha, \beta) = - \sum_{k=1}^l \text{Log}(1 - \sigma_k^4).$$

Pour cela, on définit l'information relative de deux vecteurs comme la somme de l'incertitude (ou entropie) de chacun des vecteurs moins l'incertitude conjointe :

$$(66) \quad I(Y_i^+, Y_i^-) = H(Y_i^+) + H(Y_i^-) - H\left(\frac{Y_i^+}{Y_i^-}\right).$$

Dans l'hypothèse gaussienne :

$$\begin{aligned} H(Y_i^+) &= \text{Log dét}(\Gamma^+) \\ H(Y_i^-) &= \text{Log dét}(\Gamma^-) \end{aligned}$$

et

$$\begin{aligned} H\left(\frac{Y_i^+}{Y_i^-}\right) &= \text{Log dét} \begin{pmatrix} \Gamma^+ & H \\ H^* & \Gamma^- \end{pmatrix} \\ &= \text{Log dét}(\Gamma^+) + \text{Log dét}(\Gamma^- - H^* \cdot (\Gamma^+)^{-1} \cdot H) \\ &= \text{Log dét}(\Gamma^+) + \text{Log dét}(\Gamma^-) \\ &\quad + \text{Log dét}(I_L - \Omega^* \cdot \Omega) \end{aligned}$$

soit :

$$\begin{aligned} I(Y_i^+, Y_i^-) &= - \text{Log dét}(I_L - \Omega^* \cdot \Omega) \\ \Rightarrow I(Y_i^+, Y_i^-) &= - \sum_{k=1}^l \text{Log}(1 - \sigma_k^4) \end{aligned}$$

De même :

$$\begin{aligned} I(\alpha, \beta) &= H(\alpha) + H(\beta) - H\left(\frac{\alpha}{\beta}\right) \\ &= - \text{Log dét} \begin{pmatrix} I_l & \Sigma^2 \\ \Sigma^2 & I_l \end{pmatrix} \\ &= - \text{Log dét}(I_l - \Sigma^4) \\ &= - \sum_{k=1}^l \text{Log}(1 - \sigma_k^4). \end{aligned}$$

Par ailleurs, pour $M < l$ si on a $\sigma_M^2 \gg \sigma_{M+1}^2$, alors les M premiers coefficients de corrélation canoniques résument l'essentiel de l'information et l'on peut effectuer une réduction de modèle.

4.3.2. Algorithme de réalisation canonique

On peut présenter cette méthode comme la maximisation de l'information mutuelle entre le vecteur du futur des observations Y_i^+ et le vecteur d'état du système X_i :

$$(67) \quad \text{Max}_{X_i} I(Y_i^+, X_i)$$

avec :

$$I(Y_i^+, X_i) = H(Y_i^+) + H(X_i) - H\left(\frac{Y_i^+}{X_i}\right).$$

Or, le vecteur d'état d'un système linéaire à phase minimale est relié au passé des observations : $X_i = A \cdot Y_i^-$, où A est une matrice (M, L) . Le problème revient donc à déterminer la matrice A de dimension (M, L) qui maximise $I(Y_i^+, A \cdot Y_i^-)$.

Dans l'hypothèse gaussienne :

$$\begin{aligned} H(Y_i^+) &= \text{Log dét}(\Gamma^+) \\ H(A \cdot Y_i^-) &= \text{Log dét}(A \cdot \Gamma^- \cdot A^*) \end{aligned}$$

et

$$\begin{aligned} H\left(\frac{Y_i^+}{A \cdot Y_i^-}\right) &= \text{Log dét} \left(\frac{\Gamma^+}{A \cdot H^*} \middle| \frac{H \cdot A^*}{A \cdot \Gamma^- \cdot A^*} \right) \\ &= \text{Log dét} \left(\frac{A \cdot \Gamma^- \cdot A^*}{H \cdot A^*} \middle| \frac{A \cdot H^*}{\Gamma^+} \right) \\ &= \text{Log dét}(A \cdot \Gamma^- \cdot A^*) \\ &\quad + \text{Log dét}(\Gamma^+ - H \cdot A^* \cdot (A \cdot \Gamma^- \cdot A^*)^{-1} \cdot A \cdot H^*) \end{aligned}$$

Par conséquent :

$$(68) \quad I(Y_i^+, A \cdot Y_i^-) = \text{Log dét}(\Gamma^+) - \text{Log dét}(\Gamma^+ - H \cdot A^* \cdot (A \cdot \Gamma^- \cdot A^*)^{-1} \cdot A \cdot H^*)$$

Il s'agit donc de déterminer la matrice A solution du problème de minimisation suivant :

$$(69) \quad \text{Min}_{A(M, L)} \{ \text{dét}(\Gamma^+ - H \cdot A^* \cdot (A \cdot \Gamma^- \cdot A^*)^{-1} \cdot A \cdot H^*) \}$$

puisque le log est une fonction continue croissante sur \mathbb{R}^*_{+} .

On peut remarquer la similitude entre les problèmes (47) et (69) qui cependant nécessitent une résolution tout à fait différente à cause des opérateurs qui sont utilisés (trace et déterminant).

Propriété 7 (annexe E) : La solution du problème (67) est obtenue lorsque les lignes de la matrice $(A \cdot (\Gamma^-)^{1/2})$ forment une base des M vecteurs singuliers à droite de la matrice $\{(\Gamma^+)^{-1/2} \cdot H \cdot [(\Gamma^-)^{-1/2}]^*\}$ correspondants aux M plus grandes valeurs singulières.

On notera donc :

$$(70) \quad (\Gamma^+)^{-1/2} \cdot H \cdot [(\Gamma^-)^{-1/2}]^* = [U_1 \quad U_2] \begin{bmatrix} \Sigma_1^2 & 0 \\ 0 & \Sigma_2^2 \end{bmatrix} \begin{bmatrix} V_1^* \\ V_2^* \end{bmatrix}$$

avec

$$\begin{aligned} \sigma_1^2 \geq \sigma_2^2 \geq \dots \geq \sigma_M^2 \geq \sigma_{M+1}^2 \geq \dots \geq \sigma_L^2 \\ \Sigma_1^2 = \text{diag}(\sigma_i^2)_{i=1, M} \\ \Sigma_2^2 = \text{diag}(\sigma_i^2)_{i=M+1, L} \end{aligned}$$

On a alors :

$$(71) \quad A \cdot (\Gamma^-)^{1/2} = V_1^*$$

On en déduit la matrice qui minimise le critère (69) :

$$(72a) \quad A = V_1^* (\Gamma^-)^{-1/2}$$

Comme pour le critère d'efficacité de la prédiction, on vérifie aisément que le critère (69) est invariant par changement de base. On en déduit l'expression générale du problème de minimisation :

$$(72b) \quad A = B \cdot V_1^* (\Gamma^-)^{-1/2}$$

avec B matrice (M, M), de rang M, *arbitraire*.

L'algorithme de réalisation stochastique fournit donc la matrice reliant le passé des observations à l'état du système (à phase minimale). De cette matrice, on déduit de diverses façons la matrice de transition du système.

4.3.3. Détermination de la matrice de transition du système

Méthode utilisant la matrice d'observabilité

La comparaison des deux expressions de la projection du futur sur le passé des observations :

$$Y_i^+ | Y_i^- = O \cdot A \cdot Y_i^-$$

et

$$Y_i^+ | Y_i^- = H \cdot (\Gamma^-)^{-1} \cdot Y_i^-$$

nous permet d'écrire :

$$O \cdot A = H \cdot (\Gamma^-)^{-1}$$

en utilisant (72b) on obtient :

$$(73) \quad O = H \cdot [(\Gamma^-)^{-1/2}]^* \cdot V_1 \cdot B^{-1}$$

Or : $F = (O^T)^* \cdot O^T$, on peut donc en déduire la matrice de transition du système F.

