



## SÉPARATION AVEUGLE DE SOURCES VIA UNE ANALYSE EN COMPOSANTES INDÉPENDANTES

Dinh-Tuan PHAM

Laboratoire de Modélisation et Calcul, IMAG - C.N.R.S.

B.P. 53X, 38041 Grenoble cédex, France

### ABSTRACT

Nous proposons une méthode de séparation aveugle des mélanges instantanés de sources basée sur une analyse en composantes indépendantes (ACI). On cherche une transformation linéaire sur le vecteur d'observations telle que les composantes du nouveau vecteur soient les plus indépendantes. Celles-ci sont alors interprétées comme des sources. Le critère d'indépendance adopté est la distance de Kulback-Leibner entre la loi du vecteur aléatoire étudié et celle qu'aurait ce vecteur si ses composantes sont indépendantes. Notre méthode a l'avantage d'être robuste: même si les observations ne suivent pas tout à fait le modèle de mélange de sources, l'ACI a toujours un sens et peut encore fournir une solution acceptable.

We propose a method for the blind separation of instantaneous mixture of independent sources, based on an independent components analysis (ICA). One looks for a linear transformation on the observation vector such that the components of the new vector are as independent as possible. They will then be interpreted as the sources. As independence criterion, we adopt the Kulback-Leibner distance between the distribution of the random vector under study and that it would have if its components are independent. Our method has the advantage of robustness: even if the observations do not followed exactly the model of mixtures of sources, the ICA is always meaningful and can still yield an acceptable solution.

### 1 Introduction

Les méthodes aveugles de séparation de sources sont toutes basées sur l'indépendance entre les composantes "source" à l'origine des signaux observés. C'est en effet la seule information disponible dans un contexte aveugles. Cependant la plupart de ces méthodes n'exploitent pas complètement cette indépendance: elles se contentent de la nullité de certains cumulants croisés ([1], [2], [4] [5], [7] ...) ou de la moyenne de certaines statistiques non-linéaires ([6], [8], [9], [10], ...) des sources. Dans ce travail, nous proposons d'exploiter cette indépendance dans sa globalité, à l'aide d'une mesure d'indépendance basée sur la distance de Kullback-Leibner. Notre démarche relève de l'analyse en composantes indépendantes (ACI) introduite par Comon [3]. L'idée est de chercher une transformation linéaire sur le vecteur d'observations de sorte que les composantes du nouveau vecteur soient le plus indépendantes. Ce formalisme est légèrement plus général que celui de la séparation de sources. Dans le dernier problème, on suppose disposer de  $K$  séquences d'observations  $X_1(t), \dots, X_K(t)$ , combinaisons linéaires des  $K'$  sources  $S_1(t), \dots, S_{K'}(t)$ , soit  $\mathbf{X}(t) = \mathbf{A}\mathbf{S}(t)$  où  $\mathbf{X}(t)$  et  $\mathbf{S}(t)$  sont les vecteurs de composantes  $X_1(t), \dots, X_K(t)$  et  $S_1(t), \dots, S_{K'}(t)$ . L'objectif de la séparation est de retrouver  $\mathbf{S}(t)$ . En ACI, le modèle de sources n'est

pas postulé a priori, on cherche simplement une matrice  $\mathbf{B}$  telle que les composantes de  $\mathbf{B}\mathbf{X}(t)$  soient les plus indépendantes possibles. Bien sûr, si on suppose ce modèle avec  $K' = K$ , alors  $\mathbf{B} = \mathbf{A}^{-1}$  est la solution de l'ACI, qui de plus est unique modulo une permutation et un changement d'échelle, pourvu que les sources ne sont pas toutes gaussiennes [3]. Par contre, si  $\mathbf{X}(t)$  ne suit pas tout à fait le modèle de sources (à cause d'une légère contamination de bruits par ex.), l'ACI possède toujours un sens et on peut espérer qu'elle fournit encore une solution acceptable, alors qu'une méthode de séparation de sources, conçu pour le cas des "vraies" mélange peut échouer. La robustesse de l'ACI se manifeste en fait dans un cadre plus large: certaines méthodes de séparation de sources fondés sur des équations non linéaires ([6], [10]) peuvent conduire à des solutions étrangères n'ayant rien à voir avec les vraies sources. En cherchant toujours les composantes les plus indépendantes, l'ACI évite de façon naturelle ces solutions.

