

Estimateurs du maximum de vraisemblance étendus à la séparation de sources non gaussiennes

*Maximum likelihood estimators
applied to the non gaussian source separation*



Michel GAETA

CEPHAG INPG/IEG URA 346 B.P. n° 46,
38402 Saint-Martin-d'Hère, France

1987 : DEA des dispositifs de l'électronique intégrée, option « circuit » de l'école centrale de Lyon, INSA Lyon, Université Claude Bernard, et l'Université de Saint Etienne. 1987 : Diplôme d'Ingénieur physicien électronicien de l'Institut de Chimie et Physique de Lyon, Section électronique et traitement de l'information. 1988-1990 : Thèse de doctorat nouveau régime de l'INPG en Automatique et traitement du signal : « multispectres et multicorrélation ».



J. L. LACOUME

CEPHAG INPG/IEG URA 346 B.P. n° 46,
38402 Saint-Martin-d'Hère, France

Ancien élève de l'École Normale Supérieure, Agrégé de Physique (1964), Docteur ès Sciences (1969), Professeur à l'Institut National Polytechnique de Grenoble, Domaine de recherche : géophysique : étude de l'environnement ionisé de la terre par sondage électromagnétique ; traitement du signal : analyse spectrale et interspectrale, statistiques d'ordre supérieur, applications en instrumentation, acoustique sous-marine, géophysique.

RÉSUMÉ

Une combinaison de signaux émis par plusieurs sources est reçue sur un réseau de capteurs. Dans cette situation qui se rencontre dans de nombreux domaines d'application, on cherche à séparer les sources en construisant une transformation des signaux reçus dont chaque composante est liée de manière exclusive à chacune des sources. Cette séparation des sources ne peut être réalisée sans hypothèse supplémentaire en se limitant à l'utilisation de la matrice spectrale. Nous montrons

ici que cette séparation peut être menée à bien en utilisant les statistiques d'ordre supérieur et nous établissons la forme de l'estimateur du maximum de vraisemblance du modèle de sources conduisant à la séparation.

MOTS CLÉS

Maximum de vraisemblance, non gaussien, indépendance, séparation.

ABSTRACT

A combination of signals emitted by several sources is received on an array of sensors. In this situation frequently realized in application, we want to separate the sources in obtaining a transformation of the observed signals giving components related for each of them uniquely to a specific source. This separation cannot be successful using the spectral matrix and then people use a priori information. We show here that this separation can be

done using Higher Order Statistics of the observed signals and we give the maximum likelihood estimator of the source models that defines the separation.

KEY WORDS

Maximum likelihood, non gaussian, independence, separation.

Avertissement au lecteur : convention d'écriture

Pour effectuer les calculs, nous avons utilisé la convention de sommation décrite dans ce paragraphe. Toutefois, pour faciliter la lecture et l'interprétation, chaque fois que cela était possible, certaines expressions ont été données en notation matricielle.

Ainsi on aura :

$$\underline{r}^T \underline{r} = \sum_i r_i r_i = r_i r_i \quad \text{sommation sur } i$$

$$\underline{M} \underline{r} = \sum_j M_{ij} r_j = M_{ij} r_j \quad \text{sommation sur } j$$

$$\underline{r}^T \underline{M} \underline{r} = \sum_{ij} M_{ji} r_j r_i \quad \text{sommation sur } i \text{ et } j.$$

On note c_{ijkl} un tenseur d'ordre 4 qui peut par contraction donner [15] :

— Une matrice

$$\sum_{kl} C_{ijkl} r_k r_l \quad \text{sommation sur } k \text{ et } l.$$

— Un scalaire

$$\sum_{ijkl} C_{ijkl} r_i r_j r_k r_l \quad \text{sommation sur } i, j, k \text{ et } l.$$

1. Introduction

Les méthodes, maintenant classiques, de caractérisation et de filtrage linéaire des signaux sont basées sur l'utilisation de critères quadratiques liés à l'énergie ou à l'énergie d'interaction. Pour la caractérisation, les propriétés d'un signal à une dimension sont données par la Densité Spectrale de Puissance (DSP). Pour un signal à plusieurs composantes on utilise la matrice interspectrale [1]. L'utilisation de critères quadratiques est adaptée aux signaux gaussiens.

Le problème essentiel que nous traiterons est celui de la séparation de sources. Avant de l'aborder, définissons cette opération et montrons son importance pratique.

Le problème de séparation de sources est issu de l'astronomie, de l'astrophysique et de l'acoustique sous-marine. Dans ces divers domaines, on s'intéresse à un objet (étoile, bâtiment...) qui émet des signaux (optiques, électromagnétiques, acoustiques). Ces signaux sont transmis dans un milieu de propagation et reçus sur un réseau de capteurs. A partir des signaux reçus (ou du signal multidimensionnel reçu), on veut caractériser les émetteurs que l'on appelle *sources*. Ce problème est très général et l'on fait un certain nombre d'hypothèses que nous adopterons par la suite.

Le milieu de propagation est linéaire et déterministe : les signaux reçus sont donc une combinaison linéaire des signaux émis par les sources. La combinaison la plus générale est une convolution en temps ou un produit

matriciel en fréquence. L'approche que nous présentons est adaptée à la séparation de sources combinées par un produit matriciel. Cependant, pour nous placer en fréquence, nous devrions considérer des signaux à valeurs complexes. Nous nous limiterons ici à des signaux à valeurs réelles (nous pensons pouvoir lever cette limitation dans une future étude).

L'étude limitée à des signaux à valeurs réelles permet de traiter le problème de la séparation de sources temporelles combinées linéairement *sans retard*. Le signal multidimensionnel reçu sur les capteurs s'écrit alors :

$$(1) \quad \underline{R}_{(n)}(k) = \sum_{i=1}^p A_i(k) \underline{s}_i + \underline{B}(k)$$

avec

n : nombre de capteurs

p : nombre de sources

$\underline{B}(k)$: est un vecteur parasite ou bruit, à n composantes, $\underline{B}(k)$ est un signal aléatoire.

A_i : représente l'amplitude de la source i . $A_i(k)$ est un signal aléatoire.

\underline{s}_i : est un vecteur à N composantes, la composante s_{ij} de \underline{s}_i décrit la propagation entre la source i et le capteur j . Le milieu de propagation étant déterministe \underline{s}_i est un vecteur certain.

Revenons sur la notion de source et sur la séparation de sources. La formule (1) nous montre que les signaux reçus contiennent un mélange des sources (A_i). Séparer les sources consiste, à partir des signaux reçus, à retrouver les sources A_i . Il apparaît ici une limitation insurmontable qui va nous obliger à modifier notre définition de la séparation de sources. Le terme $A_i \underline{s}_i$ sépare d'une manière arbitraire le milieu de propagation \underline{s}_i et la source A_i . On pourrait écrire de manière équivalente $(A_i \alpha)(\underline{s}_i/\alpha)$. Nous voyons donc que nous pouvons seulement espérer obtenir un signal *proportionnel* aux sources. De même, l'indice i associé à une certaine source est également arbitraire, et ne sous-entend pas d'ordre entre les sources. Celles-ci seront donc connues à *une permutation* près. Notons cependant que la technique de séparation mise en œuvre permet de retrouver le signal émis par chacune des sources et reçu par l'un quelconque des capteurs du réseau.

Finalement, nous appellerons « séparation de sources », l'opération qui, à partir des signaux reçus, nous permet d'obtenir un *ensemble* de signaux *proportionnels* aux signaux émis par chacune des sources.

Pour pouvoir espérer séparer les sources, nous devons savoir les distinguer par un caractère spécifique. Dans les études menées sur ce sujet (et nous adopterons ce point de vue), on distingue les sources en supposant que les signaux aléatoires qu'elles émettent, sont *décorrélés*. Comme l'on suppose que les sources émettent des signaux centrés, la décorrélation entraîne que l'intercorrélolation de deux sources est nulle :

$$E[A_i A_j] = \mathbb{P}_i \delta_{ij} \quad \text{avec } \delta_{ij} = 0 \text{ si } i \neq j$$

$\mathbb{P}_i = E[|A_i|^2]$ est la puissance de la source i .

Cette hypothèse est vraisemblable dans les exemples cités plus haut, où l'on peut admettre que deux étoiles différentes émettent des signaux optiques ou électromagnétiques décorrélés ou que les deux machines de deux bâtiments ne sont pas synchronisées...

On connaît des situations qui mettent cette hypothèse en défaut : c'est le cas des trajets multiples où le signal émis par une source atteint les capteurs par deux trajets différents. Ces situations ont donné lieu à des études spécifiques conduisant, par exemple, aux techniques de moyennage spatial [20] ou fréquentiel [19]. Nous admettrons ici que l'hypothèse de décorrélation des sources est vérifiée.

La condition de décorrélation exprimée sur les signaux fréquentiels est la base des méthodes de séparation introduites par H. Mermoz, utilisant la matrice spectrale [18]. Elles ont conduit à la floraison de multiples algorithmes de traitement d'antennes [21]. Cette méthode introduite « heuristiquement » par H. Mermoz, a été justifiée par le principe du maximum de vraisemblance [6] [12]. Cependant, comme l'avait montré H. Mermoz et comme le justifie l'estimateur du maximum de vraisemblance, *la matrice spectrale ne permet pas de séparer les sources* et l'on doit faire appel à *des hypothèses supplémentaires*. Heureusement, pour ces méthodes, une hypothèse supplémentaire suffisante et vraisemblable existe : c'est l'*hypothèse d'onde plane*. Si l'on suppose le milieu de propagation homogène (au voisinage des capteurs) et les sources à l'infini (à grande distance) et si, pour faciliter la résolution, on admet que l'antenne est rectiligne et à capteurs équidistants, alors on peut séparer les sources et les localiser (angulairement). Il ne faut pas être trop effrayé par cet ensemble d'hypothèses nouvelles qui sont généralement vérifiées. Notre étude va montrer comment on peut séparer les sources, sans hypothèse supplémentaire. L'idée de base consiste à élargir la notion de décorrélation des sources en admettant que les signaux émis par les sources sont des

Signaux Aléatoires Statistiquement Indépendants.

