



Séparation aveugle adaptative de mélanges convolutifs

Nathalie Delfosse
Télécom Paris, Dépt. Signal
46 rue Barrault, 75634 Paris Cedex 13

Philippe Loubaton
Télécom Paris, Dépt. Signal
46 rue Barrault, 75634 Paris Cedex 13

Résumé

Dans cet article, on montre que la séparation d'un mélange convolutif peut se faire en trois étapes - prédiction linéaire, séparation instantanée, mise en œuvre de l'inverse du filtre prédictif - et presque exclusivement à l'ordre 2 dès qu'il y a plus de capteurs que de sources. Dans ce contexte singulier, on propose une structure originale de filtre de synthèse (treillis décomposé normalisé), stable pour tout jeu des paramètres. Une application adaptative est alors possible; des simulations le confirment.

1 Introduction.

Nous disposons d'une suite d'observations réelles $y(n)$ de dimension q , obtenues par filtrage d'un vecteur $s(n)$ dont les p composantes sont des sources ARMA réelles, stationnaires, centrées, non gaussiennes, et indépendantes deux à deux:

$$y(t) = [H(z)]s(t). \quad (1)$$

Nous considérons un modèle sans bruit, d'une complexité déjà suffisante, et nous formulons les hypothèses suivantes:

- (H1): $q > p$ (il y a plus de capteurs que de sources).
- (H2): $H(z)$ est causale et rationnelle.
- (H3): $H(z)$ est de rang plein p pour tout z .

Comme $p < q$, cette dernière hypothèse sera vérifiée dès que les divers canaux de propagation entre les signaux sources et les capteurs seront suffisamment différents les uns des autres.

L'écriture (1) n'est pas unique: toute multiplication des colonnes de $H(z)$ par des fonctions de transfert scalaires, dépourvues de zéros, et compensées au niveau des signaux sources donne une écriture du même type. Dans ces conditions, chaque signal source ne peut être restitué qu'à un filtre près. Dans le cadre de cet article, nous nous fixons pour objectif de restituer la contribution $r_j(t) = [H_{j,j}(z)]s_j(t)$ de la j -ème source sur le j -ème capteur, définie intrinsèquement à partir des observations, comme le signal qu'on pourrait observer sur le j -ème capteur si les autres sources n'interféraient pas avec la source j .

2 Méthode générale.

Désignons par $(f_j(z))_{1 \leq j \leq p}$ et $(w_j(n))_{1 \leq j \leq p}$ les factorisations causales à phase minimale et les innovations normalisées des signaux sources. Si l'on pose $K(z) = H(z) \text{diag}(f_1(z), \dots, f_p(z))$, on a évidemment

$$y(n) = [K(z)]w(n).$$

L'hypothèse (H3) implique que $K(z)$, causale et rationnelle,

Abstract

In this paper, the separation of convolutive mixture problem is addressed. As soon as there are more sensors than sources, it can be solved using second order statistics through a three steps-method: analysis, instantaneous mixture separation, synthesis. A decomposed and normalized lattice synthesis filter is introduced. It solves the problem of singularity, and it is stable for any choice of its parameters. It can be adaptively computed, as shown by the simulation results.

est de rang plein p pour tout z , $|z| > 1$. On montre alors, en utilisant des résultats classiques, que $w(n)$ est une innovation normalisée de $y(n)$, tandis que $K(z)$ en est un facteur causal à phase minimale. On peut donc extraire $w(n)$ à une matrice orthogonale Θ près, en résolvant un problème de prédiction linéaire (à horizon fini en pratique), tandis que la matrice Θ peut être identifiée en utilisant une technique de séparation de mélanges instantanés.

L'hypothèse $q > p$ est capitale: en effet, si $q = p$, la condition pour que $K(z)$ soit à phase minimale est que $\det(H(z))$ ne s'annule pas pour $|z| > 1$. Cette condition est si restrictive qu'elle donnerait peu d'intérêt au problème que nous nous posons. C'est le fait de disposer de capteurs en nombre suffisant qui permet d'identifier $K(z)$ au second ordre sans que cela impose des restrictions trop importantes sur $H(z)$.