### 2 ACI basée sur la distance de Kullback-Leibner

On s'intéresse ici à l'ACI théorique où une loi de probabilité  $P_{\mathbf{X}}$  d'un vecteur aléatoire  $\mathbf{X}$  sur  $\mathbb{R}^K$  est donnée et on cherche une matrice  $\mathbf{B}$  telle que les composantes de  $\mathbf{B}\mathbf{X}$  soient les plus indépendantes. Le critère d'indépendance adopté est la distance de Kullback Leibner entre la loi  $P_{\mathbf{B}\mathbf{X}}$  de  $\mathbf{B}\mathbf{X}$  et la loi produit  $P_{\mathbf{B}\mathbf{X}}^I =$



$P_{(\mathbf{B}\mathbf{X})_1} \times \dots \times P_{(\mathbf{B}\mathbf{X})_K}$  des lois marginales  $P_{(\mathbf{B}\mathbf{X})_i}$ , des composantes  $\mathbf{B}\mathbf{X}_i$  de  $\mathbf{B}\mathbf{X}$ . Rappelons que la distance de Kullback-Leibner  $K(P, Q)$  entre deux loi de probabilité  $P$  et  $Q$  est définie comme  $-\int [\ln(dQ/dP)]dP$ . Elle possède la propriété d'être non négative est n'est nulle que si  $P = Q$ . Dans le cas où  $P_{\mathbf{X}}$  admet une densité  $f_{\mathbf{X}}$ , on montre que

$$K(P_{\mathbf{B}\mathbf{X}}, P_{\mathbf{B}\mathbf{X}}^I) = \int_{\mathbb{R}^K} \ln f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} - \ln |\det \mathbf{B}| - \sum_{i=1}^K \int_{\mathbb{R}} [\ln f_{(\mathbf{B}\mathbf{X})_i}(s) f_{(\mathbf{B}\mathbf{X})_i}(s) ds \quad (2.1)$$

où  $f_{\mathbf{B}\mathbf{X}}(\mathbf{s}) = f_{\mathbf{X}}(\mathbf{B}^{-1}\mathbf{s})/|\det \mathbf{B}|$  est la densité de  $\mathbf{B}\mathbf{X}$ . L'ACI consiste donc à minimiser le second membre de (2.1) sans le premier terme constante. C'est un contraste au sens de [3], noté  $C(\mathbf{B})$ .

Pour minimiser  $C$ , il est utile de calculer son gradient. Il s'avère plus commode de travailler avec le gradient relatif, définie comme la matrice  $C'(\mathbf{B})$  telle que la forme linéaire  $\text{Tr}[C'(\mathbf{B})\boldsymbol{\epsilon}]$  approche  $C(\mathbf{B} + \boldsymbol{\epsilon}\mathbf{B}) - C(\mathbf{B})$  avec une erreur d'ordre supérieur à  $\boldsymbol{\epsilon}$  quand  $\boldsymbol{\epsilon} \rightarrow \mathbf{0}$ .