Connaissant une estimation de la séquence des covariances de longueur 2L, on peut se restreindre à la dimension (L-1), ce qui nous permet de construire les matrices $H^T = H$ et H^T de dimension (L-1, L-1).

La matrice de transition s'écrit alors :

$$(74) \quad F = B \cdot V_1^* \cdot [(\Gamma^-)^{1/2}]^* \cdot (H^T)^{-1} \cdot H^T \cdot [(\Gamma^-)^{-1/2}]^* \cdot V_1 \cdot B^{-1}$$

où V_1 est obtenue par SVD de la matrice $\{(\Gamma^+)^{-1/2} \cdot H \cdot [(\Gamma^-)^{-1/2}]^*\}$ de dimension (L-1, L-1).

Méthode de projection

On cherche ici à extraire la matrice de transition à partir des dynamiques de l'état du système.

Soit $X_i = A \cdot Y_i^-$ le vecteur d'état du système.

La matrice de transition d'un système dynamique vérifiant : $X_{i+1} | X_i = F \cdot X_i$, on en déduit par le théorème de la projection orthogonale :

$$F = E(X_{i+1} \cdot X_i^*) \cdot [E(X_i X_i^*)]^{-1}$$

En utilisant (72b), on obtient l'expression de la matrice de transition suivante :

$$(75) \quad F = B \cdot V_1^* \cdot (\Gamma^-)^{-1/2} \cdot \overleftarrow{\Gamma}^- \cdot [(\Gamma^-)^{-1/2}]^* \cdot V_1 \cdot B^{-1}$$

Γ^- se déduit de $\overleftarrow{\Gamma}^-$ en décalant les colonnes de Γ^- d'une colonne vers la gauche.

Ici, on connaît une suite de covariance suffisamment longue pour construire exactement $\overleftarrow{\Gamma}^-$.

De l'expression (75), on déduit aisément les valeurs propres de F puisque :

$$\text{spectre}(F) = \text{spectre}(V_1^* \cdot (\Gamma^-)^{-1/2} \cdot \overleftarrow{\Gamma}^- \cdot [(\Gamma^-)^{-1/2}]^* \cdot V_1)$$

4.4. COMPARAISON DES DEUX ALGORITHMES DE RÉALISATION

Les deux méthodes proposées reposent sur la détermination de la matrice A reliant l'état du système au passé des observations : $X_i = A \cdot Y_i^-$.

L'hypothèse de phase minimale pour le modèle d'état apparaît donc fondamentale. Seul le critère pour déterminer cette matrice diffère, car ensuite pour chacune des méthodes on peut en déduire la matrice de transition du système soit à partir de la matrice d'observabilité, soit à partir des dynamiques de l'état (par projection).

De plus, en comparant les équations (47) et (69), on s'aperçoit que les critères des deux méthodes portent sur la même matrice :

$$[\Gamma^+ - H \cdot A^* \cdot (A \cdot \Gamma^- \cdot A^*)^{-1} \cdot A \cdot H^*]$$

mais les mesures d'information sont différentes (trace ou déterminant).

Il y a donc une certaine similitude entre les deux méthodes, et l'équivalence entre les résultats de simulation que nous avons obtenus pour chaque méthode conforte cette position.

4.5. LIEN AVEC LE TRAITEMENT D'ANTENNE

On suppose que le signal reçu sur chaque capteur de l'antenne est stationnaire et composé d'une partie « utile », qui est la somme de M ondes planes dont les amplitudes sont fonction de variables aléatoires, et d'un bruit additif descriptible spatialement par un modèle ARMA (Q, Q), cette modélisation du bruit étant tout à fait réaliste [15] :

$$s(i) = s_u(i) + s_b(i)$$

(signal utile) bruit

i étant le numéro de capteur.

Le bruit étant supposé statistiquement indépendant du signal utile, la fonction d'autocorrélation de leur somme est égale à la somme des deux fonctions d'autocorrélation :

$$r(k) = r_u(k) + r_b(k)$$

La densité spatiale du processus de sortie des capteurs étant définie comme la transformée de Fourier de sa fonction d'autocorrélation, on l'exprime en fonction de la variable z :

$$\Gamma(z) = \Gamma_u(z) + \Gamma_b(z)$$

avec

$$z = \exp(-j2\pi f d/c \cos \theta)$$

Le signal utile étant la somme de M ondes planes, on peut le représenter par un modèle AR d'ordre M dont les pôles sont tous de module égal à 1 et de fonction de transfert :

$$H_u(z) = \frac{1}{A_u(z)}$$

[$A_u(z)$: polynôme en z d'ordre M].

Sa densité spatiale s'écrit alors :

$$\Gamma_u(z) = H_u(z) \cdot \bar{H}_u(z^{-1})$$

$\Gamma(z) =$

$$\frac{A_b(z) \cdot A_b(z^{-1}) + \sigma_b^2 \cdot B_b(z) \cdot B_b(z^{-1}) \cdot A_u(z) \cdot A_u(z^{-1})}{A_u(z) \cdot \bar{A}_u(z^{-1}) \cdot A_b(z) \cdot \bar{A}_b(z^{-1})}$$

soit :

$$\Gamma_u(z) = \frac{1}{A_u(z) \cdot \bar{A}_u(z^{-1})}$$

Par ailleurs, le bruit de modèle ARMA (Q, Q) a pour fonction de transfert et densité spatiale

$$H_b(z) = \frac{B_b(z)}{A_b(z)}$$

$$\Gamma_b(z) = \sigma_b^2 \cdot \frac{B_b(z) \cdot \bar{B}_b(z^{-1})}{A_b(z) \cdot \bar{A}_b(z^{-1})}$$

où : $A_b(z)$, $B_b(z)$ sont deux polynômes en z d'ordre Q, et σ_b^2 est la variance du bruit blanc d'entrée du filtre formeur du bruit.

La densité spatiale du processus de sortie des capteurs s'écrit alors :

L'ensemble des pôles de la densité spatiale se trouvant sur le disque unité est donc la réunion des M pôles (sur le cercle unité) du signal utile et des Q pôles (à l'intérieur du cercle unité) du bruit.

Considérons maintenant un modèle spatial d'ordre P décrivant les sorties des capteurs :

$$\begin{cases} \mathbf{X}_{i+1} = \mathbf{F} \cdot \mathbf{X}_i + \mathbf{T} \cdot v_i \\ y_i = \mathbf{h}^* \cdot \mathbf{X}_i + v_i \end{cases}$$

(σ_v^2 = variance de v_i).

La fonction de transfert du modèle et la densité spatiale du processus de sortie du système linéaire s'exprime en fonction des paramètres du système [19].

$$H_m(z) = \mathbf{h}^* \cdot (z \cdot \mathbf{I}_p - \mathbf{F})^{-1} \cdot \mathbf{T} + 1$$

$$\Gamma_m(z) = \sigma_v^2 \cdot H_m(z) \cdot H_m(z^{-1})$$

$H_m(z)$ étant invariant par changement de base, on montre que les pôles du modèle sont les valeurs propres de la matrice de transition F [19].

Lorsque le signal reçu sur les capteurs est effectivement modélisable spatialement par le système

linéaire (4), il y a égalité entre les deux expressions de la densité spatiale : $\Gamma(z) = \Gamma_m(z)$.

Ce qui implique que les valeurs propres de la matrice de transition F sont les pôles du signal reçu, les valeurs propres de module égal à 1 correspondant aux pôles du signal utile et les autres valeurs propres correspondant au bruit. Il suffit alors d'appliquer un algorithme de réalisation pour déterminer les valeurs propres de F. On ne retient que les valeurs propres de module proche de 1 et on en déduit les gisements des sources correspondantes.

La MRC permet donc d'estimer les gisements en présence de bruit spatialement corrélé puisqu'il en est tenu compte dans la modélisation.

