



# Séparation Aveugle de Sources Temporellement Corrélées

D.T. PHAM

P. GARAT

Laboratoire de Modélisation et Calcul, I.M.A.G.  
B.P. 53X 38041 Grenoble Cedex

## RÉSUMÉ

Nous proposons une méthode de séparation aveugle des mélanges instantanés de sources indépendantes, chacune d'elles étant temporellement corrélée. L'approche du maximum de vraisemblance conduit à des équations de séparation où interviennent des filtres linéaires. Nous montrons l'existence de filtres optimaux, fonction des densités spectrales (supposées distinctes) des sources. La résolution des équations de séparation est faite selon une procédure itérative dans le cas général, ou par une simple double diagonalisation dans certains cas particuliers. Ces procédures sont testées sur signaux de parole.

## 1 Introduction

Le principe des techniques actuelles de séparation de sources est l'*indépendance mutuelle* entre les composantes "source" à l'origine des signaux observés. Les méthodes habituelles pour les mélanges instantanés ([1], [2], [4], [5] ...) exploitent seulement l'indépendance entre données observées en un même instant et non en des instants différents. Cette démarche est justifiée en séparation "aveugle" si les sources sont blanches, mais cette situation est rare. D'autre part, la performance de telles procédures de séparation se dégrade fortement quand les densités de probabilité se rapprochent de densités gaussiennes [7].

Notre suggestion est que la structure spectrale des signaux source peut être utile à améliorer la séparation (voire à la rendre possible dans le cas purement gaussien). Notre approche porte, par principe, sur l'exploitation exclusive de l'information spectrale des signaux source et conduit à une méthode de séparation applicable à des mélanges instantanés de signaux indépendants et individuellement temporellement corrélés, à condition que leurs densités spectrales soient distinctes. En d'autres termes, nous recherchons la non-corrélation des sources avec décalages temporels.

Du point de vue de la complexité algorithmique, la méthode développée fait intervenir des filtrages linéaires des signaux et des calculs de covariance. Elle se distingue à ce titre des techniques de séparation qui font usage d'opérateurs non-linéaires pour l'évaluation de moments croisés d'ordre supérieur à deux [1], [6], ou de fonctions "contraste" [3].

Nous présentons dans les sections qui suivent une justification de notre méthode par une démarche du maximum de vraisemblance, puis une étude sur son optimalité et ses performances pratiques sur signaux de parole.

## ABSTRACT

We present a method for the blind separation of instantaneous mixture of independent sources, each of them being time-correlated. Through a maximum likelihood approach, we derive the separating equations which involve linear filters. We prove the existence of optimal filters, function of the spectral densities (assumed to be distinct) of the sources. To solve the separating equations we develop a general iterative method and for certain particular cases a simple double-diagonalisation method. Both procedures are tested on speech signals.

## 2 Construction de la méthode

Soit  $\{\mathbf{X}(t)\}_{t=0}^{T-1}$  une séquence de  $T$  vecteurs "observation" (de dimension  $K$ ). Par hypothèse, ces derniers sont issus de signaux "source" inconnus  $\{\mathbf{S}(t)\}_{t=0}^{T-1}$  selon une transformation linéaire instantanée :  $\mathbf{X}(t) = \mathbf{A}\mathbf{S}(t)$ , où  $\mathbf{A}$  désigne la matrice de mélange (inconnue de rang plein) et  $\mathbf{S}(t)$  le vecteur source à l'instant  $t$  (de dimension  $K' = K$ ). La suite  $\{\mathbf{S}(t)\}_{t \in \mathbb{Z}}$  est vue comme un processus vectoriel aléatoire de loi inconnue, à valeurs réelles, centré, stationnaire au second ordre et à composantes  $\{S_i(t)\}_{t \in \mathbb{Z}}$  mutuellement indépendantes ( $i = 1, \dots, K$ ). De plus, les densités spectrales  $\{g_i(\lambda)\}_{\lambda \in \mathbb{R}}$  des  $S_i$  sont supposées distinctes.

L'estimation de  $S_i(t)$  se fait à l'aide d'une matrice  $\mathbf{C}$  selon l'équation de restitution:  $\hat{\mathbf{S}}(t) = \mathbf{C}\mathbf{X}(t)$ ,  $t = 0, \dots, (T-1)$ . Formellement,  $\mathbf{C}$  est identifiable à une matrice de la forme  $\mathbf{D}\mathbf{P}\mathbf{A}^{-1}$  où  $\mathbf{D}$  est une matrice diagonale de facteurs d'échelle,  $\mathbf{P}$  une matrice permutation et  $\mathbf{A}$  une estimation de  $\mathbf{A}$ .