Cette indépendance statistique des sources est souvent invoquée par les auteurs présentant des méthodes basées sur la matrice spectrale mais comme ils se limitent aux moments d'ordre 2, ils n'exploitent que (très) partiellement l'indépendance statistique. A la décharge des auteurs cités, il faut noter que si les signaux émis par les sources sont décorrélés et gaussiens (donc indépendants), alors l'étude faisant appel à la matrice spectrale est complète et donc, dans le cas de *signaux gaussiens*, *on ne peut séparer les sources sans hypothèse(s) supplémentaire(s)*.

Nous supposons donc que les signaux émis par les sources sont statistiquement indépendants et qu'un signal source au plus est gaussien. Nous montrerons alors que l'on peut séparer les sources en utilisant des moments d'ordre supérieur (quatre en pratique). Nous présenterons une justification théorique de ce résultat (c'est là l'apport original de cette étude) basée sur le principe du maximum de vraisemblance. Avant d'introduire notre étude, reve-

nons sur le caractère gaussien (ou non) des signaux et sur la littérature déjà existante sur ce sujet.

Malgré le théorème de la limite centrale, la nature ne produit que très rarement des signaux gaussiens et l'on peut penser que les signaux non gaussiens sont plutôt la règle commune. Peut-on dire que les signaux gaussiens sont fréquemment invoqués par suite, essentiellement, des facilités de calcul qu'ils introduisent ?

Ainsi une machine tournante émet des fréquences pures ou quasi pures, un émetteur envoie des signaux modulés en FM, PSK etc... Tous ces signaux sont non gaussiens et même déterministes. Nous renvoyons les lecteurs à [1] où ils verront comment on peut utiliser la matrice spectrale avec des signaux certains.

Même les signaux dits aléatoires sont rarement gaussiens et l'on a étudié (en acoustique sous-marine) des modèles dits de « Mixture » pour mieux approcher la réalité [2], [3].

Nous nous trouvons alors devant un nouveau problème : comment définir et caractériser les signaux non gaussiens ?

Si nous attribuons aux signaux une loi bien définie, nous traiterons un problème très particulier. Nous cherchons à préciser la description statistique des signaux tout en préservant le caractère général de notre étude.

Les propriétés statistiques des signaux sont décrites soit par leur loi de probabilité soit de manière équivalente par la densité de probabilité d'amplitude, la première ou la seconde fonction caractéristique [4]. C'est à travers la seconde fonction caractéristique que nous trouverons la solution à notre problème. La loi gaussienne conduit à une fonction caractéristique dont le développement est limité à l'ordre 2. Le développement de toute autre loi peut être limité à l'ordre 2 ce qui revient à l'approcher par une loi gaussienne. Nous proposons de préciser la caractérisation des propriétés statistiques en poussant le développement à l'ordre 4. Cette technique d'approximation des densités de probabilités est bien connue : c'est le développement de Gram-Charlier (annexe 3).

Pour décrire le caractère non gaussien des signaux nous utiliserons donc un développement de Gram-Charlier de leur loi de probabilité, qui introduit de manière naturelle les cumulants d'ordre 4. Nous montrons alors que le logarithme de la fonction de vraisemblance se décompose en la somme de deux termes. Le premier est identique au cas gaussien : sa maximisation permet d'obtenir le sous-espace source, mais les sources restent indéterminées. Nous montrons que l'indétermination sur les sources peut être ramenée à une matrice unitaire décrite par une matrice $\underline{\underline{U}}$. La maximisation de ce second terme donne une forme implicite de l'estimateur de la matrice $\underline{\underline{U}}$ et donc des sources et des cumulants des sources.

D'autres approches de l'identification de sources ont été développées. Elles abordent le problème de manière :

- Heuristique : [14]
- Analytique : [16]
- Tensorielle : [17].

2. Le modèle

Notre étude concerne la séparation d'un mélange linéaire de sources non gaussiennes, polluées par un bruit gaussien additif.

Les hypothèses introduites pour décrire un tel modèle sont :

- les signaux émis par les sources sont non gaussiens, indépendants jusqu'à l'ordre quatre ;
- les signaux sont additionnés à un bruit blanc gaussien, isotrope indépendant des signaux.

Les sources traversent le milieu par différents trajets et subissent un filtrage linéaire dépendant de la source et du capteur. Des capteurs enregistrent donc une combinaison linéaire des signaux émis par les sources.

Nous supposons, dans notre étude, l'ensemble des signaux et vecteurs à valeurs réelles. Une telle hypothèse nous place dans le cas d'un filtrage sans retard.

Le nombre p de sources indépendantes est supposé connu. On dispose de n capteurs.

En se plaçant à un k fixé et p sources indépendantes, le signal observé (1) s'écrit :

$$(2) \quad \mathbf{R}_{(n)} = \sum_{i=1}^p \mathbf{A}_i \underline{s}_i + \mathbf{B}$$

avec

$\mathbf{R}_{(n)}$: vecteur des observations de dimension n (n nombre de capteurs).

\mathbf{A}_i : amplitude aléatoire de la source i : cette variable aléatoire a pour densité de probabilité : $P_{A_i}(a_i)$. Soit \mathbf{A} le vecteur colonne des amplitudes aléatoires A_i , et \mathbf{P} la matrice définie par :

$$\mathbf{P} = E[\mathbf{A}\mathbf{A}^T]$$

\underline{s}_i : « vecteur source » décrivant la loi de propagation entre les capteurs, pour la source i . Ces vecteurs sont à composantes déterministes. La matrice $\underline{\mathbf{S}}_{(n)}$ est constituée par les vecteurs colonnes \underline{s}_i .

\mathbf{B} : vecteur bruit : vecteur aléatoire gaussien centré à composantes indépendantes, et isotrope avec : $E(\mathbf{B}) = 0$; $E(\mathbf{B}\mathbf{B}^T) = P_B \mathbf{I}$ où P_B est la puissance de bruit.

La matrice $\underline{\mathbf{S}}_{(n)}$ est définie comme la matrice obtenue en rangeant en colonnes les vecteurs sources. Une écriture matricielle du modèle (2) est : $\mathbf{R}_{(n)} = \underline{\mathbf{S}}_{(n)}\mathbf{A} + \mathbf{B}$.

Chaque composante du signal observé $\mathbf{R}_{(n)}$ est une combinaison des signaux émis par les sources et du bruit. Nous cherchons une transformation de $\mathbf{R}_{(n)}$ permettant d'obtenir p combinaisons des composantes de $\mathbf{R}_{(n)}$ dépendant chacune uniquement d'une source et du bruit. Nous aurons ainsi séparé les sources.

2.1. LIMITES DE L'ANALYSE

Nous avons montré qu'il est vain de rechercher l'amplitude, ainsi que « l'ordre » des sources.

On supposera donc les amplitudes des sources normées : $\mathbf{P} = \mathbf{I}$.

Cette hypothèse, nous permet d'avoir une égalité entre les cumulants et les kurtosis des sources [4], [22].

Les signaux des sources sont par ailleurs, supposés centrés : $E[A_i] = 0$.

Les vecteurs sources auront donc pour amplitude le gain de la transformation linéaire subie par les sources normées.

2.2. PARAMÉTRISATION DE LA TRANSFORMATION INCONNUE

Une étude au deuxième ordre, en utilisant la matrice spectrale, permet de déterminer l'espace à p dimensions contenant les sources. En projetant $\mathbf{R}_{(n)}$ sur cet espace, on obtient : (soit \mathbf{R} cette projection)

$$\mathbf{R} = \underline{\mathbf{S}}\mathbf{A} + \mathbf{B}$$

où $\underline{\mathbf{S}}$ est une matrice $p \times p$ non singulière car, par hypothèse, les sources sont supposées linéairement indépendantes. Le problème de la séparation de sources se ramène à la détermination de la matrice $\underline{\mathbf{S}}$.

La matrice $\underline{\mathbf{S}}$ des vecteurs sources, possède une « décomposition polaire à droite unique » [11], de la forme :

$$\underline{\mathbf{S}} = \underline{\mathbf{H}}\underline{\mathbf{U}}$$

où \mathbf{H} est une matrice hermitienne ; dans notre cas \mathbf{H} est une matrice symétrique ; \mathbf{U} est une matrice unitaire.

Remarque : Cette décomposition nous permet de mettre en évidence le caractère partiel d'une étude à l'ordre deux ; en effet, si l'on forme le terme quadratique :

$$\underline{\mathbf{S}}\underline{\mathbf{S}}^T = \underline{\mathbf{H}}\underline{\mathbf{U}}\underline{\mathbf{U}}^T\mathbf{H}^T = \underline{\mathbf{H}}\mathbf{H}^T$$

il apparaît que la matrice unitaire $\underline{\mathbf{U}}$ n'est pas mesurable par un tel terme.

$\underline{\mathbf{H}}$ admet une « décomposition spectrale » :

$$\underline{\mathbf{H}} = \underline{\mathbf{V}}\underline{\mathbf{\Delta}}^{1/2}\underline{\mathbf{V}}^T$$

où $\underline{\mathbf{V}}$ est une matrice unitaire. Cette matrice est formée des vecteurs propres de la matrice $\underline{\mathbf{H}}$.

$\underline{\mathbf{\Delta}}$ est une matrice diagonale réelle :

$$\underline{\mathbf{\Delta}} = \begin{pmatrix} \lambda_{11} & 0 & 0 \\ 0 & \backslash & 0 \\ 0 & 0 & \lambda_{pp} \end{pmatrix}$$

Cette matrice est formée des valeurs propres de la matrice $\underline{\mathbf{H}}$.