4.6. MRC : RÉSULTATS DE SIMULATIONS

Comme dans le cas de la MRB, on compare les performances de la MRC à celles du goniomètre (fonction de projection et pôles du goniomètre).

On considère une antenne linéaire composée de 20 capteurs équidistants à $\lambda/2$.

On estime la MIO par le périodogramme moyenné (BT=300) suivi d'une rectification (par moyennage de chaque diagonale pour rendre la matrice de forme Toeplitz).

Le bruit ici est supposé corrélé spatialement le long de l'antenne. On se propose de simuler un bruit descriptible spatialement par un modèle MA d'ordre 2 et de coefficients (0,6; -0,8). Les caractéristiques de ce bruit sont représentées sur la figure 5.

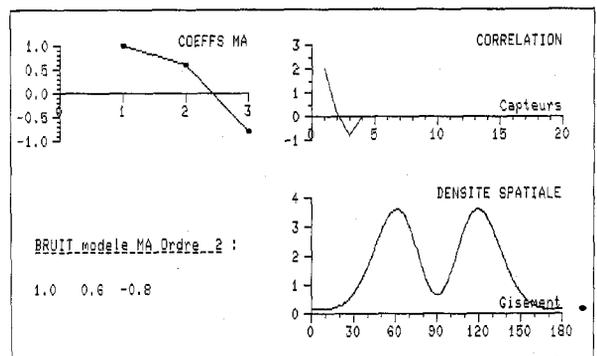


Fig. 5

La variance du bruit blanc d'entrée du filtre formeur de ce bruit MA est égale à 1. La fonction de corrélation du modèle MA représente la corrélation du bruit le long de l'antenne. On constate que ce bruit est corrélé sur 3 capteurs consécutifs, et que cette corrélation spatiale entraîne une coloration spatiale [15] qui se traduit par 2 pics à 60 et 120° dans la densité spatiale du bruit.

Simulation 1

Pour mettre en évidence les performances de la MRC, on simule 3 sources de gisements 90, 110, 160° et de niveaux respectifs -10, -3, -10 dB. La source à 110° est sur une bosse de la densité spatiale du bruit

tandis que les autres sources sont dans des creux de cette densité. Étant donné l'équivalence dans les résultats des algorithmes proposés, on présente uniquement les résultats obtenus par la méthode utilisant les coefficients de corrélation canoniques et on calcule la matrice de transition du système F par projection. Cette dernière technique d'estimation de la matrice de transition présente l'avantage de fournir un modèle stable (pôles du système à l'intérieur du cercle unité; à l'inverse, l'estimation de F à partir de la matrice d'observabilité ne garantit pas la stabilité). Cela revient donc à calculer F à partir de l'équation (75).

On applique les méthodes en supposant le nombre de sources égal à 5, et on présente 10 réalisations successives pour avoir un aspect plus statistique des résultats.

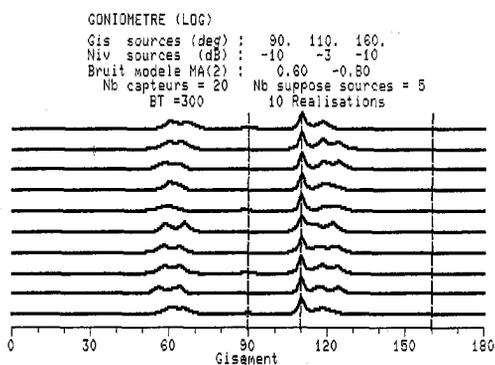


Fig. 6

Figure 6 : La fonction de projection du goniometre nous permet de détecter la source de gisement 110° mais pas les autres. De plus, il apparaît des pics parasites à 60 et 120° correspondants aux pics de la densité spatiale du bruit.

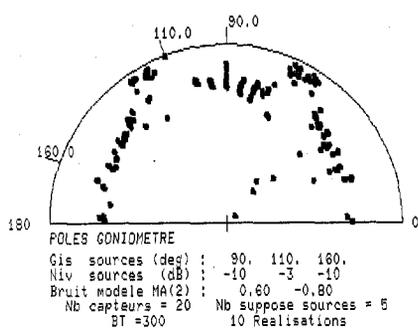


Fig. 7

Figure 7 : Les pôles du goniometre ne nous apportent pas d'informations supplémentaires. La source à 110° est bien détectée (pôles sur le cercle unité) mais les deux autres ne le sont pas. Les pôles parasites (car proches du cercle unité) à 60 et 120° sont également dus aux pics de la densité spatiale du bruit.

Figure 8 : Cette figure comme la précédente est constituée d'un empilement de 10 figures correspondantes aux 10 réalisations successives. La méthode employée

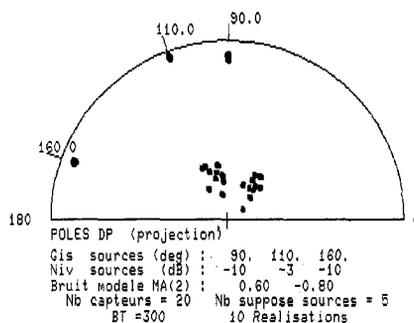


Fig. 8

est celle décrite plus haut [estimation de F par l'équation (75)]. Ici, on détecte bien les 3 sources, et les pôles parasites dus aux pics de la densité spatiale du bruit sont rejetés à l'intérieur du cercle unité.

Les gains apportés ici par la MRC sont donc les suivants :

- élimination des pôles parasites dus aux pics de la densité spatiale du bruit;
- détection des sources faibles situées dans le creux de la densité spatiale du bruit.

Simulation 2

Considérons le même modèle de bruit et intéressons-nous plus particulièrement à la détection d'une source faible située dans un creux de la densité spatiale du bruit. On simule la source de gisement 160° et de niveau -10 dB; on applique les méthodes en supposant le nombre de sources égal à 2.

- Figure 9 : fonction de projection du goniometre.
- Figure 10 : pôles du goniometre.
- Figure 11 : pôles de la MRC.

Les résultats que l'on obtient montrent que le goniometre n'arrive pas à détecter la source simulée et présente des pics (ou des pôles) parasites à 60 et 120° alors que la MRC détecte la source à 160° (pôles proches du cercle unité) et élimine les pôles parasites (rejetés à l'intérieur du cercle unité).

On peut reprendre cette simulation en augmentant le niveau de la source jusqu'à ce que le goniometre la détecte (c'est-à-dire que les pôles du goniometre correspondants à la source soient plus proches du cercle unité que les pôles parasites dus aux pics de la

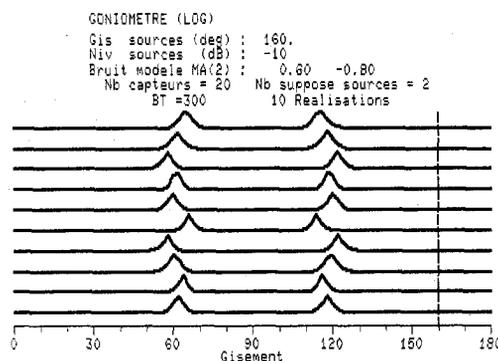


Fig. 9

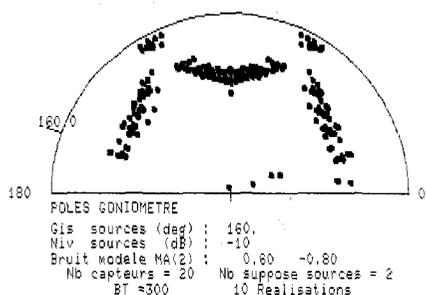


Fig. 10

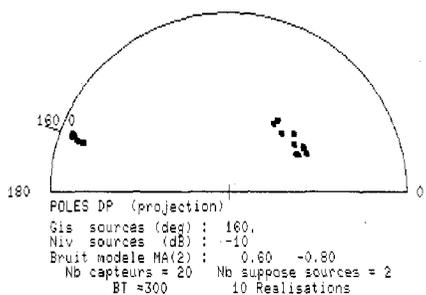


Fig. 11

densité spatiale du bruit). Ceci est atteint lorsque la source a un niveau de -7 dB.