La transformée de Fourier discrète (TFD) de  $\{\mathbf{S}(t)\}_{t=0}^{T-1}$  peut s'écrire :  $\mathbf{d}_S(n) = \frac{1}{\sqrt{T}} \sum_{t=0}^{T-1} e^{-j2\pi nt/T} \mathbf{S}(t)$ ,  $\forall n \in \mathbb{Z}$ . Les  $\mathbf{S}(t)$  étant réels, les  $\mathbf{d}_S(n)$  sont aussi réels et satisfont la symétrie hermitienne  $\mathbf{d}_S(n) = \mathbf{d}_S^*(T-n)$  où le signe \* désigne la conjugaison. Nous limitant aux seuls vecteurs  $\mathbf{d}_S(0)$  à  $\mathbf{d}_S(T/2)$  (nous ne restreignons pas la généralité en supposant que  $T$  est pair), nous avons la propriété bien connue suivante:

*Asymptotiquement, les  $\mathbf{d}_S(n)$  ( $n = 0, \dots, T/2$ ) sont des vecteurs indépendants, centrés, gaussiens et de matrice de covariance diagonale  $\mathbf{D}_g(n/T)$  d'éléments diagonaux  $g_1(n/T), \dots, g_K(n/T)$ .*

Les données observées  $\{\mathbf{X}(t)\}_{t=0}^{T-1}$  peuvent être transformées



par TFD en  $\{\mathbf{d}_X(n)\}_{n=0}^{T/2}$  satisfaisant la relation  $\mathbf{d}_S(n) = \mathbf{A}^{-1}\mathbf{d}_X(n)$ . Nous exprimons ensuite le logarithme de la densité de probabilité conjointe des  $\{\mathbf{d}_X(n)\}_{n=0}^{T/2}$  en fonction des densités approximativement gaussiennes des  $\mathbf{d}_S(n)$  pour  $T$  suffisamment grand. A une constante près, cela s'écrit :

$$L_T = -\frac{1}{2} \sum_{i=1}^K \sum_{n=0}^{T/2} \frac{|\mathbf{e}_i^T \mathbf{A}^{-1} \mathbf{d}_X(n)|^2}{g_i(n/T)} + \ln g_i(n/T) - T \ln |\det \mathbf{A}| \quad (1)$$

où  $\mathbf{e}_i^T$  désigne la  $i$ -ème ligne de la matrice unité. L'expression (1) s'interprète aussi comme le logarithme de la densité de probabilité des observations brutes  $\{\mathbf{X}(t)\}_{t=0}^{T-1}$ , car la TFD est un opérateur linéaire de Jacobien unité. C'est donc la "log-vraisemblance" des données, fonction du paramètre inconnu  $\mathbf{A}$ . Ici, les densités  $g_i$  sont supposées connues. Cette hypothèse de travail pourrait être affaiblie en supposant les  $g_i$  connues à un facteur d'échelle près, mais cela n'introduirait dans l'analyse que des paramètres d'échelle, que l'on sait inutiles dans le cadre de la séparation de sources.

Pour maximiser  $L_T$  par rapport aux coefficients de  $\mathbf{A}^{-1}$ , on est amené à calculer sa différentielle :

$$dL_T = T \text{tr}(d\mathbf{A}^{-1} \mathbf{A}) - \sum_{i=1}^K \sum_{n=0}^{T-1} \frac{\mathbf{e}_i^T d\mathbf{A}^{-1} \mathbf{d}_X(n) \mathbf{e}_i^T \mathbf{A}^{-1} \mathbf{d}_X^*(n)}{g_i(n/T)} \quad (2)$$

A l'aide de la matrice (diagonale)  $\mathbf{D}_g(n/T)$  on obtient :

$$\frac{dL_T}{T} = \text{tr} \left\{ d\mathbf{A}^{-1} \left( \mathbf{A} - \sum_{n=0}^{T-1} \mathbf{d}_X(n) \mathbf{D}_g^{-1} \left( \frac{n}{T} \right) \mathbf{d}_X^*(n) \mathbf{A}^{-T} \right) \right\} \quad (3)$$

où les symboles  $\dagger$  et  $-T$  signifient respectivement "transposée-conjuguée" et "inverse-transposée".