**Proposition 2.1** Soit  $\psi_{(\mathbf{B}\mathbf{X})_i}$ , la dérivée de  $-\ln f_{(\mathbf{B}\mathbf{X})_i}$ . Alors  $C(\mathbf{B} + \boldsymbol{\epsilon}\mathbf{B}) - C(\mathbf{B}) = \text{Tr}[C'(\mathbf{B})^T \boldsymbol{\epsilon}] + o(\boldsymbol{\epsilon})$  où  $o(\boldsymbol{\epsilon})$  désigne un terme tendant vers zéro plus rapidement que  $\boldsymbol{\epsilon}$  quand  $\boldsymbol{\epsilon} \rightarrow \mathbf{0}$  et

$$C'(\mathbf{B}) = \int_{\mathbb{R}^K} \begin{bmatrix} \psi_{(\mathbf{B}\mathbf{X})_1}[(\mathbf{B}\mathbf{x})_1] \\ \vdots \\ \psi_{(\mathbf{B}\mathbf{X})_K}[(\mathbf{B}\mathbf{x})_K] \end{bmatrix} (\mathbf{B}\mathbf{x})^T f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} - \mathbf{I}. \quad (2.2)$$

On peut vérifier que les termes diagonaux de  $C'(\mathbf{B})$  sont nuls, ce qui est attendu car  $C(\mathbf{B})$  ne change pas quand on pré-multiplie  $\mathbf{B}$  avec une matrice diagonale et donc  $\text{Tr}[C'(\mathbf{B})^T \boldsymbol{\epsilon}] = 0$  dès que  $\boldsymbol{\epsilon}$  est diagonale. Il y a donc seulement  $K(K-1)$  équations d'équilibre

$$\int_{\mathbb{R}^K} \psi_{(\mathbf{B}\mathbf{X})_i}[(\mathbf{B}\mathbf{x})_i] (\mathbf{B}\mathbf{x})_j f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = 0, \quad i \neq j = 1, \dots, K, \quad (2.2)$$

pour déterminer, à des facteurs d'échelle près, la matrice minimisant  $C$ .

### 3 Implémentation statistique de l'ACI

#### 3.1 Fonction contraste

Dans le précédent paragraphe on a proposé une ACI théorique où la densité  $f_{\mathbf{X}}$  est donnée. Ce n'est pas le cas en pratique où on dispose seulement d'un échantillon suivant cette densité. La façon la plus naturelle pour résoudre cette difficulté est remplacer la densité inconnue par son estimateur.

Nous adoptons comme méthode d'estimation de densité la méthode du noyau (voir par ex. [11]). À part du fait qu'elle est simple et très bien étudiée, elle a l'avantage de fournir une vraie densité (i.e. positive

et d'intégral un) si on prend pour noyau une densité lui même. Ceci est important car nos calculs précédents ainsi que la définition de la distance de Kullback-Leibner sont basés sur le fait qu'on a affaire à des densités. L'estimateur de noyau de  $f_{\mathbf{X}}$  est donné par ([11])

$$\hat{f}_{\mathbf{X}}(\mathbf{x}) = \frac{1}{T} \sum_{t=1}^T \frac{1}{h_T^K} \kappa\left(\frac{\mathbf{x} - \mathbf{X}(t)}{h_T}\right), \quad \mathbf{x} \in \mathbb{R}^K.$$

où  $\kappa$  est la fonction noyau et  $h_T$  est le paramètre de largeur de bande dépendant de  $T$ . L'ACI (statistique) est maintenant définie comme la minimisation de

$$\hat{C}(\mathbf{B}) = -\ln |\det \mathbf{B}| - \sum_{i=1}^K \int_{\mathbb{R}} [\ln \hat{f}_{(\mathbf{B}\mathbf{X})_i}(s_i)] \hat{f}_{(\mathbf{B}\mathbf{X})_i}(s_i) ds_i, \quad (3.1)$$

où  $\hat{f}_{(\mathbf{B}\mathbf{X})_i}$  est comme  $f_{(\mathbf{B}\mathbf{X})_i}$  avec  $\hat{f}_{\mathbf{X}}$  à la place de  $f_{\mathbf{X}}$ .