La matrice $\underline{\mathbf{S}}$ prend alors la forme :

$$\underline{\mathbf{S}} = \underline{\mathbf{V}}\underline{\mathbf{\Delta}}^{1/2}\underline{\mathbf{V}}^T\underline{\mathbf{U}}$$

On note

$$\underline{\mathbf{\Pi}} = \underline{\mathbf{V}}^T\underline{\mathbf{U}}$$

Cette matrice $\underline{\Pi}$ est une matrice unitaire, d'où :

$$\underline{S} = \underline{V} \underline{\Delta}^{1/2} \underline{\Pi}.$$

Le problème de la détermination des sources se ramène à la détermination de :

- la matrice diagonale $\underline{\Delta}$;
- des deux matrices unitaires \underline{V} et $\underline{\Pi}$.

Cette paramétrisation permettra de mener à bien le calcul de la fonction de vraisemblance.

2.3. LES DEGRÉS DE LIBERTÉ DE LA PARAMÉTRISATION

Il nous faut identifier les p^2 éléments inconnus de la matrice \underline{S} .

La matrice $\underline{\Delta}$ possède p degrés de liberté car elle est diagonale.

La matrice \underline{V} est une matrice unitaire de dimension $(p \times p)$. Ses vecteurs colonnes sont normés, ce qui impose p relations. Ils sont également orthogonaux deux à deux, ils vérifient $p(p-1)/2$ relations.

La matrice \underline{V} apporte $p(p-1)/2$ degrés de liberté.

Les transformations Π et U , unitaires, possèdent également $p(p-1)/2$ degrés de liberté.

L'estimation des transformations Π ou U , nous permettra de déterminer \underline{S} , sans hypothèse supplémentaire.

Nous allons montrer qu'une étude à l'ordre quatre permet de déterminer, dans la plupart des cas pratiques, la matrice $\underline{\Pi}$.

3. Fonction de vraisemblance pour une observation

Nous avons mené une démarche identique à [6], [12], afin de calculer les caractéristiques du sous-espace source au sens du maximum de vraisemblance. Nous présentons ici les principales étapes de calcul de ces estimateurs (le calcul est développé en annexe 1).

3.1. DESCRIPTION STATISTIQUE DES SOURCES

La fonction de vraisemblance représente la probabilité de l'observation conditionnellement aux paramètres du système.

Rappelons que ces paramètres sont :

- p : dimension de l'espace source (supposée connue).
- P_B : puissance du bruit.
- $\underline{\Pi}, \underline{\Delta}, \underline{V}$: décrivent la matrice source \underline{S} .
- cum : les cumulants d'ordre 4 des sources.

On note : $\{p, P_B, \underline{\Pi}, \underline{\Delta}, \underline{V}, \text{cum}\} = \varepsilon$; ε représente l'ensemble des paramètres inconnus.

La fonction de vraisemblance associée à une observation s'écrit :

$$P_{R/p, P_B, \Pi, \Delta, V, \text{cum}} = P_{R/\varepsilon}.$$

On dispose de M réalisations statistiquement indépendantes. La fonction de vraisemblance globale sera donnée par

le produit des fonctions de vraisemblance (identiques) associées à une observation. Les estimateurs du maximum de vraisemblance des paramètres inconnus maximisent cette fonction.

3.2. FORME GÉNÉRALE DE LA FONCTION DE VRAISEMBLANCE

Exprimons $P_{R/\varepsilon}$ conditionnellement à A (amplitude des sources). Le modèle choisi (1) induit la séparation de la fonction de vraisemblance, en deux termes :

- un premier relatif aux signaux ;
- un second relatif au bruit obtenu en retranchant les signaux à l'observation.

$$(3) \quad P_{R/\varepsilon} = \int P_{A/\varepsilon}(a) P_{\text{bruit}}(\underline{r} - \underline{S}a) da.$$

On définit :

$$(4) \quad \beta(u, r) = \text{TF}_a(P_{\text{bruit}}(\underline{r} - \underline{S}a))$$

(TF : Transformée de Fourier).

Le calcul de $\beta(u, r)$ est possible, car nous avons supposé le bruit gaussien. On définit :

$$(5) \quad \varphi_A(u) = \text{TF}_a(P_{A/\varepsilon}(a))$$

$\varphi_A(u)$ est la première fonction caractéristique de $P_{A/\varepsilon}$.

En appliquant le *théorème de Parseval* à l'expression (3), la fonction de vraisemblance s'écrit :

$$(6) \quad P_{R/\varepsilon} = \int \varphi_A(u) \beta^*(u, r) du.$$

Pour calculer la fonctionnelle de vraisemblance, nous devons exprimer la première fonction caractéristique des sources $\varphi_A(u)$.

3.3. DÉVELOPPEMENT DE LA FONCTION CARACTÉRISTIQUE DES SOURCES $\varphi_A(u)$

Le calcul de $\varphi_A(u)$, de $\beta(u, r)$ et de l'intégrale (6) est développé dans l'annexe 1. Rappelons que la forme de $\beta(u, r)$ est connue puisque l'on a supposé le bruit gaussien, alors que celle de $\varphi_A(u)$ est inconnue.

Pour calculer la fonction caractéristique $\varphi_A(u)$ des variables aléatoires A , il est nécessaire de connaître la loi de A . En général, cette loi est inconnue. Lorsque A est gaussienne, $\varphi_A(u)$ est totalement définie par sa moyenne et sa variance. Lorsque A est inconnue, on peut l'approcher au deuxième ordre par une loi gaussienne, de même moyenne et de même variance.

Comme nous l'avons indiqué dans l'introduction, nous étendrons cette démarche en supposant connus les cumulants de A jusqu'à l'ordre 4. Dans ce cas, on peut approcher la loi de A par un développement de Gram-Charlier (annexe 3) [7].

Nous supposons nuls les moments d'ordre impair (un et trois) : cette situation correspond à de nombreux cas pratiques où la loi de probabilité de A est paire.

$\varphi_A(u)$ prend alors la forme : (annexe 1).

$$(8) \quad \varphi_A(u) = \exp\left(-\frac{\sum_{ij} P_{ij} u_i u_j}{2}\right) \times \left[1 + \frac{1}{4!} \sum_{ijkl} \text{Cum}_{ijkl} u_i u_j u_k u_l\right]$$

Avec :

Cum : tenseur des cumulants d'ordre 4 des amplitudes aléatoires des sources

P : matrice d'intercorrélation des amplitudes aléatoires des sources.

Ayant exprimé $\varphi_A(u)$, nous pouvons calculer la fonction de vraisemblance à partir de l'expression (6). Le calcul est développé en annexe 1.

Vérifions tout d'abord les résultats obtenus dans le cas gaussien.

3.4. APPLICATION AU CAS GAUSSIEN

L'expression de la fonction de vraisemblance, obtenue en annexe 1, est indépendante de la transformation Π : celle-ci est *non observable* par une analyse au deuxième ordre. Cette propriété explique la nécessité d'hypothèses supplémentaires dans le cas gaussien, pour déterminer les vecteurs signaux.

Les paramètres \underline{V} , $\underline{\Delta}$, P_B sont observables. Il sera possible d'écrire les estimateurs du maximum de vraisemblance de ces grandeurs. Ces estimateurs sont identiques aux expressions classiques des paramètres d'un tel système (la démonstration est donnée en annexe 1) [6], [12].

4. La Log-vraisemblance d'ordre 4

On dispose en fait de M réalisations qui nous permettent de moyennner la fonction de vraisemblance.

Afin de calculer le maximum de la fonction de vraisemblance, il est courant de considérer le logarithme de cette fonction (LMV).

En intégrant (6), il vient alors (cf. : annexe 1) :

$$(9) \quad \text{LMV} = \log(\text{facteur}) - \frac{n-p}{2} \log(P_B) - \frac{1}{2} \log(\det(\Gamma)) - \frac{K(r)}{2} - \frac{\text{Tr}(M\hat{\Gamma})}{2P_B}$$

[-----]
I₁

$$+ \frac{1}{4!} \sum_{ijkla} \frac{\sqrt{\lambda_{ii}} \sqrt{\lambda_{jj}} \sqrt{\lambda_{kk}} \sqrt{\lambda_{ll}}}{\sqrt{\Gamma_{ii}} \sqrt{\Gamma_{jj}} \sqrt{\Gamma_{kk}} \sqrt{\Gamma_{ll}}} \times \Pi_{ia} \Pi_{ja} \Pi_{ka} \Pi_{la} \text{cum}_a \hat{K} \text{urtosis}_{ijkl}$$

[-----]
I₂

avec :

facteur : terme regroupant toutes les constantes d'intégration

n : nombre de capteurs supérieur ou égal aux nombres de sources

p : nombre de sources

P_B : puissance du bruit

$\underline{\Gamma} = \underline{\Delta} + P_B \underline{I}$: $\underline{\Gamma}$ est une matrice diagonale dont les termes non nuls sont notés Γ_{ii}

$\underline{\Delta}$: $\underline{\Delta}$ est une matrice diagonale dont les termes non nuls sont notés λ_{ii}

$K(r) = r^T \underline{V} \underline{\Gamma}^{-1} \underline{V}^T r$

$\underline{\Sigma} = \underline{\Sigma}^T \underline{\Sigma}$

$\underline{M} = \underline{I} - \underline{\Sigma} \underline{\Sigma}^{-1} \underline{\Sigma}^T$: projecteur sur l'espace bruit

$\hat{\underline{\Gamma}} = \sum r r^T$: estimation de la matrice de corrélation des signaux reçus

$\underline{\Pi}$: $\underline{\Pi}$ est une matrice unitaire

cum_a : sont les termes non nuls du tenseur des cumulants d'ordre 4 des sources. Les sources étant indépendantes ce tenseur est « diagonal ». Nous rappelons également que les sources cherchées sont supposées de puissance unité et donc les cumulants de tels signaux sont égaux aux Kurtosis de ces signaux

$\hat{K} \text{urtosis}_{ijkl}$: estimation des éléments du tenseur des kurtosis [4] des signaux reçus (cf. : annexe 2).