L'estimation  $\hat{\mathbf{A}}^{-1}$  par maximum de vraisemblance conduit à annuler le terme entre parenthèses dans (3). Cela donne le système d'équations :

$$\sum_{n=0}^{T-1} \frac{\mathbf{e}_i^T \hat{\mathbf{A}}^{-1} \mathbf{d}_X(n) \mathbf{e}_j^T \hat{\mathbf{A}}^{-1} \mathbf{d}_X^*(n)}{g_i(n/T)} = 0 \quad i \neq j = 1, \dots, K. \quad (4)$$

auquel s'ajoutent  $K$  autres équations de normalisation de  $\hat{\mathbf{A}}^{-1}$  inintéressantes ici. Pour interpréter les équations précédentes, on passe du domaine fréquentiel au domaine temporel en invoquant le théorème de Parseval. On introduit alors des filtres  $\phi_i = \{\phi_{i,k}\}_{k \in \mathbb{Z}}$  de réponse en fréquence  $g_i^{-1}(k/T)$ , pour une ré-écriture de (4) en :

$$\sum_{t=0}^{T-1} \mathbf{e}_j^T \hat{\mathbf{A}}^{-1} \mathbf{X}(t) \left( \phi_i \star \mathbf{e}_i^T \hat{\mathbf{A}}^{-1} \mathbf{X} \right) (t) = 0 \quad i \neq j = 1, \dots, K \quad (5)$$

où le symbole  $\star$  désigne la convolution.

Les équations précédentes fournissent en clair une méthode aveugle de séparation. Le terme  $\mathbf{e}_i^T \hat{\mathbf{A}}^{-1} \mathbf{X}$  s'identifiant bien sûr à la restitution de la source  $S_i$ , (5) exprime le fait que chacune des restitutions, transformée par un filtre linéaire adapté (blanchiment spectral), doit être empiriquement décorrélée avec les autres (non-corrélation par paire). La moyenne temporelle sur l'intervalle fini d'observation  $\frac{1}{T} \sum_{t=0}^{T-1}$  approxime l'opérateur *espérance mathématique*  $\mathbf{E}$  et converge vers celui-ci lorsque la longueur  $T$  de l'intervalle croît à l'infini (principe d'ergodicité). Le vrai critère de

séparation étant la non-corrélation (avec des retards temporels), le blanchiment spectral est là pour assurer une asymétrie des équations par rapport aux indices  $i$  et  $j$  des sources, ainsi qu'une certaine optimalité de la méthode (voir section suivante).

En développant la convolution dans (5), nous pouvons exprimer différemment les équations de séparation :

$$\sum_{k=1-T}^{T-1} \phi_{i,k} \mathbf{e}_i^T \hat{\mathbf{A}}^{-1} \mathbf{R}_X(k) \hat{\mathbf{A}}^{-T} \mathbf{e}_j = 0 \quad i \neq j = 1, \dots, K \quad (6)$$

où  $\mathbf{R}_X(\cdot)$  désigne la covariance empirique des observations :

$$\mathbf{R}_X(k) = \frac{1}{T} \sum_{t=\max(0,k)}^{\min(T-1,T+k-1)} \mathbf{X}(t) \mathbf{X}^T(t-k).$$

En pratique, les densités spectrales des sources sont inconnues. Par conséquent, les filtres  $\phi_i$  le sont aussi et doivent être choisis arbitrairement. Etant donné que la densité spectrale d'un processus réel est une fonction paire (et positive), nous pouvons restreindre le choix des filtres parmi les filtres symétriques ( $\phi_{i,k} = \phi_{i,-k}$ ) à réponse en fréquence positive. (De tels filtres peuvent être construits par la convolution d'un filtre causal avec son filtre "miroir"). En se restreignant encore aux filtres à réponse impulsionnelle finie (R.I.F.) d'ordre  $p$ , seules les matrices  $\mathbf{R}_X(0), \dots, \mathbf{R}_X(p)$  interviennent dans (6) et les équations deviennent :

$$\sum_{k=0}^p \phi_{i,k} \mathbf{e}_i^T \hat{\mathbf{A}}^{-1} \mathbf{R}_X^c(k) \hat{\mathbf{A}}^{-T} \mathbf{e}_j = 0 \quad i \neq j = 1, \dots, K. \quad (7)$$

où  $\mathbf{R}_X^c(k) = \mathbf{R}_X(k) + \mathbf{R}_X(-k)^T$  pour  $k \neq 0$ ,  $= \mathbf{R}_X(0)$  sinon.