La fonction  $-\hat{C}$  ressemble à la fonction log vraisemblance obtenue dans Pham et Garat [10]. Partant du modèle de séparation de sources et l'hypothèse que celles-ci sont blanches avec des densités hypothétiques  $f_i$ , Pham et Garat ont obtenu la fonction log vraisemblance pour les données  $\mathbf{X}(1), \dots, \mathbf{X}(T)$ , de la forme

$$T \left\{ \sum_{i=1}^K \frac{1}{T} \sum_{t=1}^T \ln f_i[(\mathbf{A}^{-1}\mathbf{X})_i(t)] - \ln |\det \mathbf{A}| \right\}.$$

Cette fonction est semblable à  $-\hat{C}(\mathbf{A}^{-1})$  à part le facteur  $T$  et le fait que  $\hat{f}_{(\mathbf{B}\mathbf{X})_i}$  est donnée et non estimée et que l'intégration par rapport à  $\hat{f}_{(\mathbf{B}\mathbf{X})_i}$  est remplacés par une moyenne temporelle.

L'expression pour  $\hat{C}(\mathbf{B})$  fait intervenir les densités marginales de  $\hat{f}_{\mathbf{B}\mathbf{X}}$  qui ne sont pas simple à calculer. Ce sont en fait les densités marginales de  $\mathbf{B}\mathbf{X}^* + h_T \mathbf{B}\mathbf{N}$  où  $\mathbf{X}^*$  est un vecteur aléatoire prenant les valeurs  $\mathbf{X}(1), \dots, \mathbf{X}(T)$  avec probabilité  $1/T$  and  $\mathbf{N}$  est un vecteur aléatoire de densité  $\kappa$ . Leur calcul nécessite donc l'intégration multiple de la fonction  $\kappa(\mathbf{B}^{-1}\mathbf{x})$ . Une façon simple pour éviter cela est de prendre pour  $\kappa$  une densité gaussienne car alors ses densités marginales sont aussi gaussiennes. En prenant pour  $\kappa$  la densité gaussienne centrée de matrice de covariance  $\hat{\Sigma}$ , la  $i$ -ème composante de  $\mathbf{B}\mathbf{N}$  est gaussienne centrée de variance  $\hat{\sigma}_i^2(\mathbf{B})$ , le  $i$ -ème terme diagonal de  $\mathbf{B}\hat{\Sigma}\mathbf{B}^T$ . Alors,

$$\hat{f}_{(\mathbf{B}\mathbf{X})_i}(s) = \frac{1}{T} \sum_{t=1}^T \frac{1}{h_T \hat{\sigma}_i(\mathbf{B})} k\left(\frac{s - (\mathbf{B}\mathbf{X})_i(t)}{h_T \hat{\sigma}_i(\mathbf{B})}\right) \quad (3.2)$$

où  $k(s) = \exp(-s^2/2)/\sqrt{2\pi}$  est la densité gaussienne standard. Pour  $\hat{\Sigma}$ , le choix naturel est la matrice de covariance empirique de  $\mathbf{X}(1), \dots, \mathbf{X}(T)$ . Ce choix assure la cohérence entre la méthode d'estimation et les transformation linéaires, au sens qu'estimer  $f_{\mathbf{X}}$  et en déduire  $\hat{f}_{\mathbf{B}\mathbf{X}}$  et ses densités marginales donne la même chose qu'estimer directement ces dernières par (3.2).

Toutefois, nous ne voulons pas nous restreindre aux densités gaussiennes. On peut préférer une densité à support compact pour réduire le nombre des calculs. Pour éviter l'intégration de  $\kappa$ , on remarque  $\hat{C}(\mathbf{B})$  ne dépend en fait que des densités marginales de  $\mathbf{BX}$ . On peut donc le définir directement par (3.1) avec  $\hat{f}_{(\mathbf{BX})_i}$ , donné par (3.2) avec un certain noyau  $k$  and  $\hat{\sigma}_i(\mathbf{B})$  comme ci-dessus. Comme  $\hat{\sigma}_i(\mathbf{B})$  est inchangé si les  $(\mathbf{BX})_i(t)$  sont soustraits d'une constante ou multipliés par un même facteur, notre contraste est invariant par rapport au changement d'origine et d'échelle de chaque séquence  $(\mathbf{BX})_i(t)$ . Mais la cohérence relative aux transformation linéaires n'est plus assurée: les densités (3.2) pour différentes  $\mathbf{B}$  et  $i$  ne proviennent forcément pas d'une densité multivariariée commune.