Nous proposons alors la méthode suivante:

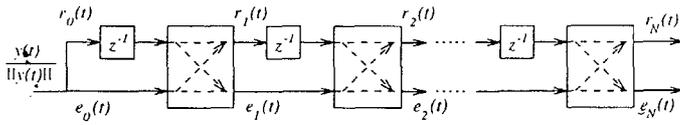
- **extraction de $w(n)$** à un facteur orthogonal près Θ par un algorithme de prédiction linéaire (ordre 2).
- **identification de Θ** pour extraire $w(n)$ (ordre 4).
Ces deux opérations permettent de déterminer un approximant FIR, $\hat{G}(z)$, d'un filtre causal $q \times q$ $G(z)$, non unique car $p < q$, pour lequel $G(z)K(z) = \begin{pmatrix} I_p \\ 0 \end{pmatrix}$.
- **estimation de $K(z)$** par $\hat{K}(z) = \hat{G}^{-1}(z) \begin{pmatrix} I_p \\ 0 \end{pmatrix}$,
et **restitution des $r_j(n) = \hat{K}_{jj}(z)w_j(n)$.**

Pour que cette démarche soit cohérente, il est nécessaire que $\hat{K}(z) = \hat{G}^{-1}(z) \begin{pmatrix} I_p \\ 0 \end{pmatrix}$ soit un filtre stable. Si $\hat{G}(z)$ était calculé via les équations de Yule-Walker associées aux corrélations exactes de l'observation, un résultat assez peu connu de Inouye [4], consacré à la prédiction linéaire dans le cas singulier, prouverait la stabilité de $K(z)$. Bien entendu, on ne dispose pas des corrélations exactes de $y(n)$, et, dans le cadre adaptatif qui est le notre, il est indispensable de contrôler à chaque instant la stabilité du filtre $\hat{G}^{-1}(z) \begin{pmatrix} I_p \\ 0 \end{pmatrix}$. Pour cela, nous proposons une structure de filtre prédictif particulière que nous détaillons dans la section suivante.



3 Filtres d'analyse et de synthèse.

Le filtre prédicteur classique en treillis a le schéma suivant:



On sait qu'avec des signaux scalaires, les coefficients de réflexion direct et rétrograde d'ordre n fixé sont égaux, de module inférieur à 1, et que le filtre de synthèse associé est stable pour tout jeu des coefficients de module inférieur à 1. Avec des signaux multivariés, l'égalité des coefficients direct et rétrograde n'est plus vérifiée et il faut considérer des signaux normalisés pour conserver cette propriété (qu'on montre alors par des arguments de géométrie et non de stationarité). Dans le cas non singulier (où $y(t)$ a pour dimension le rang de la matrice de covariance de son innovation), le filtre de synthèse associé est stable pour tout jeu des coefficients (matriciels) strictement contractants (voir [3]). Dans une optique adaptative, la contrainte de contractivité est malheureusement difficile à gérer; de plus, l'observation $y(t)$ est singulière. Nous avons choisi de nous orienter vers une structure de *filtre en treillis décomposé normalisé* qui met en jeu des coefficients de réflexion *scalaires*, aisément contrôlables, et où les singularités apparaissent localisées.

Soient $e_n(t)$ et $r_n(t)$ les erreurs normalisées directe et rétrograde d'ordre n . Elles sont définies à partir des erreurs non normalisées $w_n(t) = y(t) - y(t)/sp\{y_j(t-k), 1 \leq k \leq n, 1 \leq j \leq q\}$ et $z_n(t) = y(t-n) - y(t-n)/sp\{y_j(t-k), 0 \leq k \leq n-1, 1 \leq j \leq q\}$, (où "/" désigne l'opérateur de projection orthogonale), de sorte que leur covariance soit l'identité (en supposant que $E[w_n(t)w_n^T(t)]$ est de rang plein). Elles obéissent à la récurrence classique (cas non singulier),

$$\begin{pmatrix} r_n(t) \\ e_n(t) \end{pmatrix} = \begin{pmatrix} V_n & 0 \\ 0 & U_n \end{pmatrix} H(K_n) \begin{pmatrix} r_{n-1}(t-1) \\ e_{n-1}(t) \end{pmatrix} \quad (2)$$

$$\text{où } H(K) = \begin{pmatrix} (I - K^T K)^{-1/2} & K^T (I - K K^T)^{-1/2} \\ K (I - K^T K)^{-1/2} & (I - K K^T)^{-1/2} \end{pmatrix},$$