On observe que la log vraisemblance peut se décomposer en deux parties :

— Une première partie I_1 identique à celle obtenue dans le cas gaussien.

— Une seconde partie I_2 contenant les informations supplémentaires fournies par le quatrième ordre (nulle dans le cas gaussien).

Ayant obtenu la fonction de vraisemblance, nous en déduisons les estimateurs du maximum de vraisemblance des paramètres.

5. Les estimateurs

Nous avons vu que la matrice $\underline{\Sigma}$ peut se décomposer de la manière suivante :

$$\underline{\Sigma} = \underline{V} \underline{\Delta}^{1/2} \underline{\Pi}$$

Il nous faut donc déterminer les estimateurs des matrices \underline{V} , $\underline{\Delta}$ et $\underline{\Pi}$.

La morphologie de l'expression de la fonction de log-vraisemblance nous permet de distinguer deux cas :

5.1. Cas d'un fort rapport signal sur bruit

En l'absence de bruit, ou dans le cas d'un fort rapport signal sur bruit ($P_B \ll \lambda_{ii}$), on peut considérer que l'on a égalité entre les puissances des sources λ_{ii} et la puissance des signaux reçus décorrélés, données par les valeurs propres Γ_{ii} .

On a alors :

$$\begin{aligned} \text{LMV} &= \log(\text{facteur}) - \frac{1}{2} \log(\det(\Gamma)) - \frac{K(r)}{2} + \frac{1}{4!} \sum_{ijkla} \Pi_{ia} \Pi_{ja} \Pi_{ka} \Pi_{la} \text{cum}_a \widehat{\text{Kurtosis}}_{ijkl} \\ &= \underbrace{\hspace{10em}}_{I_1} + \underbrace{\hspace{10em}}_{I_2} \end{aligned}$$

Le terme relatif au quatrième ordre ne contient plus d'éléments des matrices \underline{V} et $\underline{\Delta}$.

On peut alors envisager une étude en deux étapes :

1. Une étude, au deuxième ordre, qui permet de déterminer les estimateurs des matrices \underline{V} et $\underline{\Delta}$.

L'analyse au second ordre, prenant seulement en compte le terme I_1 de la fonction de vraisemblance, permet de déterminer les estimateurs du maximum de vraisemblance (cf. annexe 1) :

$$\begin{aligned} &\widehat{V}_{ij} \text{ de } V_{ij} \\ &\widehat{P}_B \text{ de } P_B \text{ (dans le cas général)} \\ &\widehat{\lambda}_{ii} \text{ de } \lambda_{ii}. \end{aligned}$$

Cette étude nous donne également : $\widehat{\Gamma}_{ii} = \widehat{\lambda}_{ii} + \widehat{P}_B$ estimateur de Γ_{ii} .

2. Une étude purement au quatrième ordre qui permet de déterminer l'estimateur de la matrice $\underline{\Pi}$.

Une telle démarche, nous ramènera à l'étude de signaux sources normés, ayant subi une transformation unitaire $\underline{\Pi}$. Dans ce cas, le tenseur cumulant d'ordre 4 des signaux sources (normalisés) s'exprime en fonction du tenseur Kurtosis des signaux reçus et de la transformation $\underline{\Pi}$. Dans une première approche, nous avons imposé la même forme à l'estimateur des cumulants des sources :

$$\widehat{\text{cum}}_a = \sum_{ijkl} \Pi_{ia} \Pi_{ja} \Pi_{ka} \Pi_{la} \widehat{\text{Kurtosis}}_{ijkl}.$$

Le terme d'ordre quatre qu'il nous fait maximiser est donc égal à une somme de carrés des cum_a .

La fonction à maximiser devient donc :

$$\begin{aligned} \text{LMV} &= \log(\text{facteur}) - \frac{1}{2} \log(\det(\Gamma)) - \frac{K(r)}{2} + \\ &+ \frac{1}{4!} \sum_a \left(\sum_{ijkl} \Pi_{ia} \Pi_{ja} \Pi_{ka} \Pi_{la} \widehat{\text{Kurtosis}}_{ijkl} \right)^2. \end{aligned}$$

Cela revient à chercher une matrice $\underline{\Pi}$ qui diagonalise le tenseur Kurtosis des signaux reçus. La matrice $\underline{\Pi}$ peut être vue comme une « matrice propre » du tenseur Kurtosis. La recherche de matrice propre de tenseur d'ordre 4 est notamment abordée dans [17].

5.2. CAS D'UN RAPPORT SIGNAL SUR BRUIT QUELCONQUE

Un deuxième cas de figure peut se présenter : celui d'un rapport signal sur bruit quelconque.

Dans ce cas, le terme d'ordre 4 intervient dans la détermination de l'estimateur du maximum de vraisem-

blance de la matrice $\underline{\Delta}$. Une résolution possible du problème est, à l'instar du cas précédent, d'effectuer une étude en deux temps :

1. Une extraction des informations contenues dans le second ordre, permet de se ramener à une base orthonormée de signaux : \underline{V} et $\underline{\Delta}$, et de déterminer les estimateurs des paramètres \underline{V} , $\underline{\Delta}$, P_B (voir § 5.1). Puis l'on effectue une normalisation en puissance.

2. Une étude exclusive des cumulants d'ordre 4 permet de déterminer les degrés de liberté du problème. Cette étude tient compte de la puissance relative des sources par rapport aux signaux reçus. On pourrait alors envisager une étude itérative afin d'affiner l'estimation de la puissance des sources. Mais la convergence d'une telle méthode reste à démontrer.

Illustrons les résultats obtenus par une simulation.

6. Simulation

Cette simulation est réalisée dans le cas où l'on dispose d'un fort rapport signal sur bruit. On a donc à réaliser une étude en deux temps (voir § 5.1). La première partie de l'étude est classique [6], [12]. Nous nous sommes attachés à simuler la seconde partie. Le but de la simulation est de séparer deux sources mélangées par une transformation unitaire $\underline{\Pi}$. Dans ce cas bidimensionnel, la matrice $\underline{\Pi}$ peut s'écrire en fonction d'un seul paramètre θ :

$$\begin{aligned} \Pi_{11} &= \cos \theta & \Pi_{12} &= \sin \theta \\ \Pi_{21} &= -\sin \theta & \Pi_{22} &= \cos \theta \end{aligned}$$

A partir de deux signaux gaussiens indépendants de puissance unité u et v , on peut construire deux signaux non gaussiens, de puissance unité, indépendants a et b .

On construit les signaux a et b de la manière suivante :

— Pour tous les points $a(n)$: $a(n) = \arctangente(\text{coefficient}_1 * u(n))$

— Pour tous les points $b(n)$: $b(n) = \sinus \text{ hyperbolique}(\text{coefficient}_2 * v(n))$.

Puis les signaux a et b sont normés en puissance.

Les paramètres coefficient_1 et coefficient_2 , nous permettent de générer des signaux « plus ou moins éloignés d'une distribution gaussienne ».

On peut alors construire deux signaux x_1 , x_2 combinaison linéaire des signaux a et b .

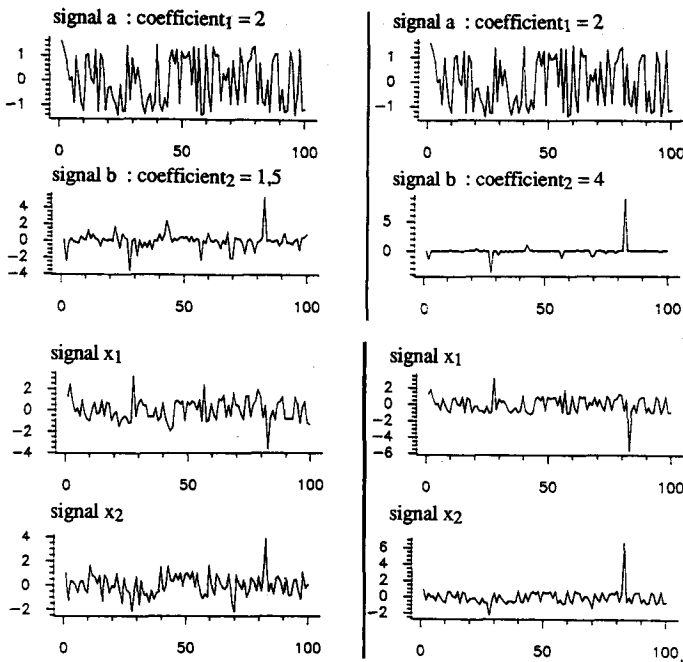
$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \cos(\theta_0) - \sin(\theta_0) \\ \sin(\theta_0) \cos(\theta_0) \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}.$$

Les signaux x_1 et x_2 sont alors non gaussiens et dépendants. Ils sont également les seules données dont on dispose.