Dans ce cas de figure, la MRC a donc un gain de 3 dB par rapport au goniomètre.

5. Conclusion

La représentation des sorties de capteurs par un système linéaire s'est avérée très propice à la prise en compte dans le traitement d'hypothèses physiques fondamentales en acoustique sous-marine.

D'une part, une meilleure utilisation de l'hypothèse onde plane par son inclusion dans le modèle même de l'état du système a conduit à une amélioration de la détection des sources faibles relativement au goniomètre.

D'autre part, la prise en compte simultanée de l'hypothèse onde plane et de la corrélation spatiale du bruit à travers un système linéaire à minimum de phase nous a permis de séparer la partie prédictive du signal reçu de ce qui ne l'est pas, et de faire la distinction entre les sources réelles et parasites en considérant le module des pôles du modèle estimé. On a ainsi obtenu une méthode d'estimation de la position angulaire de sources en présence de bruit corrélé spatialement et sans estimation du modèle de bruit, ce qui apparaît particulièrement économique en temps de calcul.

Dans ce cas, les gains obtenus par rapport au goniomètre sont les suivants :

- élimination des pôles parasites dus aux pics de la densité spatiale du bruit;
- meilleure détection des sources faibles situées dans les creux de la densité spatiale du bruit.

Par ailleurs, le coût en calcul des méthodes présentées ici est comparable à celui du calcul des pôles du goniomètre (calcul qui améliore le pouvoir de détection du goniomètre relativement à sa fonction de projection). En effet, une fois le sous-espace source estimé (volume de calcul égal à celui du goniomètre), l'estimation des gisements des sources par les méthodes de réalisation nécessite le calcul d'un produit matriciel et des valeurs propres de la matrice produit ainsi obtenue qui est de dimension égale à la dimension supposée de l'espace source (donc de dimension réduite!). La charge en calculs paraît donc tout à fait supportable au vu des gains obtenus.

Manuscrit reçu le 29 avril 1989.

BIBLIOGRAPHIE

- [1] W. S. BURDIC, *Underwater acoustic system analysis*, Prentice Hall, 1984.
- [2] J. CAPON, *Maximum-Likelihood Spectral Estimation*, S. HAYKIN, *Nonlinear Methods of Spectral Analysis*, p. 155-179, Springer-Verlag, 1983.
- [3] G. BIENVENU et L. KOPP, Optimality of high-resolution array processing using the eigenstructure approach, *IEEE Trans. on ASSP*, 31, n° 5, 1983, p. 1235-1245.
- [4] V. F. PISARENKO, The retrieval of harmonics from a covariance function, *Geophys. J.R. astr. soc.*, 33, 1973, p. 347-366.
- [5] G. BIENVENU et L. KOPP, Principe de la goniométrie passive adaptative, *VII^e Colloque GRETSI*, Nice, 1979, p. 106-110.
- [6] R. O. SCHMIDT, Multiple Emitter Location and Signal Parameter Estimation, *Proc. RADC Spectrum Estimation Workshop*, 1979, p. 243-258.
- [7] D. W. TUFTS et R. KUMARESAN, Estimating the angles of arrival of multiples plane waves, *IEEE Trans. on AES*, 19, n° 1, 1984, p. 134-139.
- [8] B. L. HO et R. E. KALMAN, Markov parameters, the moment problem, and minimal realizations, *J. Control (SIAM)*, 4, 1965.
- [9] P. FAURRE, *Stochastic realization algorithms*, in *System Identification: advances and case studies*, MEHRA and LAINOITIS éd., Academic Press, 1976, chap. 1.
- [10] S. Y. KUNG, K. S. ARUN et D. V. BHASKAR RAO, State space and singular value decomposition based approximation methods for the harmonic retrieval problem, *Journal of Optical Society of America*, 73, n° 12, Dec. 1983, p. 1789-1811.
- [11] K. S. ARUN, D. V. BHASKAR RAO et S. Y. KUNG, A new predictive efficiency criterion for approximate stochastic realization, *Proc. 22nd IEEE Conf. Decision Contr.*, San Antonio, TX, 1983, p. 1353-1355.
- [12] K. S. ARUN, A principal components approach to approximate modeling and ARMA spectral estimation, *Ph. D. dissertation*, University of Southern California, April 1984.
- [13] U. B. DESAI et D. PAL, A realization approach to stochastic model reduction and balanced stochastic realization, *Proc. 16th annual Conf. on Inf. Sciences and Systems*, Princeton University, NJ, March 1982, p. 613-620.

avec

$$\mathbf{w}^T = [w_1, w_2, \dots, w_m],$$

$$\mathbf{0}^T = [0, 0, \dots, 0] \text{ et } \mathbf{V} = (\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_M).$$

On considère la décomposition d'un vecteur \mathbf{Z} quelconque sur la base canonique $(\mathbf{E}_i)_{i=1, \dots, M}$:

$$\mathbf{Z} = \sum_{i=1}^M \beta_i \mathbf{E}_i.$$

Calculons: $\mathbf{Z}^* \cdot (\mathbf{U}_1^* \cdot \mathbf{U}_1 - (\mathbf{U}_1^{up})^* \cdot \mathbf{U}_1^{up}) \cdot \mathbf{Z}$:

Pour cela on considère tout d'abord deux éléments quelconques de la base canonique: $\mathbf{E}_i, \mathbf{E}_j$.

$$\begin{aligned} & \mathbf{E}_i^* \cdot [\mathbf{U}_1^* \cdot \mathbf{U}_1 - (\mathbf{U}_1^{up})^* \cdot \mathbf{U}_1^{up}] \cdot \mathbf{E}_j \\ &= \mathbf{E}_i^* \cdot \mathbf{U}_1^* \cdot \mathbf{U}_1 \cdot \mathbf{E}_j - \mathbf{E}_i^* \cdot (\mathbf{U}_1^{up})^* \cdot \mathbf{U}_1^{up} \cdot \mathbf{E}_j \\ &= (\bar{w}_j | \mathbf{V}_i^*) \begin{pmatrix} w_i \\ \mathbf{V}_j^* \end{pmatrix} - (\mathbf{V}_i^* | 0) \begin{pmatrix} \mathbf{V}_j \\ 0 \end{pmatrix} \\ &= \bar{w}_i \cdot w_j. \end{aligned}$$

On aura donc pour \mathbf{Z} :

$$\begin{aligned} \mathbf{Z}^* \cdot [\mathbf{U}_1^* \cdot \mathbf{U}_1 - (\mathbf{U}_1^{up})^* \cdot \mathbf{U}_1^{up}] \mathbf{Z} &= \sum_{i=1}^M \sum_{j=1}^M \beta_i \bar{w}_i \cdot \beta_j w_j \\ &= \left| \sum_{k=1}^M \beta_k w_k \right|^2 \geq 0. \end{aligned}$$

Par conséquent: $\mathbf{U}_1^* \cdot \mathbf{U}_1 - (\mathbf{U}_1^{up})^* \cdot \mathbf{U}_1^{up}$ est une matrice positive.

Or $\mathbf{U}_1^* \cdot \mathbf{U}_1 = \mathbf{I}$, donc $\mathbf{I} - (\mathbf{U}_1^{up})^* \cdot \mathbf{U}_1^{up}$ est une matrice, ce qui signifie que les valeurs propres de $(\mathbf{U}_1^{up})^* \cdot \mathbf{U}_1^{up}$ sont toutes de module inférieur ou égal à 1.

Donc F a toutes ses valeurs propres à l'intérieur du cercle unité, ce qui assure la stabilité de la solution obtenue.