### 3.2 Gradient du contraste

Pour écrire le gradient relatif de  $\hat{C}$ , on aura besoin de la notation

$$\hat{E}_{\mathbf{X}|\mathbf{B}\mathbf{X}_i}^f(s) = \frac{1}{T} \sum_{t=1}^T \left\{ \mathbf{X}(t) + \frac{(\hat{\Sigma}\mathbf{B}^T)_i}{\hat{\sigma}_i^2(\mathbf{B})} [s - (\mathbf{BX})_i(t)] \right\} \frac{1}{h_T \hat{\sigma}_i(\mathbf{B})} k\left(\frac{s - (\mathbf{BX})_i(t)}{h_T \hat{\sigma}_i(\mathbf{B})}\right),$$

où  $(\hat{\Sigma}\mathbf{B}^T)_i$  est la  $i$ -ème ligne de  $\mathbf{B}\hat{\Sigma}$ . Cette quantité s'interprète comme une estimation de l'espérance conditionnelle de  $\mathbf{X}$  sachant  $(\mathbf{BX})_i$  fois la densité  $f_{(\mathbf{BX})_i}$ . On note aussi  $\hat{E}_{\mathbf{X}^T\mathbf{B}^T|\mathbf{B}\mathbf{X}_i}^f$  pour la transposée de  $\hat{E}_{\mathbf{X}|\mathbf{B}\mathbf{X}_i}^f$ , post-multiplié par  $\mathbf{B}^T$ .

**Proposition 3.1** Avec la notation précédente et supposons que  $k$  est continûment dérivable et soit  $\hat{\psi}_{(\mathbf{BX})_i}$  la dérivée de  $-\ln \hat{f}_{(\mathbf{BX})_i}$ , alors  $\hat{C}(\mathbf{B} + \epsilon\mathbf{B}) - \hat{C}(\mathbf{B}) = \text{Tr}[\hat{C}'(\mathbf{B})^T \epsilon] + o(\epsilon)$  où  $o(\epsilon)$  désigne un terme tendant vers zéro plus vite que  $\epsilon$  quand  $\epsilon \rightarrow \mathbf{0}$  et

$$\hat{C}'(\mathbf{B}) = \begin{bmatrix} \int_{\mathbb{R}} \hat{\psi}_{(\mathbf{BX})_1}(s) \hat{E}_{\mathbf{X}^T\mathbf{B}^T|\mathbf{B}\mathbf{X}_1}^f(s) ds \\ \vdots \\ \int_{\mathbb{R}} \hat{\psi}_{(\mathbf{BX})_K}(s) \hat{E}_{\mathbf{X}^T\mathbf{B}^T|\mathbf{B}\mathbf{X}_K}^f(s) ds \end{bmatrix} - \mathbf{I}.$$