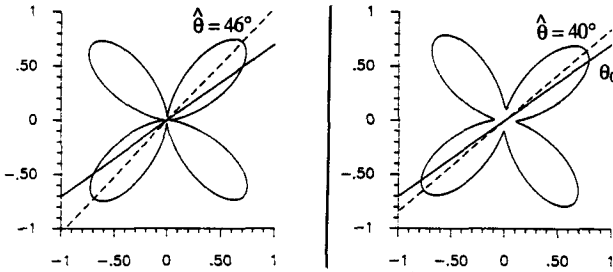
On trouve alors l'allure de la fonctionnelle de vraisemblance. Le maximum de cette fonction nous permet d'estimer un angle $\hat{\theta}$. On calcule alors les signaux « indépendants » a' et b' :

$$\begin{pmatrix} a' \\ b' \end{pmatrix} = \begin{pmatrix} \cos(-\hat{\theta}) & -\sin(-\hat{\theta}) \\ \sin(-\hat{\theta}) & \cos(-\hat{\theta}) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

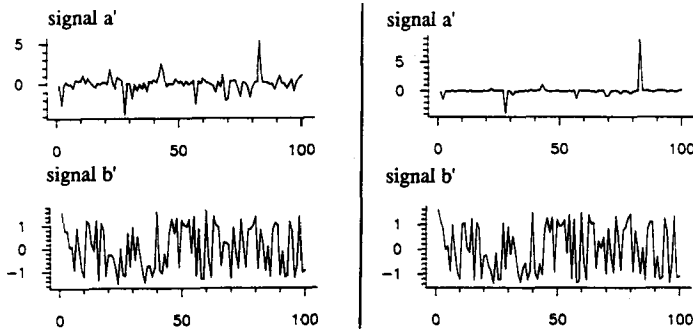
Nous présentons ici les résultats de deux simulations dont les signaux sont déformés de deux manières différentes. Pour une réalisation effectuée sur 100 points, pour un angle de mélange θ_0 égal à 35 degrés, on obtient :



Tracé de la fonction de vraisemblance



Signaux restitués

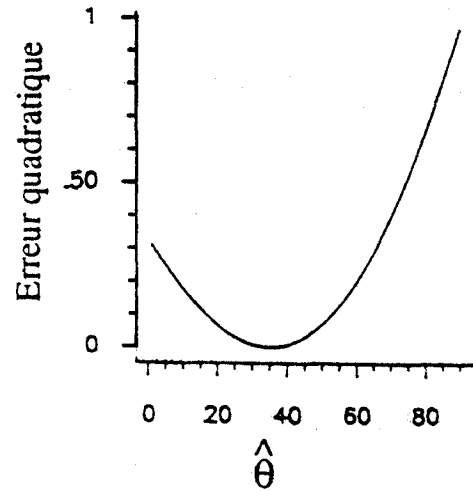


Une première approche des performances statistiques de ces estimateurs peut également être donnée :

Pour 50 réalisations de 100 points on a :

Estimation de θ_0	Moyenne	Ecart-type
coefficient ₁ = 2 coefficient ₂ = 1,5	35,1	7,5
coefficient ₁ = 2 coefficient ₂ = 4	35,0	5,5

A titre indicatif, il est intéressant d'évaluer quantitativement l'écart quadratique entre les signaux sources et leurs estimées, avec les réserves qui s'imposent s'agissant de signaux non gaussiens.



On donne ici l'écart quadratique entre les signaux a et a' pour toutes les valeurs possibles de $\hat{\theta}$, avec coefficient₁ = 2 et coefficient₂ = 1,5.

Cette erreur quadratique nous montre que l'estimation de θ peut être entachée d'erreur sans que la reconstitution des signaux ne soit significativement remise en cause.

Dans le cas général, on pourra mettre en œuvre la technique d'itération des matrices de rotation de Givens. On peut également envisager d'autres approches de ce problème d'optimisation non linéaire, par exemple par des réseaux neuromimétiques [13], ou une méthode tensorielle [17].

8. Conclusion

Dans le problème de la séparation de sources non gaussiennes combinées linéairement, nous avons effectué un

développement de Gram-Charlier de la loi de probabilité des sources, mettant en jeu les moments d'ordre supérieur. Nous avons paramétrisé l'espace source (par les vecteurs \underline{V}), la puissance des sources (par $\underline{\Delta}$) et finalement les sources (par $\underline{\Pi}$).

Nous avons montré que l'indétermination, lorsque l'on se limite aux moments d'ordre 2, est due à la non-observabilité de la matrice $\underline{\Pi}$ à cet ordre.

Nous avons construit l'estimateur des caractéristiques des sources et du bruit en deux étapes :

1^{re} étape : Utilisation des moments d'ordre 2 pour estimer l'espace source par \underline{V} , la puissance des sources par $\underline{\Delta}$ et la puissance du bruit par P_B .

2^e étape : En fixant pour \underline{V} , P_B et $\underline{\Delta}$, les valeurs obtenues à la première étape, nous utilisons les moments jusqu'à l'ordre 4 pour estimer les sources par la matrice $\underline{\Pi}$ et les cumulants des sources.

Ce travail confirme les résultats obtenus en [14] et [17] pour la séparation de sources. Il devra se poursuivre par une analyse détaillée des propriétés de l'estimateur de $\underline{\Pi}$. Ce problème est en effet assez complexe d'un point de vue analytique et la solution proposée ici doit être considérée comme une première approche.

Cette étude a été menée pour des signaux à valeurs réelles. Ce modèle décrit la séparation d'un mélange de sources soumises à des transformations instantanées. Dans le cas général, il faudrait se placer dans le domaine des fréquences et envisager des signaux à valeurs complexes. Des études ont été entreprises pour étendre la procédure à des signaux à valeurs complexes ; nous nous proposons d'étendre à ce cas les résultats obtenus ici.

Travail réalisé avec le concours de la direction des Constructions Navales, dans le cadre de la convention DCN.CEPHAG.

Annexe 1

CALCUL DE L'ESTIMATEUR DU MAXIMUM DE VRAISEMBLANCE

1. Le modèle

L'étude menée précédemment a montré que la fonction du maximum de vraisemblance pouvait se résumer, après l'application du théorème de Parseval, à l'intégrale :

$$(6) \quad P_{R/\epsilon} = \int \varphi_A(u) \beta^*(u, r) du.$$

Nous allons ici présenter les étapes du calcul de cette intégrale. Ce calcul se fera en trois étapes :

- le calcul de $\beta(u, r)$
- le calcul de $\varphi(u)$
- le calcul de l'intégrale elle-même.

2. Calcul de $\beta(U, R)$

$$\beta(u, r) = \text{TF}_a(P_{\text{Bruit}}(r - \underline{S}a)).$$

Le bruit étant supposé blanc gaussien, la loi de probabilité est :

$$P_{\text{Bruit}}\left(r - \sum a_i s_i\right) = \frac{1}{(2\pi P_B)^{n/2}} \exp\left(-\frac{\left(r - \sum a_i s_i\right)^T \left(r - \sum a_i s_i\right)}{2 P_B}\right)$$

avec : P_B la puissance du bruit.

On pose :

$$x = \left(r - \sum a_i s_i\right)^T \left(r - \sum a_i s_i\right)$$

d'où :

$$x = r^T r - 2 \sum_i a_i r^T s_i + \sum_i \sum_j a_i a_j s_i^T s_j.$$

* Soit $\underline{\underline{S}}_{(n)}$ la matrice de dimension $(n \times p)$ dite « Longue » et formée des composantes des p vecteurs de bases normés sur les n capteurs :

$$\underline{\underline{S}}_{(n)} = (\underline{s}_1, \dots, \underline{s}_p).$$

Dans notre modèle, on a imposé la norme des vecteurs de base.

* Soit $\underline{\underline{\Sigma}}$ la matrice de dimension $p \times p$ définie par :

$$\underline{\underline{\Sigma}} = \underline{\underline{S}}_{(n)}^T \underline{\underline{S}}_{(n)} = \underline{\underline{S}}^T \underline{\underline{S}}$$

$\underline{\underline{\Sigma}}$ est une matrice constituée des produits scalaires des vecteurs sources.

$\underline{\underline{\Sigma}}$ est une matrice symétrique, sa diagonale est égale à la norme des vecteurs de base. Une telle structure implique que $\underline{\underline{\Sigma}}$ est une matrice définie non négative.

$\underline{\underline{S}}$ peut se mettre sous la forme :

$$\underline{\underline{S}} = \underline{\underline{V}} \underline{\underline{\Delta}}^{1/2} \underline{\underline{\Pi}}$$

on a alors :

$$\underline{\underline{\Sigma}} = \underline{\underline{S}}^T \underline{\underline{S}} = \underline{\underline{\Pi}}^T \underline{\underline{\Delta}}^{1/2} \underline{\underline{\Delta}}^{1/2} \underline{\underline{\Pi}}$$

on définit $\underline{\underline{\Sigma}}^{1/2}$ tel que :

$$\underline{\underline{\Sigma}}_{ik} = \underline{\underline{\Sigma}}_{ji}^{1/2} \underline{\underline{\Sigma}}_{jk}^{1/2} \text{ soit } \underline{\underline{\Sigma}}^{1/2} = \underline{\underline{\Delta}}^{1/2} \underline{\underline{\Pi}}$$

on a alors :

$$\underline{\underline{\Sigma}}^{1/2} \underline{\underline{\Sigma}}^{1/2 T} = \underline{\underline{\Delta}} \text{ et } \underline{\underline{S}} = \underline{\underline{V}} \underline{\underline{\Sigma}}^{1/2}$$

$\underline{\underline{\Sigma}}^{1/2}$ est donc la matrice de passage entre $\underline{\underline{S}}$ et $\underline{\underline{V}}$.

* Soit $\underline{\rho}$ le vecteur formé par :

$$\underline{\rho} = \begin{pmatrix} \underline{s}_1^T \underline{r} \\ \dots \\ \underline{s}_p^T \underline{r} \end{pmatrix} = \underline{\underline{S}}_{(n)}^T \underline{r}$$

x prend alors la forme : $x = \underline{r}^T \underline{r} - 2 \underline{\rho}^T \underline{a} + \underline{a}^T \underline{\Sigma} \underline{a}$

d'où :

$$\beta(u, r) = \frac{1}{(2 \pi \mathbf{P}_B)^{n/2}} \int \exp \left(\frac{-x}{2 \mathbf{P}_B} - i u^T \underline{a} \right) da$$

posons :

$$\delta = \frac{\underline{\Sigma}}{\mathbf{P}_B}, \quad m = \frac{\underline{\rho}}{\mathbf{P}_B} - i u$$

on a alors :

$$\beta(u, r) = \frac{1}{(2 \pi \mathbf{P}_B)^{n/2}} \exp \left(\frac{-\underline{r}^T \underline{r}}{2 \mathbf{P}_B} \right) \int \exp \left(-\frac{\underline{a}^T \delta \underline{a}}{2} + m^T \underline{a} \right) da.$$

Appliquons le lemme 1 :

$$(7) \quad \beta(u, r) = \frac{1}{(2 \pi \mathbf{P}_B)^{n-p/2} (\det \underline{\Sigma})^{1/2}} \exp \left(-\frac{\Phi(u, r)}{2 \mathbf{P}_B} \right)$$

avec : $\Phi(u, r) = \underline{r}^T \mathbf{M} \underline{r} + \mathbf{P}_B^2 \underline{u}^T \underline{\Sigma}^{-1} \underline{u} + 2 i \mathbf{P}_B \underline{r}^T \underline{\Sigma}^{-1} \underline{u}$

$\underline{\mathbf{M}} = \underline{\mathbf{I}} - \underline{\underline{S}} \underline{\underline{\Sigma}}^{-1} \underline{\underline{S}}^T$; $\underline{\mathbf{M}}$ est le projecteur sur l'espace bruit.