A. 2. SUR LES VALEURS PROPRES DE F OBTENUE PAR PROJECTION [équation (63)]

Lorsqu'on introduit une colonne de zéros pour le calcul de (Γ^-) , cette matrice peut s'écrire:

$$(\Gamma^-) = \Gamma^- \cdot \mathbf{Z}$$

où

$$\mathbf{Z} = \begin{pmatrix} 0 & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & 0 \end{pmatrix}.$$

En utilisant la décomposition en éléments propres de $(\mathbf{H} \cdot (\Gamma^-)^{-1} \cdot \mathbf{H}^*)$, on voit que l'on peut écrire:

$$\mathbf{F} = \Sigma_1^{-1} \cdot \mathbf{U}_1^* \cdot (\mathbf{H} \cdot \mathbf{Z} \cdot \mathbf{H}^{-1}) \cdot \mathbf{U}_1 \cdot \Sigma_1,$$

d'où:

$$\text{Spectre}(F) = \text{Spectre}(\mathbf{U}_1^* \cdot (\mathbf{H} \cdot \mathbf{Z} \cdot \mathbf{H}^{-1}) \cdot \mathbf{U}_1).$$

De plus, par des calculs analogues à ceux donnés en annexe A 1, on peut montrer que:

$$\begin{aligned} & |v. p. m. (\mathbf{U}_1^* \cdot (\mathbf{H} \cdot \mathbf{Z} \cdot \mathbf{H}^{-1}) \cdot \mathbf{U}_1)|^2 \\ & \leq v. p. m. (\mathbf{U}_1^* \cdot \mathbf{H} \cdot \mathbf{Z} \cdot \mathbf{H}^{-1} \cdot \mathbf{U}_1 \mathbf{U}_1^* \cdot \mathbf{H}^{-*} \cdot \mathbf{Z}^* \cdot \mathbf{H}^* \cdot \mathbf{U}_1) \\ & \leq v. p. m. (\mathbf{H} \cdot \mathbf{Z} \cdot \mathbf{H}^{-1} \cdot \mathbf{U}_1 \cdot \mathbf{U}_1^* \cdot \mathbf{H}^{-*} \cdot \mathbf{Z}^* \cdot \mathbf{H}^*) \\ & \leq v. p. m. (\mathbf{H} \cdot \mathbf{Z} \cdot \mathbf{H}^{-1} \cdot \mathbf{H}^{-*} \cdot \mathbf{Z}^* \cdot \mathbf{H}^*) \end{aligned}$$

[où v. p. m. (M) signifie valeur propre de plus grand module de la matrice M].

Par ailleurs, pour un vecteur \mathbf{V} quelconque on a:

$$\begin{aligned} & \mathbf{V}^* \cdot (\mathbf{I} - \mathbf{H} \cdot \mathbf{Z} \cdot \mathbf{H}^{-1} \cdot \mathbf{H}^{-*} \cdot \mathbf{Z}^* \cdot \mathbf{H}^*) \mathbf{V} \\ &= \mathbf{W}^* \cdot (\mathbf{H}^{-1} \cdot \mathbf{H}^{-*} - (\mathbf{H}^{-1})^{\text{down}} \cdot [(\mathbf{H}^{-1})^{\text{down}}]^*) \cdot \mathbf{W} \end{aligned}$$

où: $\mathbf{W} = \mathbf{H}^* \cdot \mathbf{V}$ et $(\mathbf{H}^{-1})^{\text{down}} = \mathbf{Z} \cdot \mathbf{H}^{-1}$ (matrice obtenue en décalant les lignes de \mathbf{H}^{-1} d'une ligne vers le bas et en introduisant une première ligne de zéros).

De même qu'en annexe A. 1, on en déduit que la matrice $\{ \mathbf{H}^{-1} \cdot \mathbf{H}^{-*} - (\mathbf{H}^{-1})^{\text{down}} \cdot [(\mathbf{H}^{-1})^{\text{down}}]^* \}$ est une matrice positive, ce qui implique que $\{ \mathbf{I} - \mathbf{H} \cdot \mathbf{Z} \cdot \mathbf{H}^{-1} \cdot \mathbf{H}^{-*} \cdot \mathbf{Z}^* \cdot \mathbf{H}^* \}$ est positive.

On en déduit alors que les valeurs propres de $(\mathbf{H} \cdot \mathbf{Z} \cdot \mathbf{H}^{-1} \cdot \mathbf{H}^{-*} \cdot \mathbf{Z}^* \cdot \mathbf{H}^*)$ sont toutes de module inférieur ou égal à 1, et, par les inégalités successives précédentes, que les valeurs propres de F sont bien à l'intérieur du cercle unité.

Annexe B

SUR LES VALEURS PROPRES DE $[(\mathbf{O}^\dagger)^* \cdot \mathbf{O}^\dagger]$

Soit la matrice de transition du système définie par (\mathbf{O} étant définie comme précédemment)

$$\mathbf{F} = (\mathbf{O}^\dagger)^* \cdot \mathbf{O}^\dagger \Rightarrow \mathbf{F} = \mathbf{T}^{-1} \cdot (\mathbf{U}^\dagger)^* \cdot \mathbf{U}^\dagger \cdot \mathbf{T}$$

\mathbf{F} et $[(\mathbf{U}^\dagger)^* \cdot \mathbf{U}^\dagger]$ ayant le même spectre, on examine les valeurs propres de $\tilde{\mathbf{U}}_1 = (\mathbf{U}^\dagger)^* \cdot \mathbf{U}^\dagger$.

Plaçons-nous tout d'abord dans le cas $M=1$.

La matrice de transition est scalaire, $F=\lambda$, et la matrice des vecteurs propres principaux \mathbf{U}_1 , se réduit à un vecteur \mathbf{u} de norme égale à 1.

On a alors:

$$\begin{aligned} \lambda &= t^{-1} \cdot (\mathbf{u}^\dagger)^* \cdot \mathbf{u}^\dagger \cdot t \quad (t \text{ scalaire}) \\ &= (\mathbf{u}^\dagger)^* \cdot \mathbf{u}^\dagger. \end{aligned}$$

On note

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{N-1} \\ u_N \end{bmatrix}, \quad \|\mathbf{u}\| = 1$$

Donc

$$\mathbf{u}^\uparrow = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{N-1} \end{bmatrix} \quad \text{et} \quad \mathbf{u}^\downarrow = \begin{bmatrix} u_2 \\ \vdots \\ u_{N-1} \\ u_N \end{bmatrix}$$

$$(\mathbf{u}^\uparrow)^* = \frac{(\mathbf{u}^\uparrow)^*}{\|\mathbf{u}^\uparrow\|^2}$$

d'où :

$$|\lambda| = \frac{|(\mathbf{u}^\uparrow)^* \cdot \mathbf{u}^\downarrow|}{\|\mathbf{u}^\uparrow\|^2} \leq \frac{\|\mathbf{u}^\downarrow\|}{\|\mathbf{u}^\uparrow\|} = \frac{\sqrt{1-|u_1|^2}}{\sqrt{1-|u_N|^2}}$$

En général, on ne peut pas conclure quant à la majoration de $|\lambda|$ par 1. La stabilité de la solution n'est donc pas garantie. Toutefois, dans le cas asymptotique, ou, lorsque la MIO est sous forme Toeplitz, \mathbf{u} vérifie :

$$|u_1|^2 = |u_N|^2.$$

En effet :

– asymptotiquement : $\mathbf{u} = \alpha \mathbf{D}_\theta$, ce qui implique bien évidemment :

$$|u_1|^2 = |u_N|^2,$$

– cas Toeplitz : on considère la « décomplexification » de la MIO : $\Gamma = \mathbf{P} + j\mathbf{Q}$ ($j^2 = -1$, \mathbf{P} et \mathbf{Q} matrices réelles). \mathbf{P} est centrosymétrique, \mathbf{Q} est antisymétrique par rapport à sa diagonale principale et symétrique par rapport à sa diagonale secondaire.