### 3 Optimalité asymptotique de la méthode

Dorénavant,  $g_i^*$  désignera la vraie densité spectrale de la source  $S_i$  tandis que  $h_i^{-1}$  dénotera la réponse en fréquence du  $i$ -ème filtre arbitraire choisi  $\phi_i$ . L'opérateur de moyenne temporelle  $\frac{1}{T} \sum_{t=0}^{T-1}$  sera noté  $\hat{\mathbf{E}}_T$  pour simplifier.

Pour déduire les propriétés asymptotiques de l'estimateur  $\hat{\mathbf{A}}$ , nous effectuons un développement de Taylor au premier ordre de (5) au voisinage de  $\mathbf{A}$ . C'est-à-dire  $\hat{\mathbf{A}} = \mathbf{A}(\mathbf{I} + \mathcal{E})$  où  $\mathcal{E}$  est une "petite" matrice (de terme général  $\epsilon_{ij}$ ) à caractériser. Ainsi,  $\mathbf{e}_i^T \hat{\mathbf{A}}^{-1} \mathbf{X} \approx S_i(t) - \sum_{k=0}^K S_k(t) \epsilon_{ik}$ . A présent, le terme de gauche dans (5) se ré-écrit comme suit :  $T \hat{\mathbf{E}}_T [S_j(\phi_i \star S_i)] - T \sum_{k=1}^K \left\{ \epsilon_{jk} \hat{\mathbf{E}}_T [S_k(\phi_i \star S_i)] + \epsilon_{ik} \hat{\mathbf{E}}_T [S_j(\phi_i \star S_k)] \right\}$ . Par ergodicité,  $\hat{\mathbf{E}}_T [S_j(\phi_i \star S_k)] \rightarrow \mathbf{E} [S_j(t)(\phi_i \star S_k)(t)]$  qui vaut 0 si  $j \neq k$  et  $\int_{-1/2}^{1/2} g_j^*(\nu) h_i^{-1}(\nu) d\nu$  dans le cas contraire. En approximant les différents termes par leur valeur limite, le développement de Taylor s'écrit donc :

$$\hat{\mathbf{E}}_T [S_j(\phi_i \star S_i)] \approx$$

$$\epsilon_{ji} \int_{-1/2}^{1/2} g_i^*(\nu) h_i^{-1}(\nu) d\nu + \epsilon_{ij} \int_{-1/2}^{1/2} g_j^*(\nu) h_i^{-1}(\nu) d\nu$$

pour  $i \neq j = 1, \dots, K$ . L'ensemble des termes  $\{\epsilon_{ij}\}_{i \neq j}$  est assimilé à un vecteur  $\mathbf{\Delta} = (\epsilon_{12} \epsilon_{21} \dots)^T$  de taille  $K(K-1)$ .

De même  $\{S_j(t)(\phi_i \star S_i)(t)\}_{i \neq j} = \{\Psi_{ij}(t)\}_{i \neq j}$  est assimilé à  $\Psi(t)$ . Ainsi le système d'équations précédent devient:

$$\Delta \approx \mathbf{H}^{-1} \hat{\mathbf{E}}_T(\Psi) \quad (8)$$

où  $\mathbf{H}$  est la matrice bloc-diagonale dont les blocs sont:

$$\mathbf{H}_{(ij)} = \begin{bmatrix} h_{ij} & h_{ii} \\ h_{jj} & h_{ji} \end{bmatrix} \quad i \neq j, \quad (9)$$

avec  $h_{ij} = \int_{-1/2}^{1/2} g_j^*(\nu) h_i^{-1}(\nu) d\nu$ . On note que la condition d'inversibilité des  $\mathbf{H}_{(ij)}$  s'écrit  $h_{ii} h_{jj} \neq h_{ij} h_{ji}, \forall i \neq j$ . Ce qui impose que les  $h_i$  soient toutes distinctes ainsi que les  $g_i^*$ .

Le Théorème Central Limite est applicable pour le processus stationnaire vectoriel  $\Psi(t)$ : sous certaines conditions raisonnables  $\sqrt{T} \hat{\mathbf{E}}_T(\Psi)$  converge en loi, lorsque  $T \rightarrow \infty$ , vers un vecteur gaussien centré de matrice de covariance  $\mathbf{G}$  de terme général  $\lim_{T \rightarrow \infty} \mathbf{E} [\hat{\mathbf{E}}_T(\Psi_{ij} \Psi_{kl})]$ ,  $i \neq j, k \neq l$ . L'indépendance mutuelle des  $S_i$  induit en outre pour  $\mathbf{G}$  une structure bloc-diagonale. Donc, d'après (8) le vecteur  $\Delta$  est asymptotiquement gaussien centré et de matrice de covariance  $\frac{1}{T} \mathbf{H}^{-1} \mathbf{G} \mathbf{H}^{-T}$  également bloc-diagonale. On en tire une propriété asymptotique pour l'ensemble des vecteurs  $(\epsilon_{ij} \epsilon_{ji})^T$ : ils sont mutuellement indépendants, centrés, gaussiens de matrice de covariance  $\frac{1}{T} \mathbf{H}_{(ij)}^{-1} \mathbf{G}_{(ij)} \mathbf{H}_{(ij)}^{-T}$ .