3. Calcul de $\varphi_A(u)$

Afin de pouvoir poursuivre les calculs, on réalise un développement de Gram-Charlier, de la première fonction caractéristique (annexe 3) [7].

Supposant les signaux de distributions symétriques et centrés, nous nous sommes limités à l'ordre 4. Sous de telles hypothèses les moments d'ordre impair sont tous nuls.

On peut alors approximer la fonction caractéristique par (annexe 3) :

$$(8) \quad \varphi_A(u) \approx \exp\left(-\frac{\sum_{ij} \mathbb{P}_{ij} u_i u_j}{2}\right) \left[1 + \frac{1}{4!} \sum_{ijkl} \text{Cum}_{ijkl} u_i u_j u_k u_l\right]$$

Cum_{ijkl} : tenseur des cumulants d'ordre 4 des sources

\mathbb{P}_{ij} : matrice d'intercorrélation des sources, de dimension $(p \times p)$.

Le terme exponentiel permettra de calculer l'intégrale de la fonction de vraisemblance, en utilisant les propriétés de la transformée de Fourier.

Dans le cas de signaux indépendants, la matrice \mathbb{P} est diagonale et les éléments du tenseur Cum sont tous nuls, à l'exception des termes du type Cum_{iii} (autocumulants). Nous les noterons Cum_i .

On réalise ici une troncature d'un développement limité, il faudra s'assurer de la validité de cette hypothèse.

Remarque : Dans ce développement, on voit apparaître, d'une manière naturelle, les cumulants d'un signal stationnaire centré.

4. Calcul de la fonction de vraisemblance

4.1. LA FONCTION DE VRAISEMBLANCE POUR UNE RÉALISATION

La fonction de vraisemblance s'écrit :

$$P_{R/\epsilon} = \text{facteur} \int \exp\left(\frac{-1}{2} \left(\sum_{ij} \mathbb{P}_{ij} u_i u_j + P_B \sum_{ij} \Sigma_{ij}^{-1} u_i u_j - 2i \sum_{lki} r_l S_{lk} \Sigma_{ki}^{-1} u_i\right)\right) \left[1 + \frac{1}{4!} \sum_{ijkl} \text{Cum}_{ijkl} u_i u_j u_k u_l\right] du.$$

Effectuons le changement de variable :

$$u = \underline{\Sigma}^{1/2 T} w \quad \text{d'où} \quad du = \det(\underline{\Sigma}^{1/2 T}) dw$$

$$(6) \Rightarrow P_{R/\epsilon} = \text{facteur} \int \exp\left(\frac{-1}{2} \left(\sum_{ij} \left(\sum_{kl} \Sigma_{ik}^{1/2} \mathbb{P}_{kl} \Sigma_{jl}^{1/2} + P_B \sum_{kl} \Sigma_{ik}^{1/2} \Sigma_{kl}^{-1} \Sigma_{jl}^{1/2}\right) w_i w_j - 2i \sum_{ljk} r_l S_{lj} \Sigma_{jk}^{-1} \Sigma_{ik}^{1/2} w_i\right)\right) \left[1 + \frac{1}{4!} \sum_{ijkl} \left[\sum_a \text{Cum}_a \Sigma_{ia}^{1/2} \Sigma_{ja}^{1/2} \Sigma_{ka}^{1/2} \Sigma_{la}^{1/2}\right] w_i w_j w_k w_l\right] \det(\Sigma^{1/2}) dw$$

on pose $\sum_a \text{Cum}_a \Sigma_{ia}^{1/2} \Sigma_{ja}^{1/2} \Sigma_{ka}^{1/2} \Sigma_{la}^{1/2} = \text{coeff}_{ijkl}$: Coeff est un tenseur d'ordre 4 avec : $\Sigma_{ij}^{1/2} = \sqrt{\lambda_{ii}} \Pi_{ij}$

$$P_{R/\epsilon} = \text{facteur} \int \exp\left(\frac{-1}{2} \left(\sum_{ij} (\Delta_{ij} + P_B) w_i w_j - 2i \sum_{ijl} r_l S_{ij} \Sigma_{ji}^{-1/2} w_i\right)\right) \left[1 + \frac{1}{4!} \sum_{ijkl} \text{coeff}_{ijkl} w_i w_j w_k w_l\right] \det(\Sigma^{1/2}) dw$$

la fonction de vraisemblance peut se décomposer en deux parties :

$$P_{R/\epsilon} = J_1 + J_2.$$

Afin de calculer J_1 , on pose : $m^T = r^T \underline{\Sigma} \underline{\Sigma}^{-1/2} = r^T \underline{V}$ et l'application du lemme I, nous donne :

$$J_1 = \frac{\text{facteur} \det(\Sigma^{1/2})}{P_B^{\frac{n-p}{2}} (\det \Gamma \Sigma)^{1/2}} \exp\left(-\frac{r^T M r}{2 P_B}\right) \exp\left(-\frac{K(r)}{2}\right)$$

avec :

$$K(r) = r^T \underline{\Sigma} \underline{\Sigma}^{-1/2} \underline{\Gamma}^{-1/2} \underline{\Sigma}^{-1/2} \underline{\Sigma}^T r$$

$$J_1 = \frac{\text{facteur}}{P_B^{\frac{n-p}{2}} (\det \Gamma)^{1/2}} \exp\left(-\frac{r^T M r}{2 P_B}\right) \exp\left(-\frac{K(r)}{2}\right)$$

avec :
$$\mathbf{K}(r) = r^T \underline{\mathbf{V}} \underline{\mathbf{\Gamma}}^{-1} \underline{\mathbf{V}}^T r = m^T \underline{\mathbf{\Gamma}}^{-1} m$$

L'intégrale J_2 est calculée en utilisant la propriété de dérivation de la transformée de Fourier.

Si :
$$f(x) \stackrel{\text{TF}}{\rightleftharpoons} F(v)$$

 alors :
$$x^n f(x) \rightleftharpoons j^n \frac{\partial^n}{\partial v^n} (F(v)).$$

Appliquons le lemme II :

$$J_2 = \frac{\text{facteur}}{\mathbf{P}_B^{\frac{n-p}{2}} (\det \Gamma)^{1/2}} \exp \left(- \frac{\sum_{ij} r_j M_{ji} r_i}{2 \mathbf{P}_B} \right) \frac{1}{4!} \sum_{ijkl} \text{Coeff}_{ijkl} \frac{\partial^4}{\partial m_i \partial m_j \partial m_k \partial m_l} \exp \left(- \frac{\mathbf{K}(r)}{2} \right)$$

L'expression à dériver est l'exponentielle d'une forme quadratique. Les dérivées successives de telles fonctions conduisent à des polynômes d'Hermite (annexe 2) [7].

La fonction de vraisemblance peut alors se factoriser sous la forme :

$$\mathbf{P}_{R/\varepsilon} = J_1 + J_2 = J_1 \text{ poly} (\underline{\mathbf{V}}, \underline{\mathbf{\Delta}}, \mathbf{P}_B, \text{cum}, \underline{\mathbf{\Pi}}/r).$$

Les dérivées successives permettant le calcul du polynôme poly ($\underline{\mathbf{V}}, \underline{\mathbf{\Delta}}, \mathbf{P}_B, \text{cum}, \underline{\mathbf{\Pi}}/r$) sont présentées en annexe 2.

4.2. LA FONCTION DE LOG-VRAISEMBLANCE POUR « M » RÉALISATIONS

Les fonctions de vraisemblance ont été écrites pour une réalisation. On dispose en fait de M réalisations indépendantes. La fonction de vraisemblance est alors le produit des probabilités élémentaires. Afin de calculer le maximum de la fonction de vraisemblance, il est courant de considérer la moyenne du logarithme de cette fonction (LMV) :

$$\begin{aligned} \text{LMV} = & \log(\text{facteur}) - \frac{n-p}{2} \log(\mathbf{P}_B) - \frac{1}{2} \log(\det(\underline{\mathbf{\Gamma}})) - \frac{\mathbf{K}(r)}{2} - \frac{\text{Tr}(\underline{\mathbf{M}} \hat{\underline{\mathbf{\Gamma}}})}{2 \mathbf{P}_B} \\ & \text{[-----]} \\ & I_1 \\ & + \frac{1}{M} \log \left(\prod_{r=1}^M \text{poly}^{(r)} (\underline{\mathbf{V}}, \underline{\mathbf{\Delta}}, \mathbf{P}_B, \text{cum}, \underline{\mathbf{\Pi}}/r) \right) \\ & \text{[-----]} \\ & I_2 \end{aligned}$$

avec :

facteur : terme regroupant toutes les constantes d'intégration

n : nombre de capteurs supérieur ou égal aux nombres de sources

p : nombre de sources

\mathbf{P}_B : puissance du bruit.

$$\underline{\mathbf{\Gamma}} = \underline{\mathbf{\Delta}} + \mathbf{P}_B \underline{\mathbf{I}}$$

$$\underline{\mathbf{P}} = E[\mathbf{A}\mathbf{A}^T]$$

$$\mathbf{K}(r) = r^T \underline{\mathbf{V}} \underline{\mathbf{\Gamma}}^{-1} \underline{\mathbf{V}}^T r$$

$\underline{\mathbf{M}} = \underline{\mathbf{I}} - \underline{\mathbf{S}} \underline{\mathbf{S}}^{-1} \underline{\mathbf{S}}^T$: projecteur sur l'espace bruit

$$\hat{\underline{\mathbf{\Gamma}}} = \frac{1}{M} \sum r r^T : \text{estimée de la matrice } \Gamma$$

poly ($\underline{\mathbf{V}}, \underline{\mathbf{\Delta}}, \mathbf{P}_B, \text{cum}, \underline{\mathbf{\Pi}}/r$) polynôme de degré 4 provenant de la réduction d'un tenseur d'ordre 4 constitué de *cumulants et de polynômes d'Hermite.*

On observe que la log-vraisemblance peut se décomposer en deux parties :

I_1 : une première partie identique à celle obtenue dans le cas gaussien

I_2 : une seconde partie contenant les informations supplémentaires fournies par le quatrième ordre.