$\{\mathbf{s}_i\}_{i=1, \dots, N}$ étant les vecteurs propres de la matrice symétrique $[\mathbf{P} - \mathbf{J} \cdot \mathbf{Q}]$, on montre que les vecteurs propres de Γ sont définis par $\mathbf{u}_i = \mathbf{s}_i + j \cdot \mathbf{J} \cdot \mathbf{s}_i$ pour $i = 1, \dots, N$. Ce qui entraîne encore : $|u_1|^2 = |u_N|^2$.

Plus généralement, une matrice de Toeplitz peut s'écrire sous la forme :

$$\Gamma = \sum_{i=1}^P \mu_i \mathbf{D}_{\theta_i} \cdot \mathbf{D}_{\theta_i}^* + b^2 \cdot \mathbf{I}$$

où \mathbf{D}_{θ_i} est un vecteur de type « vecteur source » de gisement θ_i .

On voit donc que l'étude de la stabilité lorsque Γ est Toeplitz est équivalente à l'étude de la stabilité du cas asymptotique puisque dans les deux cas les vecteurs propres de Γ sont combinaisons linéaires de vecteurs de la forme \mathbf{D}_θ . On peut alors vérifier que la stabilité est assurée.

Annexe C

CONTRAINTES SUR L'ESTIMATION DE LA MATRICE DE TRANSITION

La résolution du problème $\mathbf{O}^\downarrow = \mathbf{O}^\uparrow \cdot \mathbf{F}$ peut se poser sous forme d'une minimisation sous contraintes.

Nous proposons ici de minimiser le critère $\|\mathbf{O}^\downarrow - \mathbf{O}^\uparrow \mathbf{F}\|$ sous la contrainte \mathbf{F} unitaire.

Deux raisons sont à l'origine de ce choix : d'une part on sait que les modules des valeurs propres de \mathbf{F} sont asymptotiquement égaux à 1, et d'autre part, le fait de rendre ceux-ci égaux à 1 nous assure la stabilité marginale de la solution.

Le problème se pose de la façon suivante :

$$\begin{cases} \text{Min} \|\mathbf{O}^\downarrow - \mathbf{O}^\uparrow \mathbf{F}\| \\ \mathbf{F} \text{ unitaire, de dimension } (M, M). \end{cases}$$

Or,

$$\|\mathbf{O}^\downarrow - \mathbf{O}^\uparrow \mathbf{F}\|^2 = \text{tr}(\mathbf{O}^{\downarrow*} \cdot \mathbf{O}^\downarrow) + \text{tr}(\mathbf{O}^{\uparrow*} \cdot \mathbf{O}^\uparrow) - \text{tr}(\mathbf{O}^{\downarrow*} \cdot \mathbf{O}^\uparrow \cdot \mathbf{F}) - \text{tr}(\mathbf{O}^{\uparrow*} \cdot \mathbf{O}^\downarrow \cdot \mathbf{F}^*)$$

Il s'agit donc de déterminer \mathbf{F} maximisant :

$$\text{tr}(\mathbf{O}^{\downarrow*} \cdot \mathbf{O}^\uparrow \cdot \mathbf{F}) + \text{tr}(\mathbf{O}^{\uparrow*} \cdot \mathbf{O}^\downarrow \cdot \mathbf{F}^*).$$

Pour cela on utilise la SVD de $(\mathbf{O}^{\uparrow*} \cdot \mathbf{O}^\downarrow)$:

$$\mathbf{O}^{\uparrow*} \cdot \mathbf{O}^\downarrow = \mathbf{V} \cdot \Delta \cdot \mathbf{W}^*$$

où \mathbf{V} et \mathbf{W} sont deux matrices (M, M) unitaires et $\Delta = \text{diag}(\delta_i)$ avec $\delta_i \geq 0$ pour $i = 1, \dots, M$.

Soit :

$$\Delta = \mathbf{V}^* \cdot (\mathbf{O}^{\uparrow*} \cdot \mathbf{O}^\downarrow) \cdot \mathbf{W}.$$

Alors :

$$\begin{aligned} \text{tr}(\mathbf{O}^{\uparrow*} \cdot \mathbf{O}^\downarrow \cdot \mathbf{F}^*) &= \text{tr}(\mathbf{V} \cdot \Delta \cdot \mathbf{W}^* \cdot \mathbf{F}^*) \\ &= \text{tr}(\Delta \cdot [\mathbf{W}^* \cdot \mathbf{F}^* \cdot \mathbf{V}]). \end{aligned}$$

Posons :

$$\mathbf{Z} = \mathbf{W}^* \cdot \mathbf{F}^* \cdot \mathbf{V},$$

\mathbf{Z} est une matrice (M, M) unitaire car produit de trois matrices unitaires.

On a donc :

$$\text{tr}(\mathbf{O}^{\uparrow*} \cdot \mathbf{O}^\downarrow \cdot \mathbf{F}^*) = \text{tr}(\Delta \cdot \mathbf{Z})$$

De même :

$$\text{tr}(\mathbf{F} \cdot \mathbf{O}^{\downarrow*} \cdot \mathbf{O}^\uparrow) = \text{tr}(\mathbf{Z}^* \cdot \Delta^*).$$

Et par conséquent :

$$\begin{aligned} \text{tr}(\mathbf{O}^{\downarrow*} \cdot \mathbf{O}^\uparrow \cdot \mathbf{F}) + \text{tr}(\mathbf{O}^{\uparrow*} \cdot \mathbf{O}^\downarrow \cdot \mathbf{F}^*) &= \text{tr}(\mathbf{Z}^* \cdot \Delta^*) + \text{tr}(\Delta \cdot \mathbf{Z}) \\ &= 2 \cdot \sum_{i=1}^M \text{Re}(Z_{ii}) \cdot \delta_i \leq 2 \cdot \sum_{i=1}^M \delta_i \end{aligned}$$

où $\text{Re}(Z_{ii})$ signifie partie réelle d'un élément de la diagonale principale de la matrice \mathbf{Z} .

Cette dernière inégalité est vérifiée car tous les δ_i sont positifs et que $\text{Re}(Z_{ii}) \leq 1$, pour $i = 1, \dots, M$, puisque \mathbf{Z} est unitaire.

$[\text{tr}(\mathbf{O}^{\downarrow*} \cdot \mathbf{O}^\uparrow \cdot \mathbf{F}) + \text{tr}(\mathbf{O}^{\uparrow*} \cdot \mathbf{O}^\downarrow \cdot \mathbf{F}^*)]$ est donc maximal pour $\mathbf{Z} = \mathbf{I}$, soit : $\mathbf{F} = \mathbf{V} \cdot \mathbf{W}^*$.

Annexe D

MINIMISATION DU CRITÈRE (47)

On veut déterminer la matrice A de dimension (M, L) qui minimise le critère :

$$\text{tr}[\Gamma^+ - H \cdot A^* \cdot (A \cdot \Gamma^- \cdot A^*)^{-1} \cdot A \cdot H^*].$$

Soit B une matrice (M, M) quelconque de rang M. On vérifie aisément l'égalité suivante :

$$\begin{aligned} \text{tr}[\Gamma^+ - H \cdot (B \cdot A)^* \cdot ((BA) \cdot \Gamma^- \cdot (BA)^*)^{-1} \cdot (BA) \cdot H^*] \\ = \text{tr}[\Gamma^+ - H \cdot A^* \cdot (A \cdot \Gamma^- \cdot A^*)^{-1} \cdot A \cdot H^*]. \end{aligned}$$

Le fait de remplacer A par (B · A), avec B inversible, ne change donc pas la valeur du critère.

Par ailleurs, la matrice (A · Γ⁻ · A*) étant hermitienne il existe une matrice B unitaire telle que (B · A · Γ⁻ · A* · B*) soit une matrice diagonale.