avec $K_n = E[e_{n-1}(t)r_{n-1}^T(t-1)]$, U_n et V_n étant des matrices orthogonales. La récurrence (2) correspond à une section du filtre d'analyse. Nous proposons de décomposer la rotation hyperbolique matricielle de cette récurrence en un produit de rotations scalaires comme dans [5], selon le schéma explicatif 1. Les signaux véhiculés sur ce schéma sont tous scalaires; il s'agit d'erreurs partielles normalisées directes (de gauche à droite) et rétrogrades (de haut en bas). A la n -ième section, l'erreur directe $e_{i,j}^{(n)}(t)$, obtenue en sortie de la rotation placée sur la i ème ligne et j ème colonne, est le résidu de la projection de $y_i(t)$ sur l'espace engendré par les $i-1$ premières composantes de $y(t)$, toutes celles de $y(t-1), \dots, y(t-n+1)$, et les j dernières composantes de $y(t-n)$. La situation est symétrique pour les erreurs partielles rétrogrades. La transformation élémentaire s'écrit:

$$\begin{pmatrix} r_{i,j}^{(n)}(t) \\ e_{i,j}^{(n)}(t) \end{pmatrix} = \theta(k_{i,j}^{(n)}) \begin{pmatrix} r_{i-1,j}^{(n)}(t) \\ e_{i,j-1}^{(n)}(t) \end{pmatrix}, \text{ où}$$

$\theta(k) = \frac{1}{\sqrt{1-k^2}} \begin{pmatrix} 1 & -k \\ -k & 1 \end{pmatrix}$. $k_{i,j}^{(n)}$ est le produit scalaire des erreurs partielles normalisées $e_{i,j-1}^{(n)}(t)$ et $r_{i-1,j}^{(n)}(t)$ en entrée de la transformation. Il s'agit d'un coefficient de réflexion partiel, de module inférieur à 1. Le filtre d'analyse décomposé

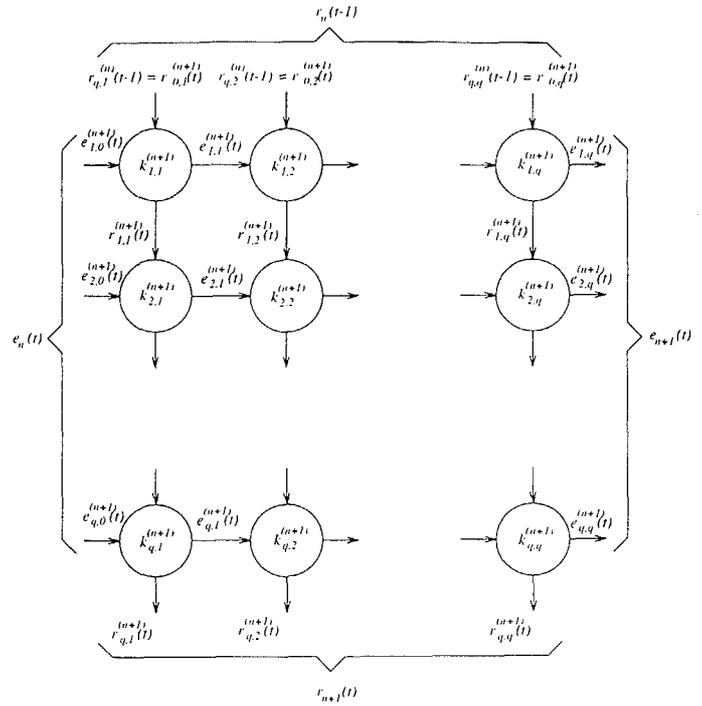


Figure 1: décomposition d'une section.

effectue une orthogonalisation de Gram-Schmidt sur les composantes de $y(t), \dots, y(t-n)$.