Un calcul direct montre le terme  $(i, j)$  de  $\hat{C}'(\mathbf{B})$  s'écrit aussi:

$$\frac{1}{T} \sum_{t=1}^T \hat{\psi}_{(\mathbf{BX})_i}[(\mathbf{BX})_i(t)] (\mathbf{BX})_j(t) + \frac{(\mathbf{B}\hat{\Sigma}\mathbf{B}^T)_{ij}}{\hat{\sigma}_i^2(\mathbf{B})} \hat{\lambda}_{(\mathbf{BX})_i}, \quad (3.3)$$

où

$$\begin{aligned} \hat{\psi}_{(\mathbf{BX})_i}(v) &= \int_{\mathbb{R}} \hat{\psi}_{(\mathbf{BX})_i}[v + h_T \hat{\sigma}_i(\mathbf{B})u] k(u) du \\ \hat{\lambda}_{(\mathbf{BX})_i} &= - \int_{\mathbb{R}} \hat{\psi}_{(\mathbf{BX})_i}(s) \frac{1}{T} \sum_{t=1}^T \tilde{k}'\left[\frac{s - (\mathbf{BX})_i(t)}{h_T \hat{\sigma}_i(\mathbf{B})}\right] ds \\ &= -h_T^2 \sigma_i^2(\mathbf{B}) \int_{\mathbb{R}} \hat{\psi}_{(\mathbf{BX})_i}(s) \tilde{f}'(s) ds, \end{aligned}$$

$\tilde{f}$  étant définie comme  $\hat{f}$  avec  $\tilde{k}(x) = -\int_{-\infty}^x uk(u) du$  à la place de  $k$ . (si  $k$  est gaussienne,  $\tilde{k} = k$  et  $\tilde{f} = f$ ).

On note que (3.3) ne change pas quand on centre les données. Dans le cas où les données sont centrées, (3.3) s'écrit encore

$$\frac{1}{T} \sum_{t=1}^T \left\{ \hat{\psi}_{(\mathbf{BX})_i}[(\mathbf{BX})_i(t)] + \frac{\hat{\lambda}_{(\mathbf{BX})_i}}{\hat{\sigma}_i^2(\mathbf{B})} (\mathbf{BX})_i(t) \right\} (\mathbf{BX})_j(t). \quad (3.4)$$

Annulant (3.3) ou (3.4) conduit à un système de  $K(K-1)$  équations d'équilibre pour déterminer  $\mathbf{B}$  à des facteurs d'échelle près. On peut noter la similitude entre ce système avec celui obtenu par Pham et Garat [10] où la seule différence est que les fonctions séparatrices sont ici  $\hat{\psi}_{(\mathbf{BX})_i}(s) + [\hat{\lambda}_{(\mathbf{BX})_i}/\hat{\sigma}_i^2(\mathbf{B})]s$  qui dépendent des données et ne sont plus données a priori. Comme on peut voir,  $\hat{\psi}_{(\mathbf{BX})_i}$  n'est d'autre qu'une version lissée de  $\hat{\psi}_{(\mathbf{BX})_i}$ , qui estime la fonction  $\psi_{(\mathbf{BX})_i}$ , que l'on sait réalisant la séparation optimale. Le terme  $\hat{\lambda}_{(\mathbf{BX})_i}$  dans (3.3) provient du lissage et a pour effet l'inclusion des statistiques au second ordre dans les équations de séparation.

*Note:* On sait que  $h_T$  doit tend vers zéro suffisamment lente quand  $T \rightarrow \infty$ , pour une estimation convergence de densité [11]. En fait, notre méthode reste valide même si  $h_T$  ne converge pas vers zéro mais vers une limite non nulle  $h$ . Pour simplifier, restons nous le cas où  $k$  est une densité gaussienne. Alors,  $\hat{f}_{\mathbf{X}}$  converge vers  $f_{\mathbf{X}}^*$ , la densité de  $\mathbf{X} + h\mathbf{N}$  où  $\mathbf{N}$  est un vecteur gaussien centré indépendant de  $\mathbf{X}$  de même matrice de covariance. On peut alors voir que l'ACI statistique revient, à la limite, à utiliser  $K(P_{\mathbf{B}\mathbf{X}+h\mathbf{B}\mathbf{N}}, P_{\mathbf{B}\mathbf{X}+h\mathbf{B}\mathbf{N}}^I)$  au lieu de  $K(P_{\mathbf{B}\mathbf{X}}, P_{\mathbf{B}\mathbf{X}}^I)$  comme critère d'indépendance. Or, rendant les composantes de  $\mathbf{B}\mathbf{X}$  indépendantes le rend aussi celles de  $\mathbf{B}\mathbf{N}$  et inversement. Le seul effet de ce nouveau critère est de donner plus importance aux statistiques du second ordres.