La preuve de l'optimalité de la méthode de séparation revient maintenant à montrer que les matrices  $\mathbf{H}_{(ij)}^{-1} \mathbf{G}_{(ij)} \mathbf{H}_{(ij)}^{-T}$  sont rendues minimales lorsque les  $h_i$  sont choisies proportionnelles aux  $g_i^*$ . Dans ce but, nous poursuivons le calcul des blocs diagonaux de  $\mathbf{G}$ . Cela requiert les égalités suivantes:

$$\sum_k \mathbf{E} [S_i(t) S_i(t+k)] \mathbf{E} [(\phi_j \star S_j)(t) (\phi_j \star S_j)(t+k)] = \int_{-1/2}^{1/2} g_i^*(\nu) g_j^*(\nu) h_j^{-2}(\nu) d\nu.$$

$$\sum_k \mathbf{E} [S_i(t) (\phi_i \star S_i)(t+k)] \mathbf{E} [S_j(t) (\phi_j \star S_j)(t+k)] = \int_{-1/2}^{1/2} g_i^*(\nu) g_j^*(\nu) h_i^{-1}(\nu) h_j^{-1}(\nu) d\nu.$$

Le  $(ij)$ -ème bloc sur la diagonale de  $\mathbf{G}$  peut être écrit matriciellement comme suit:

$$\mathbf{G}_{(ij)} = \int_{-1/2}^{1/2} \begin{bmatrix} h_i^{-1}(\nu) \\ h_j^{-1}(\nu) \end{bmatrix} [h_i^{-1}(\nu) \ h_j^{-1}(\nu)] g_i^*(\nu) g_j^*(\nu) d\nu. \quad (10)$$

L'idée est alors d'écrire  $\mathbf{H}_{(ij)}$  selon le même formalisme. De (9), il vient:

$$\mathbf{H}_{(ij)} = \int_{-1/2}^{1/2} \begin{bmatrix} h_i^{-1}(\nu) \\ h_j^{-1}(\nu) \end{bmatrix} [g_i^{*-1}(\nu) \ g_j^{*-1}(\nu)] g_i^*(\nu) g_j^*(\nu) d\nu. \quad (11)$$

La similitude des écritures (10) et (11) nous amène à l'identification matricielle remarquable:

$$\begin{bmatrix} \mathbf{G}_{(ij)} & \mathbf{H}_{(ij)} \\ \mathbf{H}_{(ij)}^T & \mathbf{J}_{(ij)} \end{bmatrix} = \int \begin{bmatrix} h_i^{-1} \\ h_j^{-1} \\ g_i^{*-1} \\ g_j^{*-1} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ h_i & h_j & g_i^* & g_j^* \end{bmatrix} g_i^* g_j^* d\nu \quad (12)$$

dans laquelle nous avons omis les bornes d'intégration et les arguments  $\nu$  pour alléger l'écriture. Le bloc  $\mathbf{J}_{(ij)}$  s'explique comme suit:

$$\mathbf{J}_{(ij)} = \begin{bmatrix} \int g_j^* g_i^{*-1} d\nu & 1 \\ 1 & \int g_i^* g_j^{*-1} d\nu \end{bmatrix}. \quad (13)$$

Nous reconnaissons dans le membre de droite de (12) une matrice semi définie positive; cela entraîne  $\mathbf{H}_{(ij)}^{-1} \mathbf{G}_{(ij)} \mathbf{H}_{(ij)}^{-T} \geq \mathbf{J}_{(ij)}$  à condition que l'inversibilité de  $\mathbf{H}_{(ij)}$  (déjà évoquée ci-dessus) soit assurée. En terme de covariance des  $(\epsilon_{ij} \epsilon_{ji})^T$ , la borne minimale  $\frac{1}{T} \mathbf{J}_{(ij)}$  est atteinte lorsque  $h_i \propto g_i^*$ . Ce que nous voulions démontrer.