Comme $y(t) = [H(z)]s(t)$, avec $H(z)$ $q \times p$ où $q > p$, les composantes du vecteur $Y_N(n) = (y(t)^T, \dots, y(t-N)^T)^T$ sont liées pour N assez grand. Il s'ensuit que la procédure d'orthogonalisation de Gram-Schmidt que nous avons décrite fait apparaître dans le filtre d'analyse "en construction", comportant un *nombre fini de sections*, un vecteur du noyau de la matrice de covariance de $Y_N(n)$. Ceci est réalisé avant qu'aucune composante de l'innovation normalisée de $y(t)$ ne soit extraite, car il faut en effet pour cela effectuer une prédiction à horizon infini, sauf configuration très particulière. Une erreur partielle directe va donc s'annuler, et elle ne pourra pas être normalisée. Supposons que l'erreur de prédiction directe $e_{i,j}^{(n)}(t)$ soit de puissance nulle (et que ce soit la première dégénérescence rencontrée). Alors, nécessairement, $e_{i,j-1}^{(n)}(t) - k_{i,j}^{(n)} r_{i-1,j}^{(n)}(t) = 0$. Ceci implique que $e_{i,j-1}^{(n)}(t)$ est proportionnelle à $r_{i-1,j}^{(n)}(t)$, et que $k_{i,j}^{(n)}$, produit scalaire de $e_{i,j-1}^{(n)}(t)$ et $r_{i-1,j}^{(n)}(t)$ vaut ± 1 , donc, que les deux erreurs en question sont égales au signe près. (Il est clair que la réciproque est vraie). Il apparaît que $r_{i,j}^{(n)}(t)$ est également identiquement nulle. Afin de poursuivre la prédiction, il faut donc remplacer, dans la n -ième section où la première dégénérescence a été observée, toutes les rotations hyperboliques élémentaires faisant suite à $\theta(k_{i,i}^{(n)})$, c'est-à-dire sur la même ligne et sur la même colonne (voir figure 1), par la matrice identité 2×2 . De cette façon, les erreurs de puissance nulle sont acheminées jusqu'en sortie de la n -ième section, puis directement jusqu'en sortie du filtre. La figure 2 donne un exemple de section dégénérée. Une section initiale, notée B , permet de passer de $y(t)$ à $e_0(t)$ et $r_0(t)$, qui sont deux versions blanchies de $y(t)$ par deux orthogonalisations de Schmidt commencées sur $y_1(t)$ et $y_q(t)$ respectivement.

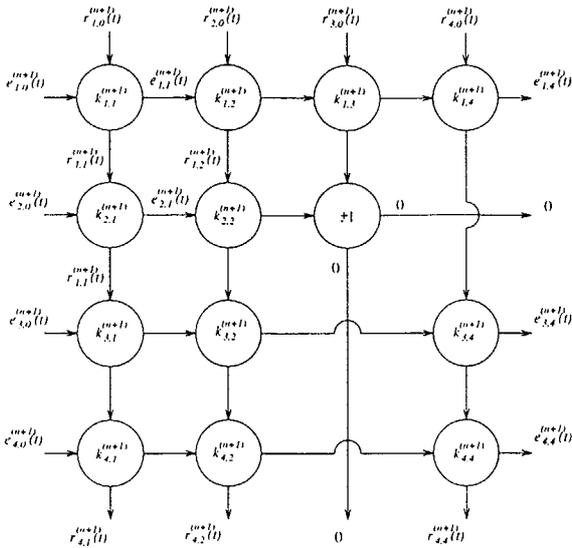


Figure 2: section dégénérée.

Un exemple de filtre d'analyse décomposé dans le cas singulier et du filtre de synthèse associé est donné en figure 4 (sachant que les sections sont décrites par les autres schémas).

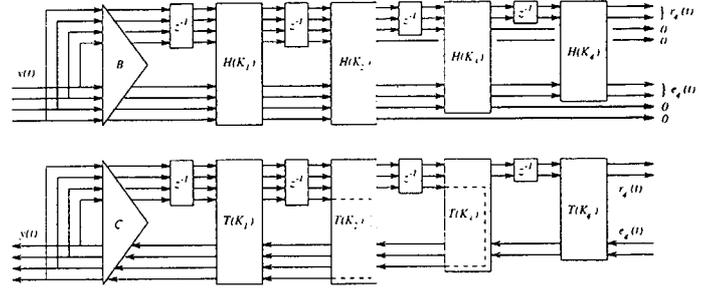


Figure 4: filtres d'analyse et de synthèse.

