

# Méthode d'estimation de la matrice de densité

## spectrale : optimisation du compromis biais-variance

### Spectral density matrix estimation

#### method : bias-variance trade-off optimization



#### B. LUMEAU

Laboratoire des Signaux et Systèmes (LSS) - CNRS-ESE-UPS - Ecole Supérieure d'Electricité, Plateau du Moulon, 91190 GIF SUR YVETTE et GRECO « Traitement du Signal et de l'Image ».

Bernard Lumeau est ingénieur ESE (1976), Docteur-Ingénieur (1980), Docteur ès Sciences (1987) et Maître de Conférences (1981) à l'Université Paris Val de Marne où il enseigne le Traitement du Signal. Depuis 1977, il effectue sa recherche au Laboratoire des Signaux et Systèmes (LSS) à Gif sur Yvette au sein de la « Division Signaux ». Ses domaines de recherche sont : le Traitement Spatial, l'Analyse Spectrale et Interspectrale avec des applications en Instrumentation Biomédicale.



#### H. CLERGEOT

Laboratoire d'Electricité, Signaux et Robotique (LESIR) - ENS - 61, avenue du Président Wilson, 94230 CACHAN.

Henri Clergeot est né à Lalpays, dans le centre de la France, en 1942. Après des études à l'ENS de Cachan, il a été reçu à l'agrégation de Physique, option physique appliquée, en 1968. Jusqu'en 1988, il a enseigné à l'ENS de Cachan, en électronique et traitement du signal. Depuis il est Professeur des Universités à l'IUT de génie électrique de Cachan, sur un enseignement de systèmes asservis.

Pour l'aspect recherche, il a travaillé en analyse spectrale au Laboratoire des Signaux et Systèmes (LSS) à Gif sur Yvette jusqu'en 1982, date de soutenance de sa Thèse d'Etat. Il travaille maintenant au Laboratoire d'Electricité, Signaux et Robotique (LESIR) à l'ENS de Cachan, en tant que responsable d'un groupe d'étude de capteurs pour la perception de l'environnement en Robotique.

### RÉSUMÉ

Nous présentons une méthode non paramétrique pour le calcul de la matrice de densité spectrale de signaux stationnaires permettant de minimiser l'erreur totale d'estimation, au moins dans le cas de signaux à spectres lentement variables. La méthode repose sur un algorithme récursif de blanchiment et de décorrélation de signaux stationnaires qui se ramène à chaque étape au calcul des autospectres et de l'interspectre de deux signaux par un lissage fréquentiel avec adaptation automatique de la fenêtre de pondération. Cette méthode tout en assurant le caractère défini non négatif de la matrice de densité spectrale permet d'optimiser le compromis biais-variance sur chacun de ses éléments.

**MOTS-CLÉS :** Autospectre, interspectre, matrice de densité spectrale, optimisation du compromis biais-variance.

### SUMMARY

*A non-parametric method is presented for the computation of the spectral density matrix of stationary signals, intended to minimize the total error, at least in the case of slowly varying spectra. The method is based on a recursive algorithm which achieves, by successive linear transforms, whitening and decorrelation of the signals. Each step involves the estimation of two spectral densities and one cross-spectral density, using frequency smoothing with an automatic choice of the smoothing window. This method preserves the definite non-negative property of the matrix and achieves an optimization of the trade-off between bias and variance.*

**KEY WORDS :** Spectral density, cross-spectral density, spectral density matrix, optimization of the trade-off between bias and variance.

**NOTATIONS :** Les grandeurs fréquentielles sont différenciées des grandeurs temporelles par un  $\circ$  sur le caractère associé. Pour rappeler qu'elles résultent de l'échantillonnage d'une fonction continue de la fréquence, cette dépendance est notée entre parenthèses et non en indice : quand une suite de valeurs fréquentielles est regroupée en vecteur, les composantes de vecteurs apparaissent sous la forme  $\hat{u}(i)$  au lieu de  $\hat{u}_i$ .

## 1. Introduction

L'une des méthodes les plus classiques en analyse spectrale non paramétrique, le lissage du périodogramme, impose quelques remarques quant au choix de la fenêtre de lissage et à sa largeur. Désignons par  $\{\hat{x}_k(m)\}$  les transformés de Fourier discrètes (T.F.D.) sur  $N$  points de  $K$  signaux étudiés, où  $m$  repère la fréquence discrète. Lorsque nous voulons estimer en discret une matrice de densité spectrale  $\hat{\Gamma}(m)$  à partir des T.F.D.  $\{\hat{x}_i(m)\}$ , nous formons la « matrice périodogramme »,  $\hat{\Pi}(m)$  dont l'espérance mathématique est voisine de la densité spectrale  $\hat{\Gamma}(m)$  [1] sous réserve que le temps de corrélation  $t_c$  soit petit devant le temps d'observation  $T$  [2] :

$$(1) \quad \begin{cases} \hat{\Pi}(m) = \hat{x}(m) \hat{x}^+(m) / N, \\ E[\hat{\Pi}(m)] = \hat{\Gamma}(m), \end{cases}$$

avec  $\begin{cases} \hat{x}(m) = [\hat{x}_1(m) \dots \hat{x}_K(m)]^T, \\ m \in \{1, N\} \text{ et } m = fT. \end{cases}$

Nous savons que les éléments de la « matrice périodogramme » ont une variance qui dans le cas gaussien est supérieure ou égale au carré de sa valeur moyenne [1]. Pour réduire cette variance une méthode classique consiste à calculer une estimée de  $\hat{\Gamma}(m)$  par la relation de lissage fréquentiel [2] :

$$(2) \quad \hat{\Gamma}(m) = \sum_{i=-L/2}^{L/2} \hat{w}(i) \hat{\Pi}(m-i) \quad \text{avec}$$

$$\text{avec} \quad \sum_{i=-L/2}^{L/2} \hat{w}(i) = 1.$$

Les matrices  $\hat{\Pi}(m)$  étant définies non négatives, la même propriété est garantie pour l'interspectre estimé si les  $\hat{w}(i)$  sont positifs. Ceci étant, plus la largeur de la fenêtre de pondération est grande plus l'effet de « lissage » est important : la variance est réduite mais les variations rapides du spectre sont perdues. Le choix de la largeur résulte donc forcément d'un compromis entre variance et biais d'estimation qui est, le plus souvent, obtenu de manière empirique.

Les matrices de densité spectrale sont d'une utilisation courante en traitement spatial. Nous savons que les méthodes développées dans ce domaine, basées sur les propriétés des éléments propres de la matrice de densité spectrale, imposent théoriquement aux valeurs propres de cette matrice d'être positives ou nulles : ce qui justifie l'importance du caractère défini non négatif de cette matrice. De plus, lors de la recherche de la dimension du sous-espace relatif

aux sources (liée au rang de la matrice) nous savons [3] que le biais doit être minimisé (sans pour autant accroître la variance d'estimation) dans le but de réduire l'effet des valeurs propres parasites qui sont à l'origine d'un accroissement injustifié du rang. A ces fins, dans le deuxième paragraphe nous développons une méthode originale qui, tout en assurant le caractère défini non négatif, optimise le compromis biais-variance sur chacun des éléments de la matrice de densité spectrale.

Dans le cas des autospectres plusieurs approches permettant d'optimiser le compromis biais-variance ont déjà été proposées dans le cadre des méthodes d'estimation non paramétriques [4-5]. Le critère MAICE d'Akaike (appliqué ou non à l'analyse spectrale) [6] répond à la même préoccupation dans le cas d'une estimation par un développement paramétrique tronqué.

Dans la première partie du deuxième paragraphe, en vue d'applications ultérieures à l'analyse spectrale, nous introduisons pour un modèle à bruit additif une généralisation de l'approche d'Akaike dans laquelle la troncature d'un développement paramétrique est remplacée par une pondération des coefficients par une fenêtre de largeur variable. Dans la deuxième partie, nous établissons une transformation sur les autospectres permettant de nous ramener au modèle à bruit additif précédent, avec une puissance de bruit connue. Nous étendons, dans la troisième partie, ces résultats au cas de l'interspectre de deux signaux en utilisant la notion de cohérence. Nous proposons, dans la quatrième partie, un algorithme récursif de blanchiment et de décorrélation spatiale permettant de généraliser les résultats précédents pour l'estimation de la matrice de densité spectrale, sans risque de perdre le caractère défini non négatif d'une telle matrice.

## 2. Estimation optimale de la matrice de densité spectrale par blanchiment et décorrélation spatiale

### 2.1. MÉTHODE DE LISSAGE OPTIMAL POUR LE MODÈLE DE BASE

La méthode de lissage optimal est introduite pour un modèle statistique d'une observation  $\hat{u}(m)$  consistant en un signal  $\hat{s}(m)$  à estimer, considéré comme déterministe mais inconnu, perturbé par un bruit blanc complexe additif  $\hat{b}(m)$  de variance connue  $\sigma^2$  et décorrélé avec le signal. Nous montrerons dans les paragraphes suivants comment on se ramène à un tel modèle à partir du périodogramme d'un signal scalaire stationnaire ou de « l'interpériodogramme » de deux signaux ; l'indice  $m$  jouera le rôle de la variable fréquence.

La partie signal est estimée de manière paramétrique : nous représentons  $\hat{s}(m)$  par un développement tronqué sur une base orthonormée, dont l'ordre est estimé à l'aide d'un critère d'erreur quadratique minimale. L'expression de cette erreur totale dans la base fait apparaître deux termes : un terme de

variance qui croît en fonction de l'ordre et un terme de biais qui décroît en fonction de l'ordre. La minimisation de cette erreur convenablement estimée nous donne la valeur optimale du rang et par suite la largeur optimale de la fenêtre de lissage. Ainsi, nous définissons un test automatique assurant le meilleur compromis biais-variance.