## 4 Résolution des équations de séparation

### 4.1 Algorithme itératif général

Pour mettre en œuvre une résolution générale des équations de séparation (5), on procède à un développement de Taylor des équations au voisinage d'une estimation initiale  $\tilde{\mathbf{A}}$  de  $\mathbf{A}$ . Cette estimation est supposée proche de  $\mathbf{A}$  et on écrit  $\tilde{\mathbf{A}} = \mathbf{A}(\mathbf{I} + \mathcal{E})$ . Selon une démarche semblable à celle de la section précédente, on obtient  $\mathcal{E}$  par résolution d'un système d'équations linéaires découplées en paires d'équations pour les couples d'inconnues  $(\epsilon_{ij}, \epsilon_{ji}), i \neq j$ . La résolution donne:

$$\epsilon_{ij} = \frac{\begin{vmatrix} \hat{\mathbf{E}}_T[\tilde{S}_i(\phi_j \star \tilde{S}_i)] & \hat{\mathbf{E}}_T[\tilde{S}_i(\phi_i \star \tilde{S}_i)] \\ \hat{\mathbf{E}}_T[\tilde{S}_i(\phi_j \star \tilde{S}_j)] & \hat{\mathbf{E}}_T[\tilde{S}_j(\phi_i \star \tilde{S}_i)] \end{vmatrix}}{\begin{vmatrix} \hat{\mathbf{E}}_T[\tilde{S}_i(\phi_j \star \tilde{S}_i)] & \hat{\mathbf{E}}_T[\tilde{S}_i(\phi_i \star \tilde{S}_i)] \\ \hat{\mathbf{E}}_T[\tilde{S}_j(\phi_j \star \tilde{S}_j)] & \hat{\mathbf{E}}_T[\tilde{S}_j(\phi_i \star \tilde{S}_j)] \end{vmatrix}}, \quad i \neq j,$$

avec  $\tilde{S}_k = \mathbf{e}_k^T \tilde{\mathbf{A}}^{-1} \mathbf{X}$ . Les termes diagonaux  $\epsilon_{ii}$  sont pris égaux à zéro. Ainsi, à partir de  $\mathcal{E}$ , on déduit la nouvelle estimation  $\tilde{\mathbf{A}}(\mathbf{I} + \mathcal{E})$  de  $\mathbf{A}$ , puis on itère le procédé jusqu'à la convergence. Il est souhaitable à chaque étape de normaliser  $\tilde{\mathbf{A}}^{-1}$  selon une méthode de son choix. On note que cette résolution algorithmique traite les données en bloc. Une illustration sera donnée plus loin.

### 4.2 Résolution directe par double diagonalisation

Il existe un cas particulier pour lequel on peut résoudre (5) directement *sans itération*. Il s'agit du cas où les filtres  $\phi_i$  sont symétriques et issus d'un sous-espace linéaire de dimension deux:  $\phi_i = \alpha_{i,1} \psi_1 + \alpha_{i,2} \psi_2$ , les deux filtres symétriques générateurs étant  $\psi_1$  et  $\psi_2$ . La symétrie des filtres rend l'expression  $\sum_{t=0}^{T-1} \sum_{u=0}^{T-1} \tilde{S}_i(t) \tilde{S}_j(u) \psi_m(t-u)$  symétrique en  $i$  et  $j$ , ce qui justifie alors l'égalité  $\hat{\mathbf{E}}_T[\tilde{S}_i(\psi_m \star \tilde{S}_j)] = \hat{\mathbf{E}}_T[\tilde{S}_j(\psi_m \star \tilde{S}_i)]$ ,  $m = 1, 2$ . Cela nous autorise à écrire les équation de séparation par paire  $\{i, j\}$  comme suit:

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{E}}_T[\tilde{S}_j(\phi_i \star \tilde{S}_i)] \\ \hat{\mathbf{E}}_T[\tilde{S}_i(\phi_j \star \tilde{S}_j)] \end{bmatrix} = \begin{bmatrix} \alpha_{i,1} & \alpha_{i,2} \\ \alpha_{j,1} & \alpha_{j,2} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{E}}_T[\tilde{S}_j(\psi_1 \star \tilde{S}_i)] \\ \hat{\mathbf{E}}_T[\tilde{S}_j(\psi_2 \star \tilde{S}_i)] \end{bmatrix}$$