La question qui reste en suspens est celle de la stabilité du treillis de synthèse quand les coefficients $k_{i,i}^{(n)}$ prennent des valeurs quelconques dans $] -1, 1[$. Un résultat sur les filtres sans perte permet de conclure:

Le filtre de synthèse s'obtient à partir du filtre d'analyse en inversant le sens de propagation des erreurs directes. Le cas singulier posant toutefois quelques problèmes, nous procéderons par étapes pour décrire sa structure. La rotation hyperbolique élémentaire s'inverse sous la forme d'une rotation: $\begin{pmatrix} r_{i,j}^{(n)}(t) \\ e_{i,j-1}^{(n)}(t) \end{pmatrix} = T(k_{i,j}^{(n)}) \begin{pmatrix} r_{i-1,j}^{(n)}(t) \\ e_{i,j}^{(n)}(t) \end{pmatrix}$, avec $T(k) = \begin{pmatrix} \sqrt{1-k^2} & -k \\ k & \sqrt{1-k^2} \end{pmatrix}$. Une section courante du treillis de

synthèse se déduit donc du schéma 1 en inversant le sens de propagation (horizontal) des erreurs décomposées directes, et en remplaçant les $\theta(k_{i,j}^{(n)})$ par les $T(k_{i,j}^{(n)})$. La section B s'inverse de manière analogue sans difficulté; on constate que son inverse se réduit à une matrice orthogonale.

Passons à l'inversion d'une section dégénérée. On a vu que les deux erreurs scalaires en entrée d'une rotation hyperbolique dégénérée sont égales au signe près. La transformation s'inverse donc par un court-circuit avec le signe approprié. Le figure 3 est un exemple d'inversion d'une section dégénérée.

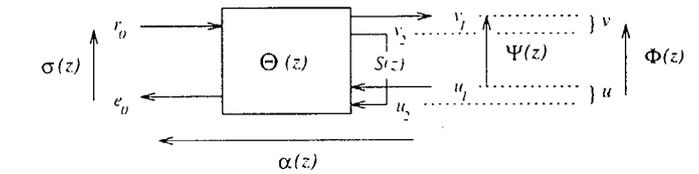


Figure 5: quadripole sans perte.

Résultat 3.1 Soit le quadripôle défini par la figure 5, supposé sans perte. Soit $\Theta(z)$ sa matrice de répartition: $\begin{pmatrix} u \\ v \end{pmatrix} = [\Theta(z)] \begin{pmatrix} r_0 \\ e_0 \end{pmatrix}$, en posant $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ et $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$. On suppose que les signaux r_0 et e_0 sont reliés par la fonction de transfert $\sigma(z)$: $r_0 = [\sigma(z)]e_0$, où $\sigma(z)$ est sans perte. Alors, v s'écrit sous la forme $v = [\Phi(z)]u$, où $\Phi(z)$ est une fonction de transfert sans perte. Supposons de plus que $\|\Phi(z)\| < 1$ pour $|z| > 1$, et que les dernières composantes de u soient une version filtrée de celles de v : $u_2 = [S(z)]v_2$, avec $S(z)$ sans perte. Alors les signaux e_0 et v_1 sont obtenus par filtrage de u_1 . Appelons $\alpha(z)$ et $\Psi(z)$ les fonctions de transfert associées: $e_0 = [\alpha(z)]u_1$ et $v_1 = [\Psi(z)]u_1$. Alors, les fonctions $\alpha(z)$ et $\Psi(z)$ sont stables; de plus, $\Psi(z)$ est sans perte.

Le filtre de synthèse peut se voir sous la forme d'un quadripôle du type 5: soient u_2 et v_2 les vecteurs formés des erreurs partielles directes et rétrogrades court-circuitées (et donc liées par une matrice orthogonale), u_1 et v_1 ceux correspondant aux autres erreurs en entrée et en sortie du filtre; comme on sait que e_0 et r_0 sont égales à une matrice orthogonale près (donc sans perte), les hypothèses de 3.1 sont vérifiées - pour la stricte-contractivité de Φ et une démonstration détaillée, voir[1].

Le filtre de synthèse ainsi paramétré est donc stable pour tout jeu des coefficients de réflexion scalaires $k_{i,j}^{(n)}$ vérifiant $|k_{i,j}^{(n)}| < 1$.