Pour une réalisation r , le terme polynomial s'écrit :

$$\text{poly}^{(r)}(\underline{V}, \underline{\Delta}, P_B, \text{cum}, \underline{\Pi}/r) = 1 + \frac{1}{4!} \sum_{ijkl} \text{coeff}_{ijkl} H_{ijkl}^{(r)}$$

$$\text{poly}^{(r)}(\underline{V}, \underline{\Delta}, P_B, \text{cum}, \underline{\Pi}/r) = 1 + \frac{1}{4!} \sum_{ijkla} \text{cum}_a \sqrt{\lambda_{ii}} \Pi_{ia} \sqrt{\lambda_{jj}} \Pi_{ja} \sqrt{\lambda_{kk}} \Pi_{ka} \sqrt{\lambda_{ll}} \Pi_{la} H_{ijkl}^{(r)}$$

Le terme $H_{ijkl}^{(r)}$ s'écrit :

$$H_{ijkl}^{(r)} = \frac{1}{\sqrt{\Gamma_{ii}} \sqrt{\Gamma_{jj}} \sqrt{\Gamma_{kk}} \sqrt{\Gamma_{ll}}} \text{Kurtosis}_{ijkl}^{(r)}$$

Les termes $H_{ijkl}^{(r)}$ et $\text{Kurtosis}_{ijkl}^{(r)}$, issus des dérivations successives de $K(r)$, sont détaillées en annexe 2.

S'agissant d'un développement limité, nous allons faire l'hypothèse suivante :

$$\frac{1}{4!} \text{coeff}_{ijkl} H_{ijkl}^{(r)} \ll 1.$$

Le terme I_2 devient alors :

$$I_2 = \frac{1}{M} \sum_r \left(\sum_{ijkla} \frac{1}{4!} \sqrt{\lambda_{ii}} \sqrt{\lambda_{jj}} \sqrt{\lambda_{kk}} \sqrt{\lambda_{ll}} \Pi_{ia} \Pi_{ja} \Pi_{ka} \Pi_{la} \text{cum}_a H_{ijkl}^{(r)} \right)$$

$$I_2 = \frac{1}{M} \sum_{ijkla} \left(\frac{1}{4!} \frac{\sqrt{\lambda_{ii}} \sqrt{\lambda_{jj}} \sqrt{\lambda_{kk}} \sqrt{\lambda_{ll}}}{\sqrt{\Gamma_{ii}} \sqrt{\Gamma_{jj}} \sqrt{\Gamma_{kk}} \sqrt{\Gamma_{ll}}} \Pi_{ia} \Pi_{ja} \Pi_{ka} \Pi_{la} \text{cum}_a \sum_r \text{Kurtosis}_{ijkl}^{(r)} \right)$$

$$I_2 = \frac{1}{4!} \sum_{ijkla} \frac{\sqrt{\lambda_{ii}} \sqrt{\lambda_{jj}} \sqrt{\lambda_{kk}} \sqrt{\lambda_{ll}}}{\sqrt{\Gamma_{ii}} \sqrt{\Gamma_{jj}} \sqrt{\Gamma_{kk}} \sqrt{\Gamma_{ll}}} \Pi_{ia} \Pi_{ja} \Pi_{ka} \Pi_{la} \text{cum}_a \widehat{\text{Kurtosis}}_{ijkl}$$

avec Kurtosis_{ijkl} , la moyenne des $\text{Kurtosis}_{ijkl}^{(r)}$ des signaux reçus.

La moyenne du logarithme de la fonction de vraisemblance s'écrit :

$$\begin{aligned} \text{LMV} = \log(\text{facteur}) - \frac{n-p}{2} \log(P_B) - \frac{1}{2} \log(\det(\Gamma)) - \frac{K(r)}{2} - \frac{\text{Tr}(M\hat{\Gamma})}{2P_B} + \\ + \frac{1}{4!} \sum_{ijkla} \frac{\sqrt{\lambda_{ii}} \sqrt{\lambda_{jj}} \sqrt{\lambda_{kk}} \sqrt{\lambda_{ll}}}{\sqrt{\Gamma_{ii}} \sqrt{\Gamma_{jj}} \sqrt{\Gamma_{kk}} \sqrt{\Gamma_{ll}}} \Pi_{ia} \Pi_{ja} \Pi_{ka} \Pi_{la} \text{cum}_a \widehat{\text{Kurtosis}}_{ijkl} \end{aligned} \quad (9)$$

5. DÉTERMINATION DES ESTIMATEURS DU MAXIMUM DE VRAISEMBLANCE

On cherche des solutions de la forme : $\underline{S} = \underline{V} \underline{\Delta}^{1/2} \underline{\Pi}$.

La morphologie de l'expression de la fonction de log-vraisemblance nous permet de distinguer deux cas dépendants du rapport signal sur bruit. L'étude pourra alors être menée en deux étapes, dont la première sera la détermination des estimateurs des paramètres V, Δ, P_B . Cela revient à se placer dans le cas de signaux gaussiens.

Nous rappelons ici les expressions de ces estimateurs.

La log-vraisemblance pour des signaux gaussiens, s'écrit alors :

$$\text{LMV} = \log(k) - \frac{n-p}{2} \log(P_B) - \frac{1}{2} \log(\Pi(\lambda_i - P_B)) - \frac{K(r)}{2} - \frac{\text{Tr}(M_{ik} \hat{\Gamma}_{kj})}{2P_B}$$

Cette fonction permet de déterminer les estimateurs des paramètres $\underline{V}, \underline{\Delta}$ et P_B .

5.1. DÉTERMINATION DE L'ESTIMATEUR DES V_i

On va maximiser la log-vraisemblance sous la contrainte $|V_i| = 1$:

$$\frac{\delta}{\delta v_i} (\text{LMV} - \beta v_i v_i^T) = 0 \Rightarrow \boxed{\frac{\hat{\Gamma} v_i}{\lambda_i + P_B} - \beta v_i = 0} .$$

Les V sont donc les vecteurs propres de $\hat{\Gamma}$:

$V_1 \dots V_p$: vecteurs propres dominants, associés aux valeurs propres maximales, estimés de $\hat{\Gamma}$;
 $V_{p+1} \dots V_n$: autres vecteurs propres estimés de $\hat{\Gamma}$.

5.2. DÉTERMINATION DES ESTIMATEURS DE λ_i ET DE P_B

Dans la base des vecteurs propres, la moyenne du logarithme de la fonction de vraisemblance prend la forme :

$$\text{LMV} = \log(k) - \frac{n-p}{2} \log(P_B) - \frac{1}{2} \sum_1^p (\log(\lambda_{ii} + P_B)) - \frac{1}{2} \sum_1^p \frac{\Gamma_{ii}}{\lambda_{ii} + P_B} - \frac{1}{2 P_B} \sum_{p+1}^n \Gamma_{ii}$$

$$\frac{\delta}{\delta \lambda_i} (\text{LMV}) = 0 \Rightarrow \boxed{\hat{\lambda}_{ii} = \hat{\Gamma}_{ii} - \hat{P}_B} .$$

D'où :

$$\text{LMV} = \log(k) - \frac{n-p}{2} \log(P_B) - \frac{1}{2} \sum_1^p (\log(\Gamma_{ii})) - \frac{p}{2} - \frac{1}{2 P_B} \sum_{p+1}^n \Gamma_{ii}$$

$$\frac{\delta}{\delta P_B} (\text{LMV}) = 0 \Rightarrow \boxed{\hat{P}_B = \frac{1}{n-p} \sum_{p+1}^n \Gamma_{ii}} .$$

LEMME I : Au cours du calcul, il est nécessaire d'intégrer des expressions de la forme :

$$I = \int \exp\left(-\frac{u^T \Gamma u}{2} + m^T u\right) du$$

donnant :

$$I = \frac{(2\pi)^{p/2}}{(\text{Det } \Gamma)^{1/2}} \exp\left(\frac{m^T \Gamma^{-1} m}{2}\right)$$

où m et u sont des vecteurs de dimension p , u est à valeurs réelles et Γ une matrice de rang plein de dimension $p \times p$.