Si nous considérons  $M$  observations, le signal observé s'écrit sous forme vectorielle dans un espace de dimension  $M$  :

$$(3) \quad \hat{\mathbf{u}} = \hat{\mathbf{s}} + \hat{\mathbf{b}} \quad \text{avec} \quad \hat{\mathbf{u}} = [\hat{u}(1), \dots, \hat{u}(M)]^T.$$

Décomposons chacun des signaux formant (3) dans une base donnée de  $M$  vecteurs orthonormés  $[\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_{M-1}]$  :

$$(4) \quad \begin{cases} \hat{\mathbf{u}} = \sum_{i=0}^{M-1} \alpha_i \mathbf{v}_i, \\ \hat{\mathbf{s}} = \sum_{i=0}^{M-1} a_i \mathbf{v}_i, \\ \hat{\mathbf{b}} = \sum_{i=0}^{M-1} \beta_i \mathbf{v}_i. \end{cases}$$

Le choix de cette base sera évoquée au sous-paragraphe suivant.

Nous en déduisons entre les projections la relation :

$$(5) \quad \alpha_i = a_i + \beta_i.$$

Les relations (4) s'interprètent comme un changement de base orthonormée dans l'espace vectoriel signal. Il en résulte entre les composantes complexes du vecteur bruit dans l'ancienne et la nouvelle base les relations suivantes :

$$(6) \quad \begin{cases} E\{|\beta_i|^2\} = E\{|\hat{b}(m)|^2\} = \sigma^2, \\ E[\beta_i \beta_j^*] = 0 \quad \text{pour} \quad i \neq j, \\ E[\beta_i \beta_j] = E[b(m) b(n)] = 0. \end{cases}$$

### 2.1.1. Développement paramétrique tronqué à l'ordre $p < M$

La transformation (4) est intéressante si, pour la famille de signaux à estimer, on connaît a priori une base telle que toute la puissance soit concentrée dans les premières composantes : on pourra alors estimer le signal par un développement tronqué à un petit nombre de termes. Ce choix de la base pourrait être fait sur un critère statistique à partir d'un échantillon représentatif de signaux, pour une application donnée.

N'oublions pas cependant que nous nous sommes d'emblée placés dans l'hypothèse de spectres à « variation lente », condition nécessaire pour la validité de (1), et qui sera utilisée ultérieurement pour introduire les propriétés statistiques du périodogramme.

Dans ces conditions un choix logique consiste à utiliser comme fonctions de base des exponentielles complexes de fréquence croissante : on retrouve la Transformée de Fourier Discrète (T.F.D.). D'autres

bases ont été essayées, notamment des polynômes orthogonaux de degré croissant, mais l'essentiel des applications traitées par les auteurs a été faite en utilisant effectivement la T.F.D. comme décomposition orthogonale.

Le choix de la base étant fait, considérons l'effet de l'estimation du signal par un développement tronqué à l'ordre  $p$ , sous la forme :

$$(7) \quad \hat{\mathbf{s}}_p = \sum_{i=0}^{p-1} \hat{a}_i \mathbf{v}_i.$$

L'estimée paramétrique au sens des moindres carrés, qui serait aussi l'estimée du maximum de vraisemblance si le bruit était gaussien, minimise la distance  $\|\hat{\mathbf{s}}_p - \hat{\mathbf{u}}\|$  à l'observation  $\hat{\mathbf{u}}$ . La solution correspondante  $\hat{a}_i$  pour les coefficients est alors simplement :

$$(8) \quad \begin{cases} \hat{a}_i = \alpha_i & \text{pour} \quad i \in \{0, \dots, p-1\}, \\ \hat{a}_i = 0 & \text{pour} \quad i \geq p. \end{cases}$$

L'écart entre l'estimée  $\hat{\mathbf{s}}_p$  des moindres carrés et le signal  $\hat{\mathbf{s}}$  peut aussi s'exprimer, compte tenu de (5) et (8), par :

$$(9) \quad \hat{\mathbf{s}}_p - \hat{\mathbf{s}} = \sum_{i=0}^{p-1} \beta_i \mathbf{v}_i - \sum_{i=p}^{M-1} a_i \mathbf{v}_i.$$

Dans ces conditions l'erreur totale d'estimation est :

$$(10) \quad Q(p) = E\{|\hat{\mathbf{s}}_p - \hat{\mathbf{s}}|^2\} = p\sigma^2 + \sum_{i=p}^{M-1} |a_i|^2.$$

Cette expression fait apparaître deux termes :

- a) un terme  $p\sigma^2$  de variance, croissant avec  $p$ ,
- b) un terme  $\sum_{i=p}^{M-1} |a_i|^2$  de biais, décroissant en fonction de  $p$ .

La valeur optimale de l'ordre  $p$  du développement serait celle minimisant  $Q(p)$ . L'inconvénient est que cette erreur ne peut être calculée exactement à partir de l'observation, les  $a_i$  étant précisément les inconnues à estimer. Nous allons maintenant en introduire une estimée  $\hat{Q}(p)$  utilisable en pratique.

### 2.1.2. Optimisation du compromis biais-variance et test

D'après (5) et compte tenu du fait que ce sont les  $\{\alpha_i\}$  et non les  $\{a_i\}$  qui sont connus, une estimation non biaisée de  $\hat{Q}(p)$  nous est donnée en remplaçant dans (10)  $|a_i|^2$  par  $(|\alpha_i|^2 - \sigma^2)$ , qui a même espérance. Nous obtenons :

$$(11) \quad \hat{Q}(p) = (2p - M)\sigma^2 + \sum_{i=p}^{M-1} |\alpha_i|^2$$

avec  $E[\hat{Q}(p)] = Q(p)$ .

En pratique  $\hat{Q}(p)$  a une variance trop grande pour donner une bonne estimation de l'erreur réelle  $Q(p)$ . Par contre le minimum de  $\hat{Q}(p)$  est atteint pour une valeur  $L'$  de  $p$  voisine de  $L$  :  $\hat{Q}(p)$  est donc utilisable comme critère d'estimation de la largeur optimale de la fenêtre de troncature.

Le résultat fourni par (11) peut être comparé au critère MAICE d'Akaike [6]. En effet pour des  $\{\beta_i\}$  gaussiens, il est possible de vérifier que le critère MAICE appliqué à l'estimation paramétrique de  $\hat{s}$  par le modèle tronqué à l'ordre  $p$  de  $\hat{s}_p$ , pour le modèle statistique considéré, conduit au même résultat que  $\hat{Q}(p)$ .

Le critère  $\hat{Q}(p)$  défini par (11) peut être calculé de manière récursive. En effet :

$$(12) \quad \hat{Q}(p+1) - \hat{Q}(p) = 2\sigma^2 - |\alpha_p|^2.$$

Pour des raisons de simplicité, dans la mesure où ce n'est que la valeur  $L'$  que l'on veut déterminer, nous avons utilisé le critère équivalent :

$$(13) \quad \hat{Q}'(p) = \sum_{i=1}^{p-1} [|\alpha_i|^2 - 2\sigma^2],$$

égal à  $-\hat{Q}(p)$  à une constante près. En effet,  $\hat{Q}'(p)$  se calcule récursivement à partir de  $\hat{Q}'(0) = 0$  par la relation :

$$(14) \quad \hat{Q}'(p+1) - \hat{Q}'(p) = |\alpha_p|^2 - 2\sigma^2.$$

La valeur  $L'$  correspond donc cette fois au *maximum* de  $\hat{Q}'(p)$ .

Notons, que les parties croissantes de  $\hat{Q}'(p)$  correspondent à  $|\alpha_p|^2 > 2\sigma^2$  soit, en moyenne, à  $|\alpha_p|^2 > \sigma^2$ , les parties décroissantes à  $|\alpha_p|^2 < \sigma^2$ . Le maximum du critère peut aussi s'interpréter comme le point au-delà duquel le « rapport signal à bruit »  $|\alpha_p|^2/\sigma^2$  reste en moyenne inférieur à 1.

A partir de la connaissance du  $L'$  optimum nous avons introduit un test qui va nous permettre de « juger » de la validité des hypothèses sur le bruit. En escomptant que pour  $i$  supérieur à  $L'$  les  $\{a_i\}$  sont en moyenne très inférieurs au bruit, la variance de celui-ci peut être estimée et comparée à la valeur escomptée  $\sigma^2$  à partir du test

$$(15) \quad T(L') = \frac{\hat{\sigma}^2}{\sigma^2} = \frac{1}{\sigma^2} \frac{1}{(M-L')} \sum_{i=L'}^{M-1} |\alpha_i|^2$$

qui doit être voisin de 1. Ce test, que nous n'avons pas cherché particulièrement à affiner, est surtout utile en pratique pour déceler d'éventuelles erreurs ou inadéquations du modèle.

Notons que tronquer l'ordre du développement équivaut à appliquer une fenêtre de pondération rectangulaire sur le signal  $\alpha_i$  dans la base des vecteurs  $\{v_i\}$ . Dans le sous-paragraphe qui suit nous allons envisager l'effet d'une fenêtre de forme plus générale (le choix de la forme de la fenêtre devant en outre permettre de maintenir le caractère défini non négatif de la matrice de densité spectrale).

### 2.1.3. Détermination plus générale d'une fenêtre de largeur optimum

Un estimateur de  $\hat{s}$  plus général peut être obtenu par pondération des paramètres  $\{\alpha_i\}$  par une fenêtre

$w(i)$ ; au lieu de (7) l'estimée de  $\hat{s}$  est prise de la forme :

$$(16) \quad \hat{s} = \sum_{i=0}^{M-1} w(i) \alpha_i v_i.$$

En fonction du choix de la fenêtre  $w(\cdot)$ , l'erreur totale d'estimation est donnée d'après (4) et (16) par :

$$(17) \quad Q(w(\cdot)) = E\{|\hat{s} - \hat{s}|^2\} = \sigma^2 \sum_{i=0}^{M-1} w^2(i) + \sum_{i=0}^{M-1} |\alpha_i|^2 (w(i) - 1)^2.$$

Compte tenu du fait que les  $\{a_i\}$  sont inconnus et de la relation (5), on remplace (17) par une estimation non biaisée de  $Q(w(\cdot))$ , fournie par :

$$(18) \quad \hat{Q}(w(\cdot)) = \sigma^2 \sum_{i=0}^{M-1} w^2(i) + \sum_{i=0}^{M-1} (|\alpha_i|^2 - \sigma^2) (w(i) - 1)^2.$$

Ainsi, dans une famille de fenêtres de largeur croissante, la minimisation du critère  $\hat{Q}(w(\cdot))$  permet de déterminer une largeur de fenêtre optimum.