Les filtres n'étant pas choisis proportionnels les uns aux autres, la matrice entre crochet ci-dessus est non singulière et les équations de la séparation sont ramenées sur  $\psi_1, \psi_2$  uniquement:  $\hat{\mathbf{E}}_T[\tilde{S}_j(\psi_m \star \tilde{S}_i)] = 0, m \in \{1, 2\}, i > j = 1, \dots, K$ . En fonction des covariances empiriques  $\mathbf{R}_X$ , on a:

$$\mathbf{e}_i^T \hat{\mathbf{A}}^{-1} \begin{bmatrix} T-1 \\ \sum_{k=1-T} \psi_{m,k} \mathbf{R}_X(k) \end{bmatrix} \hat{\mathbf{A}}^{-T} \mathbf{e}_j = 0 \quad i > j = 1, \dots, K, \quad m = 1, 2 \quad (14)$$

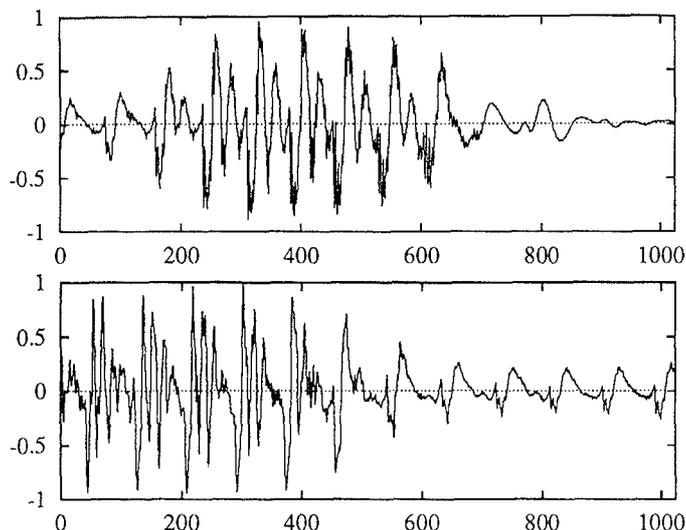
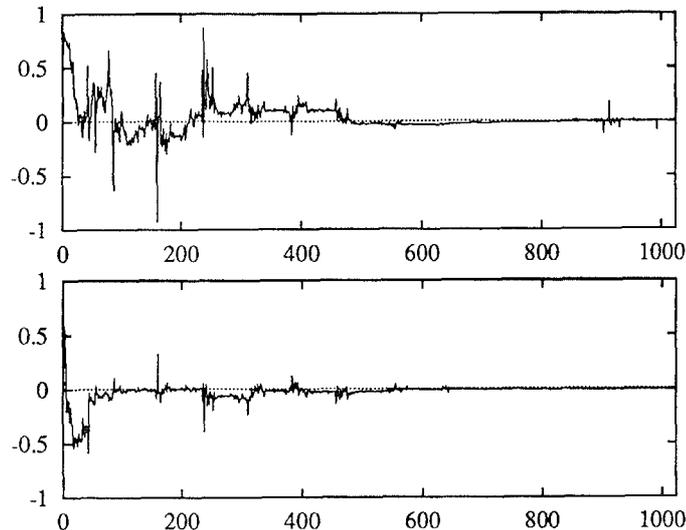


Figure 1: Signaux de parole non mélangés.

Figure 2: Evolution des écarts relatifs  $r_{12}$  et  $r_{21}$ .

On note  $\mathbf{Q}_X^{(m)} = \sum_k \psi_{m,k} \mathbf{R}_X(k)$  ( $m = 1, 2$ ). Tenant compte de leur symétrie, (14) signifie que les deux matrices  $\hat{\mathbf{A}}^{-1} \mathbf{Q}_X^{(m)} \hat{\mathbf{A}}^{-T}$   $m = 1, 2$  doivent être diagonales. En conclusion,  $\hat{\mathbf{A}}^{-1}$  est la matrice qui diagonalise simultanément les matrices symétriques  $\mathbf{Q}_X^{(1)}$  et  $\mathbf{Q}_X^{(2)}$ . Une telle matrice existe assurément si au moins l'une des deux matrices symétriques est définie positive. Cette condition est elle-même remplie si les filtres  $\psi_1$  et  $\psi_2$  sont choisis ayant une réponse en fréquence positive. Si de plus ils sont R.I.F. d'ordre  $p$ , l'évaluation de  $\mathbf{Q}_X^{(m)}$  n'exige que le calcul de  $\mathbf{R}_X(0), \dots, \mathbf{R}_X(p)$ .

La séparation par double diagonalisation de  $\mathbf{Q}_X^{(1)}$  et  $\mathbf{Q}_X^{(2)}$  peut être rendue adaptative, car les  $\mathbf{R}_X(k)$  sont aisément réévaluables après acquisition de nouvelles données. A chaque étape, il faut toutefois procéder à la diagonalisation.