LEMME II : L'intégrale :

$$J = \int \left(\sum_{ijkl} \text{Cum}_{ijkl} u_i u_j u_k u_l \right) \exp\left(\frac{-\sum_{ij} \Gamma_{ij} u_i u_j}{2} + \sum_k m_k u_k\right) du$$

donne :

$$J = I \times \left[\sum_{ijkl} \text{Cum}_{ijkl} \frac{\delta^4}{\delta m_i \delta m_j^T \delta m_k \delta m_l^T} \exp\left(\frac{-\sum_{ab} \Gamma_{ab}^{-1} m_a m_b^T}{2}\right) \right] .$$

Annexe 2

CALCUL DE DÉRIVÉES ET DÉFINITION DES KURTOSIS

Calcul des dérivées successives nécessaires à l'intégration du développement de la fonction de vraisemblance, dans le cas où $f(m)$ peut être mis sous la forme :

$$f(m) = \exp\left(-\frac{1}{2} \sum_i \Gamma_{ii}^{-1} m_i m_i\right) \quad \text{avec} \quad \underline{m} = \underline{V}^T \cdot r$$

$$\frac{\partial^4}{\partial m_i \partial m_i \partial m_i \partial m_i} f(m) = ((\Gamma_{ii}^{-2} m_i m_i)^2 - 6 \Gamma_{ii}^{-3} m_i m_i + 3 \Gamma_{ii}^{-2}) f(m) = H_{iiii} f(m) = \Gamma_{ii}^{-2} \text{Kurtosis}_{iiii} f(m)$$

avec Kurtosis_{iiii} = $(\Gamma_{ii}^{-2} m_i^4 - 6 \Gamma_{ii}^{-1} m_i^2 + 3)$

$$\frac{\partial^4}{\partial m_i \partial m_i \partial m_j \partial m_j} f(m) = (\Gamma_{jj}^{-1} m_j) (-3 \Gamma_{ii}^{-2} m_i + \Gamma_{ii}^{-1} m_i \Gamma_{ii}^{-2} m_i m_i) f(m) = H_{iijj} f(m) =$$

$$= \Gamma_{ii}^{-3/2} \Gamma_{jj}^{-1/2} \text{Kurtosis}_{iijj} f(m)$$

avec Kurtosis_{iijj} = $(\Gamma_{jj}^{-1/2} m_j) (-3 \Gamma_{ii}^{-1/2} m_i + \Gamma_{ii}^{-3/2} m_i^3)$

$$\frac{\partial^4}{\partial m_i \partial m_i \partial m_j \partial m_j} f(m) = (\Gamma_{jj}^{-2} m_j m_j - \Gamma_{jj}^{-1}) (\Gamma_{ii}^{-2} m_i m_i - \Gamma_{ii}^{-1}) f(m) = H_{ijjj} f(m) = \Gamma_{ii}^{-1} \Gamma_{jj}^{-1} \text{Kurtosis}_{ijjj} f(m)$$

avec Kurtosis_{ijjj} = $(\Gamma_{jj}^{-1} m_j^2 - 1) (\Gamma_{ii}^{-1} m_i^2 - 1)$

$$\frac{\partial^4}{\partial m_i \partial m_j \partial m_k \partial m_k} f(m) = (\Gamma_{kk}^{-2} m_k m_k - \Gamma_{kk}^{-1}) (\Gamma_{jj}^{-1} m_j) (\Gamma_{ii}^{-1} m_i) f(m) = H_{ijkk} f(m) =$$

$$= \Gamma_{ii}^{-1/2} \Gamma_{jj}^{-1/2} \Gamma_{kk}^{-1} \text{Kurtosis}_{ijkk} f(m)$$

avec Kurtosis_{ijkk} = $(\Gamma_{kk}^{-1} m_k^2 - 1) (\Gamma_{jj}^{-1/2} m_j) (\Gamma_{ii}^{-1/2} m_i)$

$$\frac{\partial^4}{\partial m_i \partial m_j \partial m_k \partial m_l} f(m) = (\Gamma_{ll}^{-1} m_l) (\Gamma_{kk}^{-1} m_k) (\Gamma_{jj}^{-1} m_j) (\Gamma_{ii}^{-1} m_i) f(m) = H_{ijkl} f(m) =$$

$$= \Gamma_{ii}^{-1/2} \Gamma_{jj}^{-1/2} \Gamma_{kk}^{-1/2} \Gamma_{ll}^{-1/2} \text{Kurtosis}_{ijkl} f(m)$$

avec Kurtosis_{ijkl} = $(\Gamma_{ii}^{-1/2} m_i) (\Gamma_{jj}^{-1/2} m_j) (\Gamma_{kk}^{-1/2} m_k) (\Gamma_{ll}^{-1/2} m_l)$.

Les polynômes H_{ijkl} sont définis grâce aux dérivées successives de la fonction de Gauss multidimensionnelles. On retrouve une extension des polynômes d'Hermite monodimensionnels [7].

Il est alors possible de définir les Kurtosis multidimensionnels pour une réalisation (Kurtosis_{ijkl}). En moyennant pour M réalisations, on obtient un estimateur du tenseur des kurtosis des signaux reçus. Nous proposons cette extension de la définition des kurtosis, qui correspond à une normalisation des cumulants multidimensionnels. Cette définition correspond dans le cas unidimensionnel à celle proposée dans [4].

Annexe 3

DÉVELOPPEMENT DE GRAM CHARLIER

Soit une variable aléatoire \underline{X} , à n dimensions, de densité de probabilité $P_{\underline{X}}(\underline{x})$ et de seconde fonction caractéristique $\psi_{\underline{X}}(\underline{u})$.

Les moments d'ordre 1 et 3 de \underline{X} sont supposés nuls. Soient :

P_{ij} : la covariance de \underline{X}

C_{ijkl} : le tenseur des cumulants d'ordre 4 de \underline{X} .

La seconde fonction caractéristique de \underline{X} est :

$$\psi_{\underline{X}}(\underline{u}) = -\frac{1}{2!} \sum_{ij} P_{ij} u_i u_j + \frac{1}{4!} \sum_{ijkl} C_{ijkl} u_i u_j u_k u_l + \dots$$

On l'approche par :

$$\psi_{\underline{X}}(\underline{u}) \approx \psi_{\underline{Y}}(\underline{u}) + \frac{1}{4!} \sum_{ijkl} C_{ijkl} u_i u_j u_k u_l$$

\underline{Y} est une variable aléatoire gaussienne de même moyenne et de même variance que \underline{X} .

La première fonction caractéristique de \underline{X} est alors :

$$\varphi_{\underline{X}}(\underline{u}) \approx \varphi_{\underline{Y}}(\underline{u}) \left(1 + \frac{1}{4!} \sum_{ijkl} C_{ijkl} u_i u_j u_k u_l \right).$$

La densité de probabilité de \underline{X} est :

$$P_{\underline{X}}(\underline{x}) \approx P_{\underline{Y}}(\underline{x}) \left(1 + \frac{1}{4!} \sum_{ijkl} C_{ijkl} H_4(x_i, x_j, x_k, x_l, P_{ij}) \right)$$

H_4 est le polynôme d'Hermite d'ordre 4 et l'expression (1) donne le développement de Gram-Charlier de la densité de probabilité de \underline{X} [7], [23].

BIBLIOGRAPHIE

- [1] J. L. LACOUME, Modèle et traitement des signaux multidimensionnels. *Traitement du signal*, vol. 5, n° 2, 1988, pp. 87-106.
- [2] M. BOUVET, *Étude de certains problèmes de détection et normalisation adaptative*. Application à l'acoustique sous-marine : Thèse de doctorat d'état, Orsay, 1987.
- [3] H. CHERIFI, Modélisation, identification et système de détection de bruit non gaussien. *Rapport CEPHAG*, n° 30-86.
- [4] D. R. BRILLINGER, *Times series Data Analysis and Theory*. Holdenday 1981.
- [5] C. L. NIKIAS, *Bispectrum estimation*. A digital signal processing framework Proceeding of the IEEE, July 1987.
- [6] C. LATOMBE, *Détection et caractérisation des signaux à plusieurs composantes à partir de la matrice interspectrale*. Thèse de doctorat d'état INPG, 1983.
- [7] A. ANGOT, *Compléments de mathématiques*, sixième édition. Masson, 1982.
- [8] A. M. GOODBODY, *Cartesian tensor*. Ellis Horwood series in mathématiques and its applications, 1982.
- [9] Mc CULLAG, Tensor notation and cumulants of polynomials. *Biometrika* (1984) 71.3, pp. 461-76.
- [10] G. H. GOLUB, C. F. VAN LOAN, *Matrix computation*. North Oxford academic, 1983.
- [11] L. H. SIBUL, G. R. L. SOHIE, Application of operator theory to signal modelling and processing IX. *Colloque GRETSI* (1983), pp. 63-68.
- [12] G. BIENVENU, L. KOOP, Optimality of high resolution array processing using the eigensystem approach. *IEEE Transactions on ASSP*, vol. ASSP-31, n° 5, October 1983.
- [13] J. HERAULT, Brain and Neurocomputing. *EUSIPCO*, 1988, Grenoble, pp. 7-14.
- [14] P. RUIZ, J. L. LACOUME, Extraction of independent sources from correlated inputs : A solution base on cumulants. *Workshop on Higher-order spectral analysis*. Vail Colorado, June 1989, pp. 146-152.
- [15] L. BRILLOUIN, *Les tenseurs en mécanique et en élasticité*. Édition Jacques Gabay, 1987.
- [16] P. COMON, Separation of stochastic processes whose a linear mixture is observed. Workshop on Higher-order spectral analysis, Vail Colorado, June 1989, pp. 174-179.
- [17] J. F. CARDOSO, Eigen-structure of the fourth-order cumulant tensor with application to the blind source separation problem. *ICASSP 90*, April 1990, pp. 2655-2658.
- [18] H. MERMOZ, Imagerie, corrélation et modèles. *Annales des télécommunications*, 31, n° 1-2, 1976, pp. 17-36.
- [19] H. WANG, M. KAVEH, On the performance of signal subspace processing. Part II : Coherent wideband systems. *IEE ASSP*, Vol. 35, n° 11, November 1987.
- [20] M. WAX, T. J. SHAN, T. KAILATH, Detection of signals by information theoretic criteria. *IEEE ASSP*, Vol. 32, pp. 817-827, August 1985.
- [21] N. L. OWSLEY, Sonar array processing in array signal processing. S. Haykin editor, Prentice Hall. *Signal Processing*, serie 1985, pp. 115-193.
- [22] R. F. DWYER, Use of the Kurtosis statistic in the frequency domain as an aid in detecting random Signals. *IEEE J. of Oceanic Engineering*, Vol. OE-9, n° 2, April 1984, pp. 85-92.
- [23] M. G. KENDALL, A. STUART, *The advanced theory of statistics*. Vol. 1. Charles Griffin & Company limited, London.

Manuscrit reçu le 1^{er} février 1990