On peut alors se restreindre à chercher la solution du problème de maximisation sous contrainte suivant :

$$(P) \left\{ \begin{array}{l} \text{Max}_A \{ \text{tr}[H \cdot A^* \cdot A \cdot H^*] \} \\ \mathcal{C} : A \cdot \Gamma^- \cdot A^* = I \\ \mathcal{H} : A \cdot \Gamma^- \cdot A^* \text{ est une matrice diagonale.} \end{array} \right.$$

Sous l'hypothèse \mathcal{H} , déduite de ce que nous avons trouvé précédemment, la contrainte \mathcal{C} s'écrit : $A_i^* \cdot \Gamma^- \cdot A_i = 1$, pour $i = 1, \dots, M$, A_i^* étant le i -ième vecteur ligne de A.

En remarquant que

$$\text{tr}[(H \cdot A^*) \cdot (A \cdot H^*)] = \text{tr}[(A \cdot H^*) \cdot (H \cdot A^*)],$$

il est équivalent d'écrire :

$$(P) \left\{ \begin{array}{l} \text{Max}_A \{ \text{tr}[A \cdot H^* \cdot H \cdot A^*] \} \\ \mathcal{C} : A_i^* \cdot \Gamma^- \cdot A_i = 1, \quad \text{pour } i = 1, \dots, M. \end{array} \right.$$

On note E_i le vecteur composé d'un 1 sur la i -ième ligne et de 0 partout ailleurs, et $\Delta_i = E_i \cdot E_i^*$, matrice (M, M).

Il vient :

$$\begin{aligned} A_i^* \cdot \Gamma^- \cdot A_i &= \text{tr}[A_i^* \cdot \Gamma^- \cdot A_i] \\ &= \text{tr}[\Gamma^- \cdot A_i \cdot A_i^*] \\ &= \text{tr}[\Gamma^- \cdot (A^* \cdot E_i) \cdot (E_i^* \cdot A)] \\ &= \text{tr}[\Gamma^- \cdot A^* \cdot \Delta_i \cdot A] \end{aligned}$$

(P) est alors équivalent à :

$$(P) \left\{ \begin{array}{l} \text{Max}_A \{ \text{tr}[A \cdot H^* \cdot H \cdot A^*] \} \\ \mathcal{C} : \text{tr}[\Gamma^- \cdot A^* \cdot \Delta_i \cdot A] = 1. \end{array} \right.$$

Une condition nécessaire pour que (P) ait une solution est qu'il existe des multiplicateurs de Lagrange λ_i , $i = 1, \dots, M$, tels que :

$$\begin{aligned} D_A \{ \text{tr}[A \cdot H^* \cdot H \cdot A^*] \} (K) \\ - \sum_{i=1}^M \lambda_i \cdot D_A \{ \text{tr}[\Delta_i \cdot A \cdot \Gamma^- \cdot A^*] \} (K) = 0 \end{aligned}$$

où $D_A \{ \text{tr}[\dots] \} (K)$ signifie différentielle en A de l'application $\text{tr}[\dots]$, appliquée au point K.

Soit :

$$\text{tr}[A \cdot H^* \cdot H \cdot K^*] - \sum_{i=1}^M \lambda_i \cdot \text{tr}[\Delta_i \cdot A \cdot \Gamma^- \cdot K^*] = 0,$$

K quelconque

$$\Leftrightarrow \text{tr} \left[\left(A \cdot H^* \cdot H - \sum_{i=1}^M \lambda_i \cdot \Delta_i \cdot A \cdot \Gamma^- \right) K^* \right] = 0,$$

K quelconque.

L'égalité précédente étant vérifiée pour tout K et étant donné que l'application :

$$K \rightarrow \text{tr} \left[\left(A \cdot H^* \cdot H - \sum_{i=1}^M \lambda_i \cdot \Delta_i \cdot A \cdot \Gamma^- \right) K^* \right]$$

est une forme linéaire de $\mathcal{M}_C(M, L)$, on a en particulier pour

$$K = \left(A \cdot H^* \cdot H - \sum_{i=1}^M \lambda_i \cdot \Delta_i \cdot A \cdot \Gamma^- \right) : \text{tr}[K \cdot K^*] = 0.$$

Or, l'application $K \rightarrow \text{tr}[K \cdot K^*]$ est une forme hermitienne définie positive sur $\mathcal{M}_C(M, L)$. Ce qui implique :

$$A \cdot H^* \cdot H - \sum_{i=1}^M \lambda_i \cdot \Delta_i \cdot A \cdot \Gamma^- = 0$$

$$\Leftrightarrow A \cdot H^* \cdot H = \Lambda \cdot A \cdot \Gamma^-$$

où Λ est la matrice (M, M) diagonale des multiplicateurs de Lagrange.

Posons $A \cdot \Gamma^- \cdot H^{-1} = \Phi$, l'égalité précédente devient :

$$\Phi \cdot [H \cdot (\Gamma^-)^{-1} \cdot H^*] = \Lambda \cdot \Phi$$

Φ étant de dimension (M, L), $M < L$, la meilleure approximation de rang M consiste à ne retenir que les M composantes principales de la matrice $(H \cdot (\Gamma^-)^{-1} \cdot H^*)$. Soit :

$$H \cdot (\Gamma^-)^{-1} \cdot H^* = [U_1 | U_2] \begin{bmatrix} \Sigma_1^2 & O \\ O & \Sigma_2^2 \end{bmatrix} \begin{bmatrix} U_1^* \\ U_2^* \end{bmatrix}$$

On obtient

$$\begin{cases} \Phi = \Sigma_1^{-1} \cdot U_1^* \\ \Lambda = \Sigma_1^2 \end{cases}$$

On en déduit alors la matrice A qui minimise le critère (47) :

$$A = \Sigma_1^{-1} \cdot U_1^* \cdot H \cdot (\Gamma^-)^{-1}$$

Remarque : Plus généralement, on a vu que si A est solution du problème de minimisation, (B, A), avec B matrice (M, M) de rang M, l'est aussi. On en déduit une solution plus générale du problème :

$$(50) \quad A = B \cdot \Sigma_1^{-1} \cdot U_1^* \cdot H \cdot (\Gamma^-)^{-1}.$$

Annexe E

MINIMISATION DU CRITÈRE (69)

On veut déterminer la matrice A de dimension (M, L) qui minimise le critère :

$$\det[\Gamma^+ - HA^*(A\Gamma^-A^*)^{-1}AH^*].$$

Ce problème est équivalent au problème de minimisation sous contrainte suivant :

$$(E.1) \quad \begin{cases} \text{Min}_{A(M,L)} \{ \det(\Gamma^+ - H \cdot A^* \cdot A \cdot H^*) \\ \text{Contrainte : } A \cdot \Gamma^- \cdot A^* = I_M. \end{cases}$$

Pour cela, on considère la SVD de A :

$$(E.2) \quad A = P \cdot \Delta \cdot Q$$

P, Q matrices (M, M) et (L, L) unitaires, Δ matrice (M, L) des valeurs singulières.