Nous particularisons à présent deux situations simples pour lesquelles les filtres symétriques  $\phi_i$  sont issus d'un espace linéaire de dimension deux. *Cas 1:* Les filtres  $\phi_i$  sont tous R.I.F. d'ordre  $p = 1$ . Une base génératrice de ces filtres est donnée par deux filtres RIF d'ordre un quelconques. Ici, la méthode de séparation ne dépend pas du choix des filtres. *Cas 2:* Le nombre de sources est  $K = 2$ . Là on peut simplement choisir  $\psi_1 = \phi_1$  et  $\psi_2 = \phi_2$ .

## 5 Exemples d'application

Nous testons notre méthode de séparation sur des signaux de parole réels <sup>1</sup>(échantillonnés à 10kHz) que nous mélangeons artificiellement. La Fig. 1 montre l'allure des signaux non mélangés (deux blocs de 1024 échantillons) traités dans la suite. Nous nous plaçons dans le cas  $K = 2$  et donc les deux procédés de résolution (itératif ou direct) décrits en 4.1 et 4.2 sont possibles. La matrice de restitution  $\mathbf{C}$  est prise de la forme:

$$\begin{bmatrix} c_{11} = \sqrt{1 - c_{12}^2} & c_{12} \\ c_{21} & c_{22} = \sqrt{1 - c_{21}^2} \end{bmatrix} \text{ avec } c_{12}^2 + c_{21}^2 \leq 1.$$

Les deux filtres symétriques choisis sont RIF d'ordre  $p = 2$ , de réponse impulsionnelle  $(\dots, 0, 0, 1, 2, 3, 2, 1, 0, 0, \dots)$  et  $(\dots, 0, 0, 1, -2, 3, -2, 1, 0, 0, \dots)$ .

<sup>1</sup>fournis par C. Jutten et H.L. Nguyen du TIRF/INPG, Grenoble.

La résolution itérative nécessite une initialisation de  $c_{12}$  et  $c_{21}$ . Volontairement, nous prenons des valeurs initiales assez éloignées des valeurs théoriques attendues  $c_{ij}^*$ . Nous mesurons à chaque itération les écarts  $r_{ij} = \mathbf{e}_i^T \mathbf{C} \mathbf{C}^{*-1} \mathbf{e}_j$ . Ainsi, typiquement, nous observons une convergence des  $c_{ij}$  en moins de 10 itérations, les valeurs finales de  $r_{12}$  et  $r_{21}$  étant inférieures à 1%. Nous reproduisons dans le tableau ci-dessous un exemple de convergence en sept itérations:

	1	2	3	4	5	6	7
$r_{12}$	.584	.481	.235	.013	-.0080	-.0078	-.0078
$r_{21}$	.730	.636	.355	.038	-.0011	-.0014	-.0014

En ce qui concerne le procédé direct par double diagonalisation, nous nous plaçons dans un contexte adaptatif (avec facteur d'oubli égal à 0.999) et initialisons les matrices de covariances par des matrices nulles. La Fig. 2 reproduit alors l'évolution des erreurs  $r_{12}$  et  $r_{21}$  au cours du temps. On voit que la séparation est correcte au bout de cinq cents pas de temps (50ms).

## Références

- [1] J.-F. Cardoso, "Super-symmetric decomposition of the 4-th order cumulant tensor. Blind identification of more sources than sensors", *Proc. ICASSP 91*, Toronto, Canada, May 91, 5, pp. 3109-3112.
- [2] P. Comon, "Blind Separation of Sources, Part II", *Signal Processing*, 24, 1, 1991, pp. 11-20.
- [3] P. Comon, "Blind Identification in presence of noise", *Proc. EUSIPCO 92*, Brussels, Aug. 1992, 2, pp. 835-838.
- [4] P. Duvaut, "Principe des méthodes de séparation de sources fondées sur les moments d'ordre supérieur", *Traitement du Signal*, 7, 5, 1990, pp. 407-418.
- [5] C. Jutten, J. Herault, "Blind Separation of Sources, Part I", *Signal Processing*, 24, 1, 1991, pp. 1-10.
- [6] J.-L. Lacoume, P. Ruiz, "Separation of Independent Sources from Correlated Inputs", *IEEE Trans. on Signal Proc.*, 40, 12, 1992, pp. 3074-3078.
- [7] D.T. Pham, P. Garat, "Separation of a mixture of independent sources through a maximum likelihood approach", *Proc. EUSIPCO 92*, Brussels, Aug. 1992, 2, pp. 771-774.