Dans le cas où la T.F.D. est utilisée comme transformation orthogonale, au produit de  $\alpha_i$  par la *fenêtre de pondération*  $w(i)$ , correspond le produit de convolution des T.F.D. : convolution de  $\hat{\alpha}_i$  par la *fenêtre de lissage*  $\hat{w}(i)$ . A la fenêtre de pondération  $w(i)$  de largeur optimale correspond par T.F.D. la fenêtre de lissage  $\hat{w}(i)$  de largeur optimale. Resterait à choisir la *forme* des fonctions de la famille considérée pour obtenir une fenêtre de lissage sans lobes et si possible positive.

Une autre approche pour optimiser la fenêtre consisterait à chercher les  $\{w(i)\}$  qui minimisent l'erreur  $Q(w(\cdot))$ . Dès lors, il est facile de vérifier d'après (17) que le minimum serait obtenu pour la valeur théorique suivante des facteurs de pondération :

$$(19) \quad w(i) = \frac{|\alpha_i|^2}{|\alpha_i|^2 + \sigma^2}.$$

Cette relation a un intérêt purement théorique dans la mesure où les  $\{a_i\}$  sont évidemment inconnus. Elle appelle cependant un certain nombre de remarques importantes pour la suite. Nous voyons que les facteurs de pondération  $\{w(i)\}$  sont compris entre zéro et un. Ils tendent vers un dès que  $|\alpha_i|^2 \gg \sigma^2$  et vers zéro pour  $|\alpha_i|^2 \ll \sigma^2$ . La mi-hauteur ( $w(i) = \frac{1}{2}$ ) correspond à  $|\alpha_i|^2 = \sigma^2$ . Remarquons aussi qu'imposer a priori la forme de la fenêtre de pondération en n'ajustant que sa largeur est assez arbitraire et risque de donner une solution loin de « l'optimal » correspondant à (19).

Dans les paragraphes suivants, compte tenu de ces remarques, nous proposons deux méthodes pratiques

simples de détermination des fenêtres, basées sur la recherche d'intervalles où l'on peut à partir de l'observation, considérer que les  $\{|a_i|^2\}$  sont soit significativement plus grands que  $\sigma^2$ , soit plus petits que  $\sigma^2$ .

**2.1.4. Détermination pratique d'une fenêtre à largeur et « raideur » variable**

Restons dans l'hypothèse où le signal est concentré sur les premières composantes  $\{a_i\}$ . Pour simplifier, imaginons de plus que  $|a_i|$  soit fonction non croissante de  $i$ . Dans ces conditions la fenêtre optimale (19) est elle-même fonction non croissante de  $i$ . Comme nous l'avons déjà remarqué, ce n'est que pour les valeurs de  $|a_i|^2$  de l'ordre de  $\sigma^2$  que la valeur de  $w(i)$  est indéterminée car, soit pour  $|a_i|^2 \gg \sigma^2$ , soit pour  $|a_i|^2 \ll \sigma^2$ ,  $w(i)$  se confond respectivement avec 1 ou 0. Le passage de  $w(i)$  de 0,9 à 0,1 correspond à l'intervalle de transition pour  $|a_i|^2$ :

$$(20) \quad 9 \sigma^2 \geq |a_i|^2 \geq 1/9 \sigma^2.$$

Nous proposons, pour réaliser une « apodisation » de la fenêtre rectangulaire du paragraphe précédent, une fenêtre  $w(i)$  avec une transition de 0 à 1 en « cosinus ajusté » (figure 1) avec une largeur à mi-hauteur  $L'$  et un palier à 1 de largeur  $L'' < L'$ :

$$(21) \quad w(i) = 1 \quad \text{pour } i \leq L''.$$

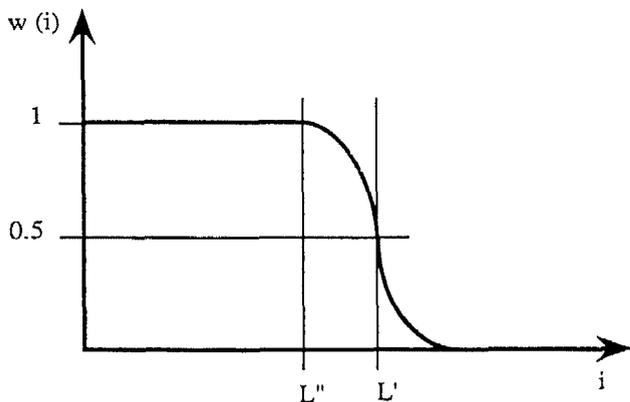


Figure 1. — Fenêtre en cosinus ajusté.

Ce gabarit de fenêtre, continu et à dérivée continue, assure une atténuation des lobes (« apodisation ») de la fenêtre de lissage correspondante. Le choix de  $L''$  indépendant de  $L'$  permet une adaptation de la « raideur » de la coupure en fonction du signal.

Pour ce choix de  $L'$  nous avons gardé, pour sa simplicité, le critère général (13), qui donne la largeur du filtre rectangulaire équivalent. Pour déterminer  $L''$  nous allons essayer d'estimer l'indice correspondant à la limite de l'intervalle de transition (20),  $|a_i|^2 \approx 9 \sigma^2$ . C'est-à-dire l'indice au-delà duquel, en moyenne

$$(22) \quad |a_i|^2 - 9 \sigma^2 < 0.$$

Pour ceci nous pouvons introduire un critère moyen analogue à  $Q'(p)$ , en recherchant le maximum de

$$(23) \quad Q''(p) = \sum_{i=1}^{p-1} (|a_i|^2 - 9 \sigma^2).$$

Au-delà du maximum absolu, les accroissements de  $Q''(p)$  seront bien en moyenne négatifs et l'inégalité (22) satisfaite.

En pratique, à partir de  $\alpha_p$  au lieu de  $a_p$ , la valeur  $L''$  sera obtenue en maximisant  $\hat{Q}''(p)$  calculé récursivement par :

$$(24) \quad \hat{Q}''(p+1) - \hat{Q}''(p) = |\alpha_p|^2 - 10 \sigma^2.$$

Par cette méthode de choix de  $L'$  et  $L''$ , il est donc possible de déterminer, relativement simplement, une fenêtre de lissage en cosinus ajusté permettant d'optimiser le compromis biais-variance, si nous connaissons la puissance du bruit blanc additif.

**2.1.5. Généralisation : signal non localisé autour de zéro**

Nous avons supposé jusqu'ici que le signal  $\hat{s}$  décomposé dans la base des  $\{v_i\}$  était tel que les  $\{a_i\}$  les plus grands étaient centrés (localisés) autour de zéro (voir fig. 1). En pratique, nous trouverons des situations différentes, en particulier lors du passage des autospectres aux interspectres. En effet, si nous choisissons pour base celle des exponentielles complexes, le signal obtenu dans cette base caractérise une fonction de corrélation : nous savons que dans le cas de l'intercorrélation cette fonction n'est pas, dans le cas le plus général, maximale en zéro, en particulier en présence de retard entre voies.

Ainsi et de manière générale, nous n'utiliserons plus une fenêtre unimodale au voisinage de  $i = 0$ , mais une fenêtre « multimodale », avec autant de modes que d'intervalles « significatifs », c'est-à-dire, d'après (19), d'intervalles où le signal est significativement supérieur au bruit. Pratiquement ces intervalles sont définis comme ceux pour lesquels la puissance du signal est supérieure à la variance du bruit : ceci afin d'être en accord avec les remarques du sous-paragraphe 2.1.2. Pour ce faire nous utilisons un critère somme  $\hat{S}(p)$  (voir (A.1) dans l'annexe A) analogue à (13) en sélectionnant les intervalles où cette somme est en moyenne croissante. Ces intervalles étant déterminés, chacune des fenêtres (en particulier leur « raideur ») est définie comme indiqué au sous-paragraphe 2.1.4. L'algorithme permettant de déterminer les intervalles significatifs est décrit en détail dans l'annexe A.

Dans les deux sous-paragraphe suivants nous allons montrer comment, par une transformation appropriée, nous nous ramenons au modèle à bruit additif blanc, dans le cas de l'estimation d'un autospectre (par transformation logarithmique) ou dans le cas d'un interspectre (par blanchiment). Après transformation, nous pourrions appliquer la méthode de lissage optimal que nous venons de discuter.