## 4 Quelques résultats de simulation

Nous donnons ci-dessous quelques résultats de simulations illustrant la bonne performance de notre méthode. Nous considérons la situation de séparation de source avec deux sources (de longueur 200) mélangées selon la matrice

$$\mathbf{A} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$$

et dont les lois sont décrites dans la table ci-dessous

Cas	Source 1	Source 2
1	$\mathcal{U}(-.5, .5)$	$\mathcal{U}(-.5, .5)$
2	$\frac{1}{2}\mathcal{N}(-2, .64) + \frac{1}{2}\mathcal{N}(2, .64)$	$\mathcal{U}(-1, 5, 1.5)$

où  $\mathcal{U}(-a, a)$  désigne la loi uniforme sur  $[-a, a]$ ,  $\mathcal{N}(m, \sigma^2)$  la loi gaussienne de moyenne  $m$  et variance  $\sigma^2$  et  $\frac{1}{2}\mathcal{N}(-m, \sigma^2) + \frac{1}{2}\mathcal{N}(m, \sigma^2)$  la loi "gaussienne mixte"



ayant pour densité la moyenne des densités gaussienne de moyenne  $-m$  et  $m$  et variance  $\sigma^2$ .

Les résultats sont illustrés graphiquement par les figures 1 et 2. Les observations sont représentées par des points dans un repère cartésien. La transformation  $\mathbf{X} \mapsto \mathbf{BX}$  est vue comme un changement de repère et donc les sources reconstruites se lisent comme les coordonnées dans un nouveau repère, dont les axes sont parallèles aux colonnes de  $\hat{\mathbf{B}}^{-1}$ . Ceux-ci sont représentés en solide sur les figures. Il ne sont pas orthogonaux, contrairement aux axes principaux représentés en pointillés. Ces figures montrent la bonne séparation obtenue: les axes de l'ACI reproduisent assez bien les axes théoriques et on voit apparaître le contour de la loi produit des deux sources.

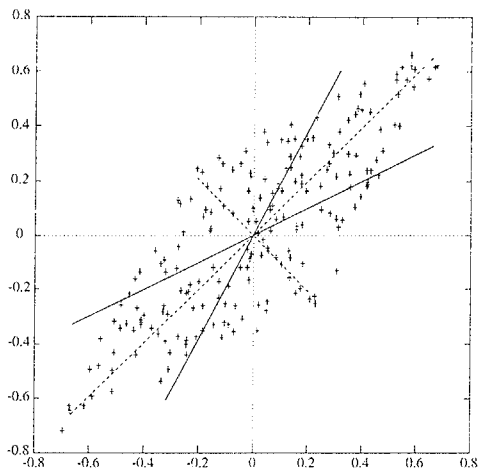


Figure 3: ACI de deux loi uniformes

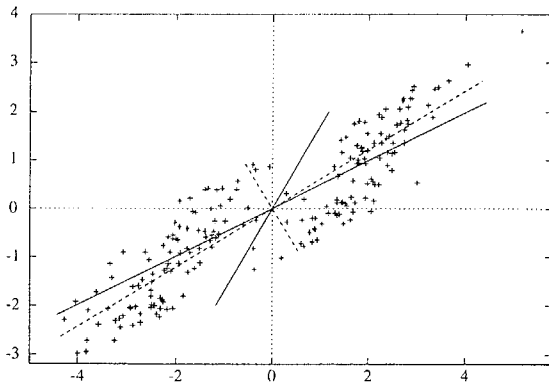


Figure 4: ACI d'une loi uniforme et une loi "gaussienne mixte"