La contrainte s'écrit :

$$\begin{aligned} P \cdot \Delta \cdot Q \cdot \Gamma^- \cdot Q^* \cdot \Delta^* \cdot P^* &= I_M \\ \Leftrightarrow [\Delta \cdot Q \cdot (\Gamma^-)^{1/2}] [\Delta \cdot Q \cdot (\Gamma^-)^{1/2}]^* &= I_M. \end{aligned}$$

Donc, il existe $V_1(L, M)$ telle que :

$$(E.3) \quad \begin{cases} V_1^* = \Delta \cdot Q \cdot (\Gamma^-)^{1/2} \\ V_1^* \cdot V_1 = I_M. \end{cases}$$

Par ailleurs :

$$\begin{aligned} \det(\Gamma^+ - H \cdot A^* \cdot A \cdot H^*) &= \det(H \cdot H^*) \cdot \det(H^{-1} \cdot \Gamma^+ \cdot H^{-*} - A^* \cdot A) \\ &= \det(H \cdot H^*) \cdot \det(H^{-1} \cdot \Gamma^+ \cdot H^{-*} \\ &\quad - [(\Gamma^-)^{-1/2}]^* \cdot V_1 \cdot V_1^* \cdot (\Gamma^-)^{-1/2}) \\ &= \det(H \cdot H^*) \cdot \det((\Gamma^-)^{-1}) \\ &\quad \times \det[\Omega^{-1} \cdot (\Omega^{-1})^* - V_1 \cdot V_1^*] \end{aligned}$$

avec

$$\Omega = (\Gamma^+)^{-1/2} \cdot H \cdot [(\Gamma^-)^{-1/2}]^*.$$

Le problème revient donc à déterminer la matrice de rang M solution de :

$$\begin{aligned} \text{Min}_{V_1(L,M)} \{ \det[\Omega^{-1} \cdot (\Omega^{-1})^* - V_1 \cdot V_1^*] \\ \text{Contrainte : } V_1^* \cdot V_1 = I_M. \end{aligned}$$

Avec pour hypothèse $[\Omega^{-1} \cdot \Omega^{-*} - V_1 \cdot V_1^*]$ matrice positive (car issue d'une matrice de covariance).

Posons :

$$V_{(m)} = (v_1, v_2, \dots, v_m), \quad V_{(M)} = V_1.$$

La solution est obtenue par récurrence sur m ([15], chap II, § 4.1).

En effet : pour m = 1, on doit minimiser :

$$(E.4) \quad \begin{aligned} \det(\Omega^{-1} \cdot \Omega^{-*} - v_1 \cdot v_1^*) \\ = \det(\Omega^{-1} \cdot \Omega^{-*}) \cdot (1 - v_1^* \cdot (\Omega^* \cdot \Omega) \cdot v_1) \end{aligned}$$

La matrice $[\Omega^{-1} \cdot \Omega^{-*}]$ étant positive, ce minimum est atteint lorsque v_1 est vecteur propre de $(\Omega^* \cdot \Omega)$ associé à la plus grande valeur propre.

De plus, la contrainte est vérifiée puisqu'on peut choisir v_1 normé. Dans ces conditions, on a :

$$\det(\Omega^{-1} \cdot \Omega^{-*} - v_1 \cdot v_1^*) = \det(\Omega^{-1} \cdot \Omega^{-*}) \cdot (1 - \sigma_1^4)$$

où σ_1^2 est la plus grande valeur singulière de $\Omega = E[\varepsilon^+ \cdot \varepsilon^-]^*$.

Alors, $\Omega^{-1} \cdot \Omega^{-*} - v_1 \cdot v_1^*$ est positive si $\sigma_1^4 \leq 1$.

On montre aisément que ceci est vérifié en utilisant le lemme suivant :

Lemme 3 [18], p. 166 : Soient ξ et η des matrices de même dimension (pas nécessairement carrées) à éléments aléatoires. Si la matrice $E(\eta \cdot \eta^*)$ est inversible, alors :

$$E(\xi \cdot \xi^*) \geq E(\xi \cdot \eta^*) \cdot [E(\eta \cdot \eta^*)]^{-1} \cdot E(\eta \cdot \xi^*).$$

L'égalité n'est possible que si $\xi = \zeta \cdot \eta$,

$$\zeta = E(\xi \cdot \eta^*) \cdot [E(\eta \cdot \eta^*)]^{-1}.$$

En posant : $\xi = \varepsilon^+$ et $\eta = \varepsilon^-$, on obtient :

$$I_L \geq \Omega \cdot \Omega^*.$$

Ce qui entraîne que la plus grande valeur propre de $(\Omega \cdot \Omega^*)$ est inférieure à la plus grande valeur propre de I_L . Soit : $\sigma_1^4 \leq 1$.

— supposons la propriété vérifiée à l'ordre m : les $\{v_i\}$ sont engendrés par les m vecteurs propres de $(\Omega^* \cdot \Omega)$ associés aux m plus grandes valeurs propres.

La contrainte est vérifiée (on peut choisir un système orthonormé pour système propre d'une matrice hermitienne), et $\det(\Omega^{-1} \cdot \Omega^{-*} - V_{(m)} \cdot V_{(m)}^*)$ est positif.

— à l'ordre (m + 1) on a :

$$V_{(m+1)} \cdot V_{(m+1)}^* = V_{(m)} \cdot V_{(m)}^* + v_{m+1} \cdot v_{m+1}^*$$

et

$$(E.5) \quad \begin{aligned} \det(\Omega^{-1} \cdot \Omega^{-*} - V_{(m+1)} \cdot V_{(m+1)}^*) \\ = \det(\Omega^{-1} \cdot \Omega^{-*} - V_{(m)}^*) \\ \times (1 - v_{m+1}^* [\Omega^{-1} \cdot \Omega^{-*} - V_{(m)} \cdot V_{(m)}^*]^{-1} \cdot v_{m+1}). \end{aligned}$$

La matrice $(\Omega^{-1} \cdot \Omega^{-*} - V_{(m)} \cdot V_{(m)}^*)$ étant par construction positive, le minimum de (E.5) relativement à v_{m+1} est atteint lorsque v_{m+1} est vecteur propre de

$[\Omega^{-1} \cdot \Omega^{-*} - V_{(m)} \cdot V_{(m)}^*]^{-1}$ associé à la plus grande valeur propre. Pour que la contrainte soit vérifiée il faut et il suffit que la famille $\{v_1, \dots, v_{m+1}\}$ forme un système orthonormé.

On en déduit donc que v_{m+1} est un vecteur propre de $(\Omega^* \cdot \Omega)$ associé à la valeur propre de $(\Omega^* \cdot \Omega)$ suivant immédiatement les valeurs propres associées aux vecteurs v_1, \dots, v_m .

La matrice $(\Omega^{-1} \cdot \Omega^{-*} - V_{(m+1)} \cdot V_{(m+1)}^*)$ ainsi construite est bien évidemment positive.

On en conclut donc que la matrice V_1 solution du problème est constituée des M vecteurs propres associés au M plus grandes valeurs propres de $(\Omega^* \cdot \Omega)$.

Considérons la SVD de Ω [équation (64)] :

$$\Omega = U \cdot \Sigma^2 \cdot V^*$$

$$\Rightarrow \Omega^* \cdot \Omega = V \cdot \Sigma^4 \cdot V^* = [V_1 \quad V_2] \left[\begin{array}{c|c} \Sigma_1^4 & 0 \\ \hline 0 & \Sigma_2^4 \end{array} \right] \left[\begin{array}{c} V_1^* \\ V_2^* \end{array} \right]$$

On peut donc déduire aisément la matrice V_1 de la décomposition en valeurs singulières de $\Omega = (\Gamma^+)^{-1/2} \cdot H \cdot [(\Gamma^-)^{-1/2}]^*$:

$$(\Gamma^+)^{-1/2} \cdot H \cdot [(\Gamma^-)^{-1/2}]^* = [U_1 \quad U_2] \left[\begin{array}{c|c} \Sigma_1^2 & 0 \\ \hline 0 & \Sigma_2^2 \end{array} \right] \left[\begin{array}{c} V_1^* \\ V_2^* \end{array} \right]$$

A partir de (E.2), (E.3) on détermine la matrice $A(M, L)$ reliant l'état du système au passé des observations :

$$(E.6) \quad A = P \cdot V_1^* (\Gamma^-)^{-1/2}$$

avec P matrice (M, M) unitaire arbitraire.

Le critère étant invariant par changement de base, on en déduit une solution plus générale :

$$A = B \cdot V_1^* (\Gamma^-)^{-1/2}$$

avec B matrice (M, M) , de rang M , arbitraire.