**2.2. LISSAGE OPTIMAL POUR LES AUTOSPECTRES**

Les composantes  $\hat{x}(m)$  de la T.F.D. d'un signal réel présentent la symétrie hermitienne. Pour

$0 < m < \frac{N}{2}$ , si le signal est blanc gaussien, nous savons que les parties réelles et imaginaires des  $\hat{x}(m)$  constituent N variables aléatoires gaussiennes indépendantes de même variance. Les modules  $|\hat{x}(m)|^2$  sont donc des variables en khi-deux indépendantes à deux degrés de liberté. Cette propriété reste vraie asymptotiquement pour un signal stationnaire non blanc si l'horizon d'observation tend vers l'infini [7]. Elle demeure une bonne approximation à horizon N fini, pourvu que N reste grand devant le temps de corrélation du signal [8, 9]. Ceci équivaut au fait que le spectre est « lentement variable », dans le sens où les variations relatives de  $\hat{\gamma}(m)$  à  $\hat{\gamma}(m+1)$  sont faibles devant 1 :

$$(25) \quad \left| \frac{\hat{\gamma}(m+1) - \hat{\gamma}(m)}{\hat{\gamma}(m)} \right| \ll 1.$$

Dans les mêmes conditions, l'espérance mathématique de  $|\hat{x}(m)|^2$  se relie directement à  $\hat{\gamma}(m)$  [1] par :

$$(26) \quad E[|\hat{x}(m)|^2] \cong N\hat{\gamma}(m).$$

Il est donc naturel d'introduire le « périodogramme » par la relation :

$$(27) \quad \hat{\pi}(m) \triangleq \frac{1}{N} |\hat{x}(m)|^2.$$

Compte tenu de la discussion précédente, pour un spectre lentement variable, le périodogramme est une variable en khi-deux que l'on peut normaliser par son espérance  $\hat{\gamma}(m)$  pour faire apparaître une variable normée  $\varepsilon(m)$  :

$$(28) \quad \hat{\pi}(m) = E[\hat{\pi}(m)] |\varepsilon(m)|^2 \cong \hat{\gamma}(m) |\varepsilon(m)|^2.$$

Le périodogramme s'interprète donc comme le spectre à estimer  $\hat{\gamma}(m)$  perturbé par le bruit blanc « multiplicatif »  $|\varepsilon(m)|^2$  de variance unité.

En prenant le logarithme de (26) nous revenons au modèle proposé dans le paragraphe précédent. Cette approche est celle de Wahba [5] mais elle conduit à une augmentation notable de la variance, à fenêtre donnée, si le spectre est estimé directement par lissage du Log-périodogramme. Pour éviter cet inconvénient Clergeot [8] a proposé de faire un « pré-moyennage » sur le périodogramme avant de prendre le logarithme.

Le périodogramme pré-moyenné sur P termes est défini par :

$$(29) \quad \hat{\pi}_p(m) = \frac{1}{NP} \sum_{p=m-P/2}^{m+P/2} |\hat{x}(m-p)|^2.$$

Les composantes sont prises de P en P pour conserver l'indépendance. D'après (26) :

$$(30) \quad E[\hat{\pi}_p(m)] = \frac{1}{P} \sum_{p=m-P/2}^{m+P/2} \hat{\gamma}(m-p), \\ \cong \hat{\gamma}(m) \quad (\text{si } P \text{ est petit}).$$

La transformation logarithmique proposée est alors (voir annexe B) :

$$(31) \quad \hat{u}(m) = \text{Log } \hat{\pi}_p(m) - \eta_p = \hat{s}(m) + \hat{b}(m)$$

où  $\eta_p$  est un terme de correction de biais tel que :

$$(32) \quad \begin{cases} E[\hat{u}(m)] &= \hat{s}(m) = \text{Log } \hat{\gamma}(m), \\ E[\hat{b}(m)] &= 0, \\ E[|\hat{b}(m)|^2] &= \sigma_b^2 = \frac{1}{P} (1 + \delta_p). \end{cases}$$

La transformation (31) nous ramène bien à un modèle signal inconnu (le logarithme du spectre) plus bruit blanc de variance connue. Le facteur  $(1 + \delta_p)$  sur la variance dans (32) correspond à un accroissement de variance lié à la transformation logarithmique. Il est égal à 1,645 sans pré-moyennage ( $P = 1$ ), il tend rapidement vers 1 quand P augmente (1,185 pour  $P = 3$ ). Notons que le pré-moyennage fixe une limite inférieure à la largeur de la fenêtre de lissage finale : il n'y a pas intérêt à l'augmenter au-delà de 3, ou 5 au maximum ( $\delta_p = 0,107$  pour  $P = 5$ ), sous peine de limiter la résolution.

La démarche suivie pratiquement pour estimer de manière optimale l'autospectre d'un signal « à variations lentes » est résumée sur l'organigramme de la figure 2.

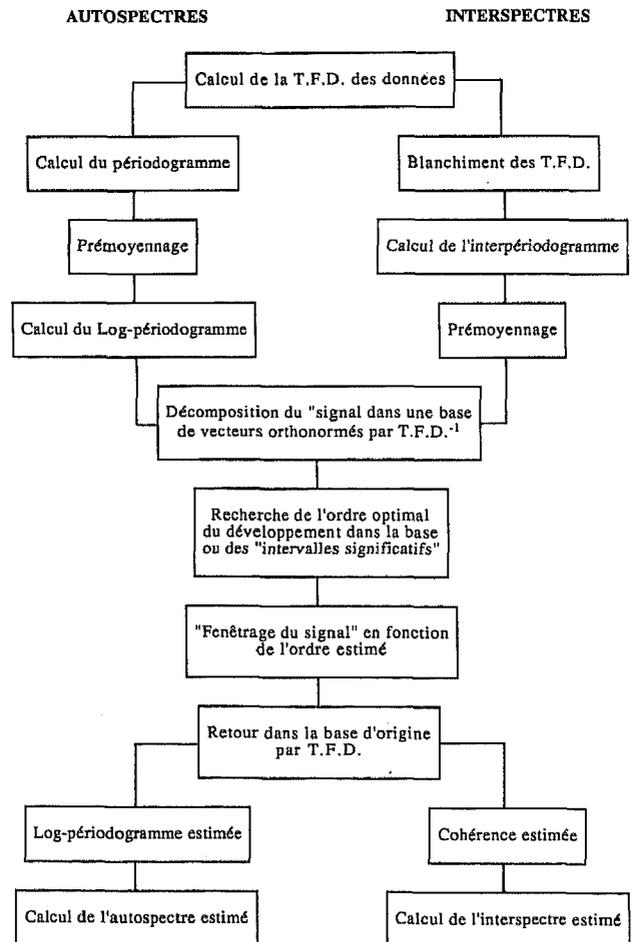


Figure 2. — Organigramme résumant la méthode du lissage optimal pour les autospectres et les interspectres.

### 2.3. LISSAGE OPTIMAL POUR LES INTERSPECTRES

Dans ce sous-paragraphe nous montrons comment, à partir de « l'interpériodogramme » et pour estimer un interspectre, il est possible de se ramener au modèle défini par la relation (3). Pour ceci nous introduisons les notions de blanchiment et de cohérence, les autospectres étant supposés connus.

Si nous considérons deux signaux  $\hat{x}_i(m)$  et  $\hat{x}_j(m)$ , leurs autospectres respectifs  $\hat{\gamma}_{ii}(m)$  et  $\hat{\gamma}_{jj}(m)$  et leur interspectre  $\hat{\gamma}_{ij}(m)$ , nous savons que la cohérence de ces signaux est définie par :

$$(33) \quad \hat{c}_{ij}(m) = \hat{\gamma}_{ij}(m) / [\hat{\gamma}_{ii}(m) \hat{\gamma}_{jj}(m)]^{1/2}.$$

Au lieu d'estimer l'interspectre, nous allons estimer la cohérence  $\hat{c}_{ij}(m)$  : l'interspectre s'en déduisant d'après (33).

L'estimation de la cohérence  $\hat{c}_{ij}(m)$  peut aussi s'interpréter comme l'estimation de l'interspectre des signaux  $x_i(m)$  et  $x_j(m)$  blanchis. Dans le domaine fréquentiel, une bonne approximation des signaux blanchis  $\hat{x}'_i(m)$  et  $\hat{x}'_j(m)$  est donnée par [2] :

$$(34) \quad \begin{aligned} \hat{x}'_i(m) &= \hat{x}_i(m) / \hat{\gamma}_{ii}^{1/2}(m) \\ \text{et} \quad \hat{x}'_j(m) &= \hat{x}_j(m) / \hat{\gamma}_{jj}^{1/2}(m). \end{aligned}$$

Soit  $\hat{\pi}'_{ij}(m)$  l'interpériodogramme des signaux blanchis :

$$(35) \quad \hat{\pi}'_{ij}(m) = \frac{1}{N} \hat{x}'_i(m) \hat{x}'_j{}^*(m).$$

De (26) et (34) nous déduisons :

$$(36) \quad E[\hat{\pi}'_{ij}(m)] = \tilde{\gamma}_{ij}(m) / [\hat{\gamma}_{ii}(m) \hat{\gamma}_{jj}(m)]^{1/2} \triangleq \tilde{c}_{ij}(m)$$

où  $\tilde{c}_{ij}(m)$  est voisin de  $\hat{c}_{ij}(m)$  dans le cadre des hypothèses que nous avons faites.

Si nous posons :

$$(37) \quad \hat{\pi}_{ij}(m) = \tilde{c}_{ij}(m) + \hat{\varepsilon}(m),$$

$\hat{\varepsilon}(m)$  a une moyenne nulle d'après (36) et est décorrélé de  $\hat{\varepsilon}(m')$  pour  $m$  différent de  $m'$  ; nous pouvons montrer, sous l'hypothèse gaussienne, que (voir annexe C, relation (C.8)) :

$$(38) \quad E[|\hat{\varepsilon}(m)|^2] = 1.$$

Par suite, si nous faisons un lissage de « l'interpériodogramme  $\hat{\pi}'_{ij}(m)$  » sur  $P$  composantes et prenons les valeurs de  $P$  en  $P$ , nous sommes ramenés, d'après (37) et (38), à un modèle de signal analogue à celui du sous-paragraphe précédent, à savoir :

$$(39) \quad \hat{u}(m) = \hat{s}(m) + \hat{b}(m),$$

où  $\hat{u}(m)$  représente une estimation de la cohérence à la fréquence  $m$  après prémoyennage sur  $P$  points,  $\hat{s}(m)$  la cohérence moyenne que nous voulons esti-

mer et  $\hat{b}(m)$  un bruit blanc centré qui tend vers un bruit gaussien de variance  $1/P$  si  $P$  est assez grand.