Comme mesure de performance d'une méthode de séparation de source, nous adoptons les coefficients de contamination  $c_{ij} = (\hat{\mathbf{B}}\mathbf{A})_{ij}\sigma_i / [(\hat{\mathbf{B}}\mathbf{A})_{ii}\sigma_j]$ . Ils représentent la contribution relative contribution de la  $j$ -ème source à la  $i$ -ème source reconstruites, toutes sources étant normalisées pour avoir la même énergie. Les coefficients de contamination sont donnés dans la table ci-dessous. Pour être complet, les coefficients de contamination  $c'_{12}$ ,  $c'_{21}$  correspondant à l'analyse en composantes principales sont aussi listés.

Cas	$c_{12}$	$c_{21}$	$c'_{12}$	$c'_{21}$
1	-0.0136	-0.0187	0.9866	-0.9866
2	-0.0530	-0.0186	0.2429	-1.5028

Sur le plan calcul, l'algorithme converge très rapidement. On peut aussi utiliser une forme discrétisée du contraste (dont nous ne donnons de détails ici, faute de place) qui réduit encore le coût des calculs.

## 5 Conclusion

Nous avons introduit une nouvelle méthode de séparation de sources, basée sur une analyse en composantes indépendantes, qui s'adapte automatiquement aux différentes formes de la densité des sources. Nos résultats de simulations semblent indiquer que cette méthode est performante, robuste et ne souffre pas du problème de solutions étrangères mentionnés dans l'introduction.

## RÉFÉRENCES

- [1] Cardoso, J.P. "Source separation using higher order moments". *proc. ICASSP*, 1989, 2109-2112.
- [2] Cardoso, J.P. "Iterative technique for blind source separation using only fourth order cumulants". *Signal Processing VI: Theory and Application (Proc. EUSIPCO 92)*, J. Vandewalle, R. Boite, M. Moonen & A. Oosterlink eds, 1992, 739-742.
- [3] Comon, P. "Independent components analysis, a new concept". *Signal Processing* **36**, 3, 1994, 287-314.
- [4] Duvaut, P. "Principe des méthodes de séparation de sources fondées sur les moments d'ordre supérieur". *Traitement du Signal*, **7**, 5 (numéro spécial non linéaire non gaussien), 1990, 407-418.
- [5] Gaeta, M., Lacoume, J. L. "Source separation without a priori knowledge: the maximum likelihood approach". *Signal Processing V (Proc. EUSIPCO 90)*, L. Tores, E. Mas-Grau and M. A. Lagunas eds, 1990, 621-624.
- [6] Jutten, C. Herault, J. "Blind separation of sources, Part I: an adaptive algorithm based on neuromimetic structure". *Signal Processing* **24**, 1991, 1-10.
- [7] Lacoume, J. L. "Sources identification: a solution based on the cumulants". *IEEE workshop on spectrum analysis and modeling*. August 1988, Mineapolis.
- [8] Loubaton, Ph., Delfosse, N. "Séparation adaptative de sources indépendantes par une approche de déflation". Dans *XIV Colloque sur le Traitement du Signal et des Images (GRETSI)*, Juan-Les-Pins, Sept. 1993, 325-328
- [9] Moreau, E., Macchi, O. "New self-adaptive algorithm for sources separation based on contrast functions". *Proc. IEEE SP Workshop on Higher-Order-Stat.*, Lake Tahoe, U.S.A., 1993, 215-219.
- [10] Pham, D. T., Garat, Ph., Jutten, C. "Separation of a mixture of independent sources through a maximum likelihood approach". Dans *Signal Processing VI: Theory and Application (Proc. EUSIPCO 92)*, J. Vandewalle, R. Boite, M. Moonen & A. Oosterlink eds, 1992, 771-774.
- [11] Silverman, B. W. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall: London & New-York, 1986.