Dans le cadre de cette estimation le signal  $\hat{u}(m)$  est généralement non « localisé », au sens du paragraphe 2.1.5 (présence de retard entre les signaux  $\hat{x}'_i(m)$  et  $\hat{x}'_j(m)$ ). De ce fait, la recherche des intervalles significatifs (intervalles où le signal  $\hat{s}(m)$  est présent) devra se faire selon la méthode proposée au sous-paragraphe 2.1.5.

La démarche suivie pratiquement pour estimer de manière optimale l'interspectre de deux signaux « à variations lentes » est résumée sur l'organigramme de la figure 2.

### 2.4. LISSAGE OPTIMAL POUR LA MATRICE DE DENSITÉ SPECTRALE

Dans ce sous-paragraphe nous présentons un algorithme itératif d'estimation spectrale multivariable. Cet algorithme a pour tâche d'optimiser l'estimation de chacun des éléments d'une matrice de densité spectrale  $\hat{\Gamma}(m)$  en utilisant les résultats précédents (obtenus dans le cas scalaire) tout en maintenant le caractère défini non négatif de la matrice.

Dans le cas scalaire la recherche de la densité spectrale d'un signal  $\hat{x}(m)$  peut se ramener à la recherche du filtre blanchisseur  $\hat{g}(m)$  transformant  $\hat{x}(m)$  en un bruit blanc normé  $\hat{\varepsilon}(m)$  ;  $\hat{g}(m)$  vérifie donc :

$$(40) \quad \hat{\gamma}_{xx}(m) = \hat{\gamma}_{\varepsilon\varepsilon}(m) [|\hat{g}(m)|^2]^{-1} = |\hat{g}(m)|^{-2}.$$

Dans le cas multivariable la recherche de la matrice de densité spectrale d'un signal vectoriel  $\hat{\mathbf{x}}(m)$  peut se ramener à la recherche de la matrice « blanchisseuse »  $\hat{\mathbf{G}}(m)$  telle que si  $\hat{\mathbf{\varepsilon}}(m)$  est égal à  $\hat{\mathbf{G}}(m) \hat{\mathbf{x}}(m)$  :

$$(41) \quad \begin{aligned} \hat{\Gamma}_{xx}(m) &= E[\hat{\mathbf{x}}(m) \hat{\mathbf{x}}^+(m)] = \\ &= \hat{\mathbf{G}}^{-1}(m) \hat{\Gamma}_{\varepsilon\varepsilon}(m) [\hat{\mathbf{G}}^+(m)]^{-1}, \end{aligned}$$

où  $\hat{\Gamma}_{xx}(m)$  est définie par  $E[\hat{\varepsilon}(m) \hat{\varepsilon}^+(m)]$  et est égale à la matrice identité.

En fait, il existe une infinité de matrices  $\hat{\mathbf{G}}(m)$  répondant à la question, cette infinité correspondant à l'ensemble des factorisations possibles de la matrice  $\hat{\Gamma}_{xx}(m)$ . L'algorithme que nous proposons revient à factoriser  $\hat{\Gamma}_{xx}(m)$  sous la forme du produit de deux matrices triangulaires dans lesquelles les termes diagonaux sont réels et positifs ce qui assure l'unicité de la factorisation.

Formellement l'algorithme est équivalent à un algorithme d'orthogonalisation de Gram-Schmidt pour une série de  $K$  vecteurs [10].

La méthode consiste à blanchir et à décorréler spatialement les signaux par transformations linéaires successives comme le montre la figure 3 (l'indice fréquence est omis et le sera dans la suite pour alléger les notations).

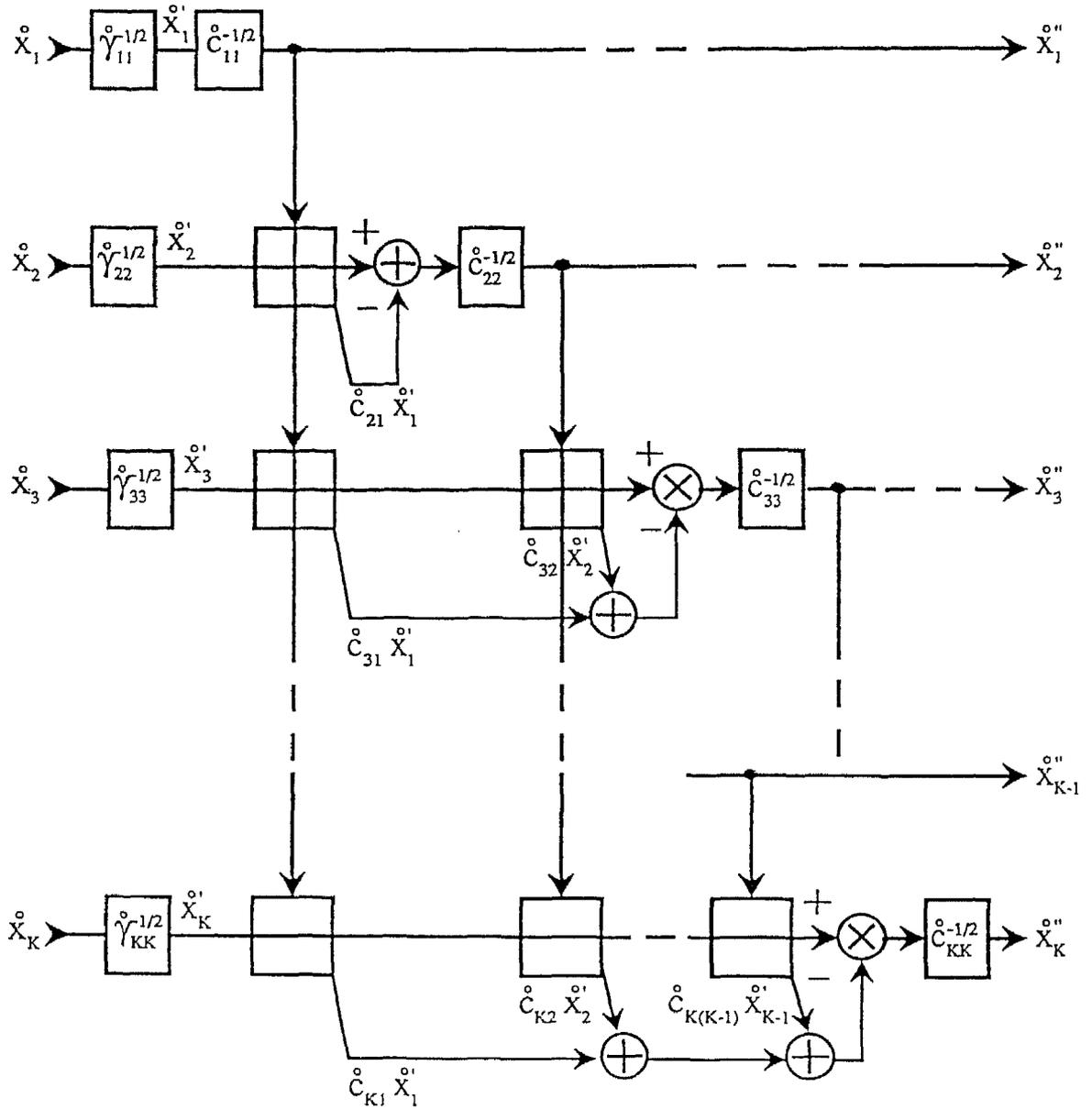


Figure 3. — Algorithme de blanchiment matriciel et de décorrélation spatiale.

La première étape de cet algorithme consiste à blanchir les K signaux obtenus par T.F.D. :

$$(42) \quad \hat{x}_i' = \hat{x}_i / \hat{\gamma}_{ii}^{1/2} \text{ avec } i \in \{1, K\},$$

l'estimation des  $\hat{x}_i'$  se déduisant de celle des  $\hat{\gamma}_{ii}$  selon la méthode proposée au sous-paragraphe 2.2.

La deuxième étape consiste à extraire de chaque signal  $\hat{x}_i'$  la partie décorrélée de l'ensemble des autres signaux à l'aide des cohérences. Par exemple, à partir de la connaissance de la cohérence entre les signaux  $\hat{x}_2'$  et  $\hat{x}_1'$  :

$$(43) \quad \hat{c}_{21} = E \{ \hat{x}_2' \hat{x}_1'^* \},$$

cette cohérence étant estimée selon la méthode

proposée au sous-paragraphe 2.3, nous pouvons écrire :

$$(44) \quad \hat{x}_2' = \hat{c}_{21} \hat{x}_1'' + \hat{x}_{i2}' ,$$

expression dans laquelle  $\hat{x}_{i2}'$  est la partie du signal  $\hat{x}_2'$  décorrélée avec  $\hat{x}_1'$ .

La troisième étape consiste à blanchir les signaux ainsi décorrélés :

$$(45) \quad \hat{x}_i'' = \hat{x}_{di}' / \hat{\gamma}_{di}^{1/2} \text{ avec } i \in \{1, K\} .$$

En réitérant ce raisonnement sur l'ensemble des signaux d'entrée nous obtenons :

$$(46) \quad \hat{x}_j' = \hat{c}_{jj}^{1/2} \hat{x}_j'' + \sum_{i=1}^{j-1} \hat{c}_{ji} \hat{x}_i'' \text{ avec } \hat{c}_{jj} = \hat{\gamma}_{jj} .$$

Ainsi, en posant :

$$(47) \quad \begin{cases} \hat{\mathbf{x}} = [\hat{x}_1, \dots, \hat{x}_K]^T, \\ \hat{\mathbf{x}}' = [\hat{x}'_1, \dots, \hat{x}'_K]^T, \\ \hat{\mathbf{x}}'' = [\hat{x}''_1, \dots, \hat{x}''_K]^T, \end{cases}$$

il vient, d'après (46) :

$$(48) \quad \hat{\mathbf{x}} = \begin{bmatrix} \hat{\gamma}_{11}^{1/2} & (0) \\ & \ddots \\ (0) & \hat{\gamma}_{KK}^{1/2} \end{bmatrix} \hat{\mathbf{x}}' = \hat{\mathbf{D}} \hat{\mathbf{x}}'$$

et

$$(49) \quad \hat{\mathbf{x}}' = \begin{bmatrix} \hat{c}_{11}^{1/2} & & (0) \\ \vdots & \ddots & \\ \hat{c}_{j1} & \dots & \hat{c}_{jj}^{1/2} & \dots \\ \vdots & & \ddots & \\ \hat{c}_{K1} & \dots & \dots & \hat{c}_{KK}^{1/2} \end{bmatrix} \hat{\mathbf{x}}'' = \hat{\mathbf{C}} \hat{\mathbf{x}}'' \text{ avec } \hat{c}_{11} = 1.$$

Par suite,

$$(50) \quad \hat{\mathbf{x}} = \hat{\mathbf{D}} \hat{\mathbf{C}} \hat{\mathbf{x}}'' \triangleq \hat{\mathbf{A}} \hat{\mathbf{x}}''$$

et la matrice de densité spectrale  $\hat{\Gamma}_{xx}$  se met sous la forme factorisée :

$$(51) \quad \hat{\Gamma}_{xx} = E \{ \hat{\mathbf{x}} \hat{\mathbf{x}}^+ \} = E \{ \hat{\mathbf{A}} \hat{\mathbf{x}}'' \hat{\mathbf{x}}''^+ \hat{\mathbf{A}}^+ \} = \hat{\mathbf{A}} E \{ \hat{\mathbf{x}}'' \hat{\mathbf{x}}''^+ \} \hat{\mathbf{A}}^+ = \hat{\mathbf{A}} \hat{\mathbf{A}}^+,$$

ce qui assure à cette matrice son caractère défini non négatif tout en ayant optimisé l'estimation de chacun de ses éléments.

## 2.5. CONCLUSIONS ET RÉSULTATS DE SIMULATIONS

Pour l'estimation d'une matrice de densité spectrale la méthode que nous venons de décrire présente deux principaux avantages par rapport aux méthodes classiques : elle est automatique (pas de choix empirique cas par cas pour l'estimation de la largeur de la fenêtre de lissage) et elle est optimale dans le cas de signaux gaussiens à spectres continus lentement variables.

Si ces hypothèses de spectre à variations lentes ne sont plus vérifiées (en présence de raies étroites ou pour un modèle sinusoïdes plus bruit par exemple), nous ne pouvons plus prétendre à l'optimalité, mais les résultats ne sont pas aberrants dans la mesure où le critère aura tendance à sélectionner la résolution la plus fine. Ce cas sort cependant du cadre de cet article. Nous ne présentons ci-après que des exemples qui correspondent aux hypothèses de départ et justifient l'intérêt de la méthode.

Dans les exemples que nous présentons ci-après, la base associée aux  $\{v(i)\}$  de la relation (4) est celle des exponentielles complexes : la T.F.D. est utilisée comme transformation orthogonale. Dans ces conditions le fenêtrage de  $\hat{\mathbf{u}}$  par les  $\{w(i)\}$  dans (16) équivaut, d'une manière générale, au fenêtrage de la fonction de corrélation par la fenêtre d'apodisation :

la largeur optimale de cette fenêtre étant définie à partir de (14) et (19).

Sur la figure 4, nous montrons :

- a) en A un bruit blanc après passage dans un filtre dont le carré du module caractérise l'autospectre vrai (pointillés blancs) ;
- b) en B la comparaison entre notre estimation et l'autospectre vrai de A, et nous donnons la valeur du test, défini par (15), qui théoriquement devrait être égal à un ;
- c) en C et D la comparaison entre la méthode classique du lissage du périodogramme (pour deux valeurs de la fenêtre de lissage) et l'autospectre vrai. Sur la figure 5, à partir de deux signaux définis comme en figure 4 A, nous montrons :

- a) en A l'interpériodogramme (module) de deux bruits blancs filtrés ;
- b) en B la comparaison entre notre « estimation directe » de l'interspectre (estimation faite selon la procédure décrite au sous-paragraphe 2.3) et la vraie valeur de ce dernier, et nous donnons la valeur du test qui, ici aussi, devrait être égal à un ;
- c) en C la comparaison entre notre « estimation matricielle » de l'interspectre (estimation faite selon la procédure décrite au sous-paragraphe 2.4) et la vraie valeur de ce dernier ;
- d) en D la comparaison entre la méthode classique du lissage de l'interpériodogramme (pour une valeur de la fenêtre de lissage) et l'interspectre vrai.

Ces deux figures justifient l'intérêt de notre méthode (optimisation du compromis biais-variance) par rapport aux méthodes classiques où le choix de la largeur équivalente de la fenêtre de lissage est fait de manière plus ou moins empirique.

L'intérêt de notre méthode n'est pas dégradé par la durée des calculs. En effet, sur un PC-AT, la méthode que nous proposons est seulement trois fois moins rapide que celle du lissage du périodogramme classique. Notre algorithme d'estimation des autospectres est actuellement implanté sur une carte de calcul vectoriel travaillant en flottant et pouvant s'intégrer dans un PC-AT ou un COMPAQ 386/20. Dès lors, pour estimer 64 autospectres à partir de 64 échantillons de 1024 points il faut sur un PC-AT 4 minutes, sur un COMPAQ 386/20 40 secondes et par l'intermédiaire de la carte de calcul vectoriel moins de 3 secondes : un résultat qui atténue le degré de complexité de notre algorithme.

Cette méthode a fait l'objet de dépôts de brevet [11] compte-tenu de l'application directe qu'elle a trouvée dans le domaine de la « Cartographie Cérébrale » [12].

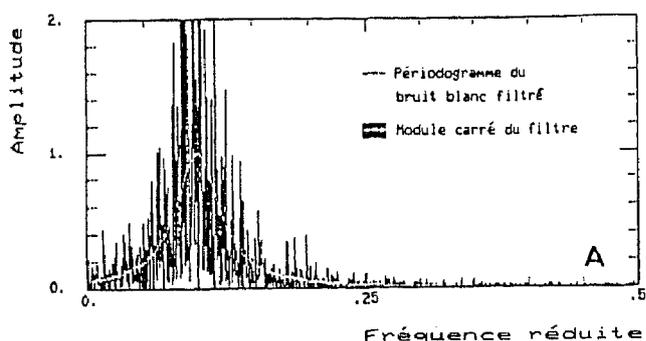
*Manuscrit reçu le 20 décembre 1988.*

## Remerciements

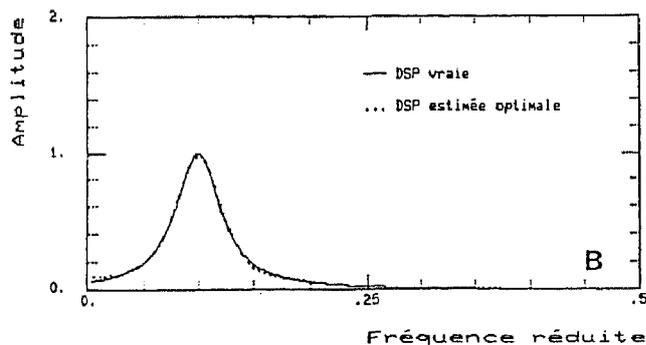
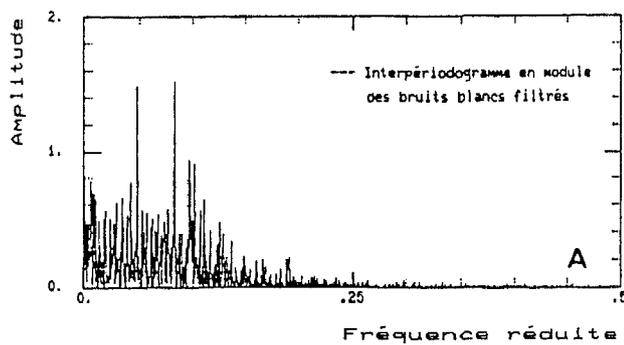
Nous tenons, pour l'optimisation des temps de calcul et pour l'implantation sur la carte de calcul vectoriel de notre algorithme d'estimation des autospectres, à remercier MM. J. STRUB (Ingénieur du LSS) et A. MAUPIN (Stagiaire de DEA) qui grâce à leur ténacité ont pu nous permettre de justifier pratiquement tout l'intérêt de la méthode d'estimation de la

matrice de densité spectrale optimisant le compromis biais-variance.

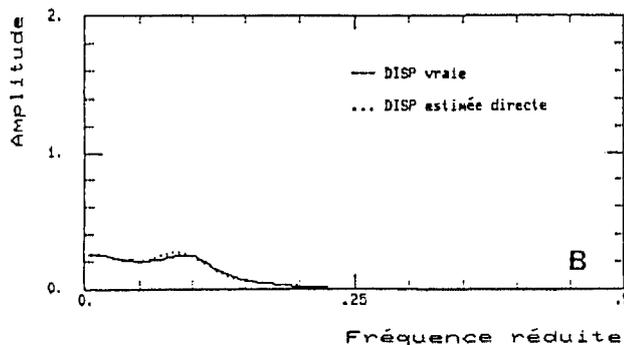
Manuscrit reçu le 20 décembre 1988, Version révisée le 5 mars 1990.



Horizon temporel: 2048



Test: 1.07 Largeur équivalente: 50



Test: 0.99 Largeur équivalente: 150

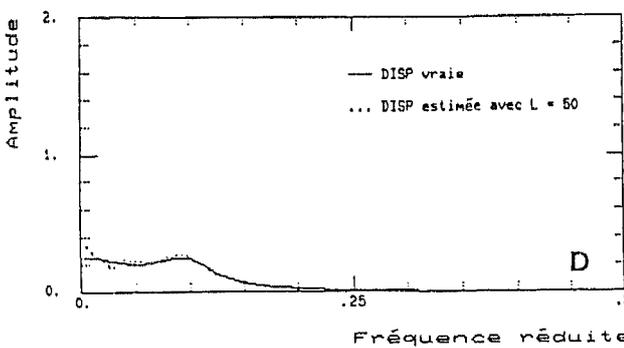
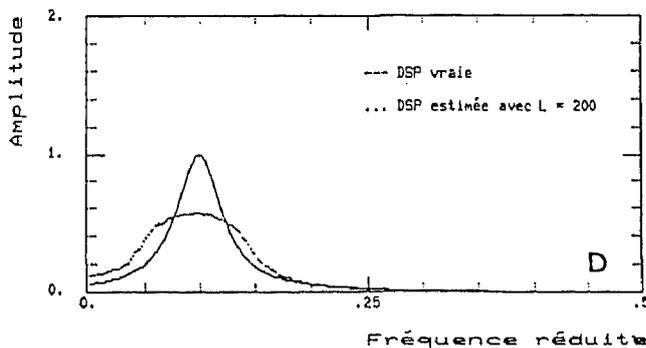
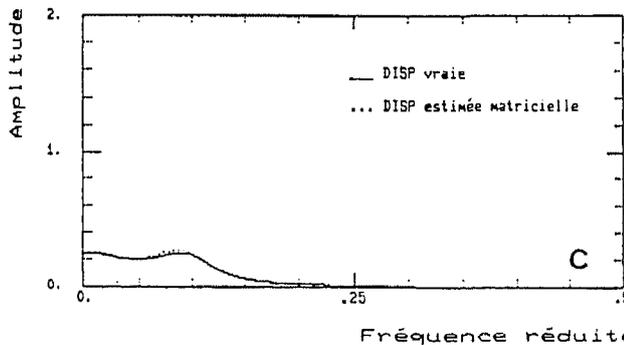
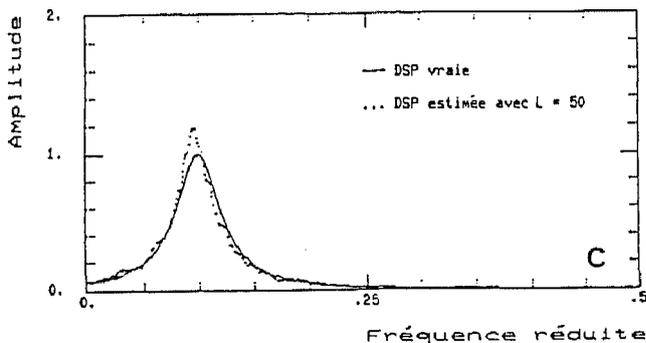


Figure 4.— Comparaison de notre méthode à la méthode classique dans le cas des autospectres.

Figure 5.— Comparaison de notre méthode à la méthode classique dans le cas des interspectres.

**ANNEXE A : ALGORITHME DE RECHERCHE DES INTERVALLES SIGNIFICATIFS**

Comme nous l'avons dit au sous-paragraphe 2.1.5. il nous faut, dans la base des vecteurs  $\{v(i)\}$ , déterminer les intervalles significatifs pour les composantes  $\{a_i\}$ . C'est-à-dire ceux où  $|a(i)|^2$  est supérieur à  $\sigma^2$ .

La décision se faisant sur les  $\{\alpha_i\}$ , affectés d'un bruit  $\{\beta_i\}$  de variance  $\sigma^2$  d'après (6), nous considérons le critère somme  $\hat{S}(p)$ , analogue à (13), défini à partir d'une origine  $p_0$  qui n'est plus 1 nécessairement :

$$(A.1) \quad \hat{S}(p) = \sum_{i=p_0}^{p-1} [|\alpha(i)|^2 - 2\sigma^2].$$

La recherche des intervalles significatifs se ramène à la recherche des intervalles où  $\hat{S}(p)$  est en moyenne croissant. A chaque intervalle croissant correspondra en principe un « mode » de la fenêtre de pondération. Cependant nous allons passer par deux étapes intermédiaires destinées à réduire l'effet des fluctuations statistiques de  $\hat{S}(p)$ .

**A.1. RECHERCHE DES INTERVALLES CROISSANTS « EN MOYENNE »**

Du fait notamment des fluctuations statistiques,  $\hat{S}(p)$  peut présenter des extrema secondaires à l'ensemble des minima et maxima que nous recherchons, comme le montre la figure A.1. L'effet en serait de morceler la fenêtre de pondération de façon préjudiciable.

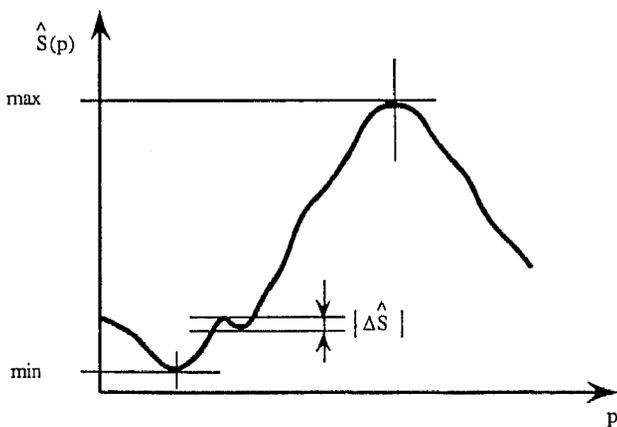


Figure A.1. — Intervalle croissant « en moyenne » de  $\hat{S}(p)$ .

Nous devons donc introduire un test permettant d'éliminer des extrema ayant une forte probabilité de correspondre soit à des valeurs isolées  $\{|a_i|^2 < \sigma^2\}$ , soit à l'effet du bruit additif.

Le test choisi est le suivant :

$$(A.2) \quad \begin{matrix} H_0 \\ |\Delta \hat{S}| \geq 2\sigma^2 \\ H_1 \end{matrix}$$

Sous l'hypothèse  $H_1$  on décide qu'il y a du signal et les extrema sont éliminés. Sous l'hypothèse  $H_0$  on décide qu'il n'y a pas de signal et on retient les extrema comme limites d'intervalle.

Si  $i_1$  et  $i_2$  désignent les indices du maximum et du minimum, nous vérifions qu'à partir de l'expression (A.1) :

$$(A.3) \quad \Delta \hat{S} = \hat{S}(i_2) - \hat{S}(i_1) = \sum_{i=i_1}^{i_2} [|\alpha_i + \beta_i|^2 - 2\sigma^2]$$

- les extrema isolés ( $i_2 = i_1 + 1$ ) sont automatiquement éliminés ( $|\Delta \hat{S}| < 2\sigma^2$  dans ce cas),
- la probabilité de  $H_0$  est très faible dès que  $|a_i| > \sigma$ ,
- pour  $a_i = 0$ ,  $\Delta \hat{S}$  est une variable en chi-deux à  $2I = 2(i_2 - i_1)$  degrés de liberté, de valeur moyenne  $-I\sigma^2$  et d'écart type  $\sqrt{I}\sigma^2$ . Dès que  $I = 4$  la probabilité de  $H_0$  est notable et elle devient très voisine de 1 si  $I \geq 9$ .

Nous pouvons donc considérer que le test satisfait les conditions que l'on s'était fixées.

**A.2. SÉLECTION DES INTERVALLES SIGNIFICATIFS**

Nous avons dans une première étape éliminé les parties décroissantes accidentelles de  $\hat{S}(p)$ . Parmi les intervalles croissants « en moyenne » ainsi retenus, certains peuvent être liés non à du signal mais à des valeurs anormalement grandes du bruit. D'où la nécessité d'un deuxième test pour sélectionner les intervalles réellement « significatifs ».

Le test retenu (voir fig. A.2) est le suivant :

$$(A.4) \quad \begin{matrix} H_1 \\ \Delta \hat{S} > 9\sigma^2 \\ H_0 \end{matrix}$$

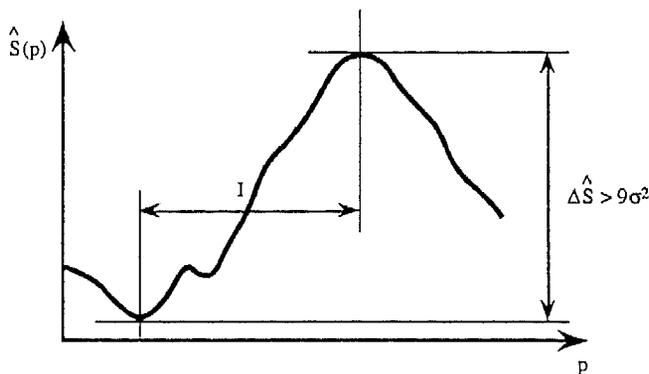


Figure A.2. — Sélection d'un intervalle significatif.

Sous  $H_1$  on décide qu'il y a du signal : l'intervalle est retenu. Sous  $H_0$  l'accroissement est considéré comme non significatif.

En reprenant la définition (A.3) de  $\Delta \hat{S}$  nous vérifions les points suivants :

— en l'absence de signal nous avons vu que  $\Delta \hat{S}$  a une loi en khi-deux d'espérance  $-I\sigma^2$  et d'écart type  $\sqrt{I\sigma^2}$  : la probabilité de décider  $H_1$  est donc très faible,

— en présence de signal  $E[\Delta \hat{S}] = I[\overline{|a_i|^2} - \sigma^2]$ , le seuil de détection moyen correspondant est donc donné par :

$$(A.5) \quad \overline{|a_i|^2} = \sigma^2 \left( 1 + \frac{9}{I} \right).$$

Ce seuil tend rapidement vers  $\sigma^2$  en présence de composantes adjacentes de signal.

Pour une composante isolée, le seuil est au pire à  $|a_i| = 3\sigma$ .

On peut donc considérer que le test joue bien le rôle demandé.

A.3. TRACE DE LA FENÊTRE OPTIMALE

Les intervalles significatifs étant déterminés, il reste à construire les fenêtres de pondération correspondantes. Pour les raisons déjà évoquées nous gardons une fenêtre en cosinus ajusté.

Conformément à la discussion sur la formule (19), nous plaçons les points à mi-hauteur aux extrémités des intervalles significatifs, qui correspondent en moyenne à  $|a_i|^2/\sigma^2 = 1$ . Par analogie avec la discussion du paragraphe 2.1.4, les points sont mis à 1 dans la partie centrale de la fenêtre où en moyenne  $|a_i|^2/\sigma^2 \geq 9$ . Cette partie est détectée par un critère analogue à  $Q''(p)$ , le critère utilisé est défini par :

$$(A.6) \quad \hat{S}_9(p) = \sum_{i=p_0}^{p-1} [|\alpha(i)|^2 - 10\sigma^2].$$

La partie où en moyenne  $|a_i|^2/\sigma^2 \geq 9$  est la partie croissante de  $\hat{S}_9(p)$ , située entre le minimum et le maximum absolu de  $\hat{S}_9(p)$  dans l'intervalle significatif correspondant (voir fig. A.3). Dans l'éventualité

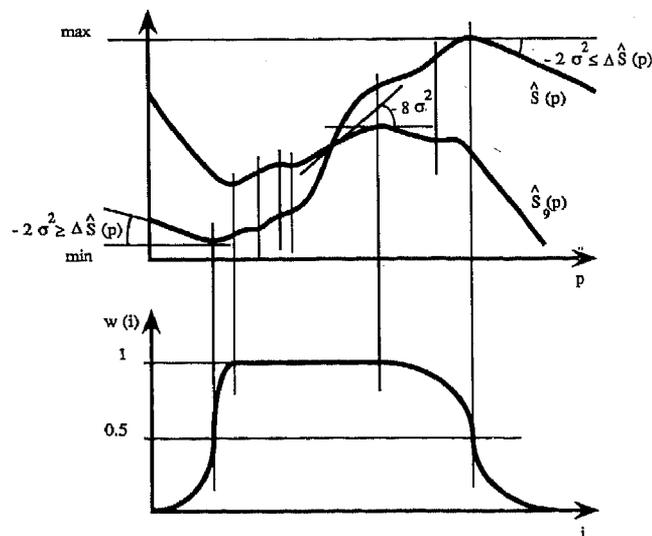


Figure A.3. — Fenêtre optimale.

où il n'y aurait pas d'extréma (si  $|\alpha(i)|$  reste inférieur à  $3\sigma$  dans l'intervalle) nous utiliserons une fenêtre en cosinus surélevé centrée.

ANNEXE B : TRANSFORMATION LOGARITHMIQUE DU PÉRIODOGRAMME PRÉMOYENNE

A partir du périodogramme le périodogramme pré-moyenné est défini par :

$$(B.1) \quad \hat{\pi}_P(m) = \frac{1}{NP} \sum_{p=m-P/2}^{m+P/2} |\hat{x}(m-p)|^2$$

qui se met sous la forme :

$$(B.2) \quad \hat{\pi}_P(m) = \tilde{\gamma}_P(m) \hat{\chi}_{2P}^2(m)$$

où  $\hat{\chi}_{2P}^2(m)$  est une variable en khi-deux d'ordre  $2P$  et où  $\tilde{\gamma}_P(m)$  reste voisin de  $\hat{\gamma}(m)$  si  $P$  est petit.

Le Log-périodogramme s'écrit alors :

$$(B.3) \quad \text{Log } \hat{\pi}_P(m) = \text{Log } \tilde{\gamma}_P(m) + \text{Log } \hat{\chi}_{2P}^2(m)$$

avec

$$(B.4) \quad \begin{cases} E[\text{Log } \hat{\chi}_{2P}^2(m)] = \eta_P, \\ E[(\text{Log } \hat{\chi}_{2P}^2(m) - \eta_P)^2] = \frac{1}{P} (1 + \delta_P), \end{cases}$$

le coefficient  $\delta_P$  représente l'augmentation de variance liée à l'utilisation du Log-périodogramme, il tend rapidement vers zéro (ainsi que  $\eta_P$ ) lorsque  $P$  augmente ( $P=1$ ,  $\delta_P=0,645$  et  $\eta_P=-0,577$ ;  $P=3$ ,  $\delta_P=0,185$  et  $\eta_P=-0,176$ ).

Ainsi, pour obtenir le modèle du paragraphe précédent il suffit de poser :

$$(B.5) \quad \begin{cases} \hat{u}(m) = \text{Log } \hat{\pi}_P(m) - \eta_P \\ \quad = \hat{s}(m) + \hat{b}(m), \\ \hat{s}(m) = \text{Log } \tilde{\gamma}_P(m), \\ \hat{b}(m) = \text{Log } \hat{\chi}_{2P}^2(m) - \eta_P \\ \quad \text{avec } \sigma_b^2 = \frac{1}{P} (P > 3). \end{cases}$$

ANNEXE C : ESTIMATION DU MOMENT D'ORDRE DEUX DE LA COHÉRENCE ESTIMÉE

D'après les relations (33) et (34) l'élément  $\hat{c}_{ij}(m)$  de la matrice de cohérence estimée est définie par :

$$(C.1) \quad \hat{c}_{ij}(m) = \frac{1}{N} \hat{x}_i(m) \hat{x}_j^*(m) / (\hat{\gamma}_{ii}(m) \hat{\gamma}_{jj}(m))^{1/2},$$

dont nous nous proposons de mesurer la covariance.

En variables réduites, en supprimant les indices pour alléger les notations, le problème peut être reformulé de la façon suivante. Introduisons les variables normalisées

$$(C.2) \quad \begin{cases} \hat{u} = \hat{x}_i(m) / \sqrt{N\hat{\gamma}_i(m)}, \\ \hat{v} = \hat{x}_j(m) / \sqrt{N\hat{\gamma}_j(m)}. \end{cases}$$

On remarque que

$$(C.3) \quad \hat{c} \triangleq \hat{c}_{ij}(m) = \hat{u}\hat{v}^* = \hat{c} + \hat{\varepsilon}(m)$$

où nous avons posé :

$$(C.4) \quad \hat{c} \triangleq E[\hat{c}_{ij}(m)] = E[\hat{u}\hat{v}^*].$$

Les propriétés du périodogramme entraînent que  $\hat{u}$  et  $\hat{v}$  sont deux variables gaussiennes circulaires vérifiant

$$(C.5) \quad E[\hat{u}^2] = E[\hat{v}^2] = E[\hat{u}\hat{v}] = 0$$

et d'après la normalisation (C.2) :

$$(C.6) \quad E[|\hat{u}|^2] = E[|\hat{v}|^2] = 1.$$

En utilisant les propriétés des variables gaussiennes complexes il vient

$$(C.7) \quad \begin{aligned} E[|\hat{c}|^2] &= E[\hat{u}\hat{v}^* \hat{u}^* \hat{v}] \\ &= |E[\hat{u}\hat{v}^*]|^2 + E[|\hat{u}|^2] E[|\hat{v}|^2] \\ &= |\hat{c}|^2 + 1. \end{aligned}$$

D'où la covariance

$$(C.8) \quad E[|\hat{c}|^2] - |E[\hat{c}]|^2 = E[|\hat{\varepsilon}(m)|^2] = 1.$$

Un calcul plus complet est donné dans [3] pour le cas où dans (C.1), nous utilisons non pas les valeurs exactes de  $\hat{\gamma}_i(m)$ , mais des valeurs estimées par moyennage sur  $L_e$  composants du périodogramme. Nous obtenons alors, au premier ordre en  $\frac{1}{L_e}$ , la relation :

$$(C.9) \quad E[|\hat{\varepsilon}(m)|^2] \cong 1 + \frac{1}{4L_e} (1 + |\hat{c}|^2).$$

Pour les valeurs usuelles de  $L_e$ , nécessairement grandes devant un si nous voulons une bonne précision sur le spectre, le terme correctif pourra être négligé.

### BIBLIOGRAPHIE

- [1] A. PAPOULIS, Signal analysis, McGraw-Hill, Auckland, 1984, p. 378-380.
- [2] B. LUMEAU, Identification de sources d'activités cérébrales, Thèse de docteur-ingénieur, Orsay, 1980.
- [3] B. LUMEAU et H. CLERGEOT, Spatial Localization. Spectral matrix bias and variance. Effects on the source subspace, *Signal Processing*, Vol. 4, 1982, pp. 103-123.
- [4] N. L. OWSLEY, Spectral signal set extraction, *Aspects of Signal Processing Part 2*, G. Tacconi Ed., D. Reidel Publishing Company, Dordrecht, 1976, p. 469-475.
- [5] G. WAHBA, Automatic smoothing of the Log-periodogram, *Journal of the American Statistical Association*, Vol. 75, 1980, p. 122-131.
- [6] H. AKAIKE, A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, Vol. 19, 1974, p. 716-723.
- [7] H. L. VAN TREES, Detection, estimation and modulation theory, part I, John Wiley and Sons, New York, p. 231-232.
- [8] H. CLERGEOT, Estimation du spectre du signal aléatoire gaussien par le critère du maximum de vraisemblance ou du maximum de probabilité a posteriori, Thèse de docteur ès sciences physiques, Orsay, 1982.
- [9] A. PAPOULIS, Signal analysis, McGraw-Hill, Auckland, 1984, p. 302-303 et p. 374-377.
- [10] H. L. VAN TREES, Detection, estimation and modulation theory, part I, John Wiley & Sons, New York, 1968, p. 258 et 380.
- [11] B. LUMEAU et H. CLERGEOT, Dispositif d'analyse spectrale mono et multidimensionnelle optimisant le compromis biais-variance, Dépôt de brevet français N° 8707600, 29 mai 1987. Dépôt de brevet européen N° 88 401310.3, 30 mai 1988. Publication N° EP 0298 786 A1, 11 janvier 1989.
- [12] B. LUMEAU, Traitement spatial et analyse spectrale. Applications neurophysiologiques, Thèse de docteur ès sciences physiques, Orsay, 1987.