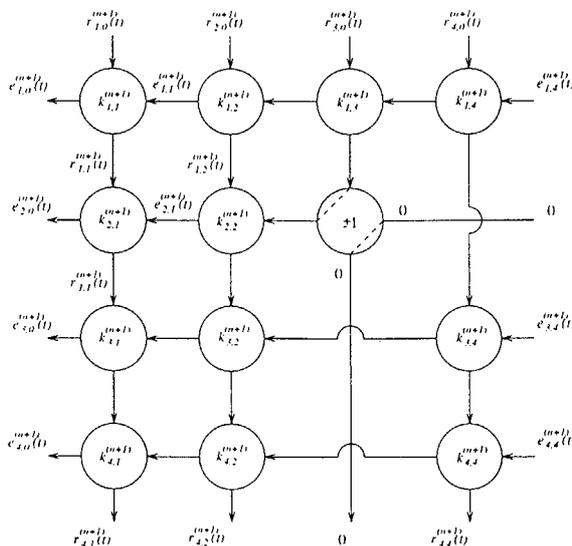


Figure 3: inversion d'une section dégénérée.



4 Simulations en adaptatif.

Seules les deux premières étapes de la méthode de séparation que nous avons décrite sont réalisées de façon adaptative. Il est nécessaire de fixer *a priori* la structure du prédicteur (en particulier le nombre de sources p , et la place des dégénérescences). Il peut paraître illusoire de tenter une séparation adaptative quand des paramètres d'une telle importance pour le modèle de factorisation causale à phase minimale doivent être choisis *a priori*. Il faut comprendre toutefois que ces indéterminations sont inhérentes au problème de séparation comme à celui de l'égalisation. On peut remarquer que les dégénérescences apparaîtront la plupart du temps sur les dernières cellules scalaires de certaines sections. En fixant *a priori* la place des court-circuits à de tels endroits, nous minimiserons le risque d'erreur. Il nous faut aussi fixer un ordre maximum N à la prédiction. L'approximation ainsi réalisée est d'autant meilleure que les signaux sources sont peu dépendants de leurs passés.

Remarquons que le coefficient $k_{i,j}^{(n)}$ minimise la fonction

$$E \left((e_{i,j-1}^{(n)}(t) - k_{i,j}^{(n)} r_{i-1,j}^{(n)}(t))^2 \right). \quad (3)$$

La méthode optimale serait alors de résoudre le système

$$\frac{\partial}{\partial k_{i,j}^{(n)}} \sum_{i,j,n} E \left((e_{i,j-1}^{(n)}(t) - k_{i,j}^{(n)} r_{i-1,j}^{(n)}(t))^2 \right) = 0$$

par un gradient stochastique par exemple, mais la complexité de la structure des filtres rend la dérivée de la fonction de coût incalculable. Il faut donc se résoudre à employer une procédure sous-optimale. Or, la dépendance entre les $k_{i,j}^{(n)}$ est triangulaire, ce qui permet d'identifier ces coefficients en minimisant de façon adaptative les critères (3) *en cascade*, avec $k_{i,j}^{(n)} = \sin \phi_{i,j}^{(n)}$ pour garantir $|k_{i,j}^{(n)}| \leq 1$ (la convergence de l'ensemble se montre comme dans [2]).

Les simulations ont été réalisées avec 2 sources AR(3) et 3 capteurs, avec pour matrice de mélange $H(z) = B(z)A^{-1}(z)$, où $A(z) = \begin{pmatrix} 2 - 4/3 z^{-1} & 1 + 1/2 z^{-1} \\ 3/4 - 9/16 z^{-1} - 1 & -1/3 z^{-1} \end{pmatrix}$, et $B(z) = \begin{pmatrix} 1 - 1/2 z^{-1} & 2 + 3/2 z^{-1} \\ 2 + 8/5 z^{-1} & 3 - z^{-1} \\ -1 + 1/4 z^{-1} & 2 + 10/7 z^{-1} \end{pmatrix}$. Dans ce cas, il existe

un prédicteur exact de degré 3, possédant une dégénérescence sur la dernière cellule de la seconde section (voir [1]). L'algorithme a été testé sur les signaux séparés bien sûr, et, de façon plus visuelle, par la convergence des paramètres angulaires du prédicteur. Faute de place, nous ne présentons ici que l'évolution des paramètres des première et dernière sections, avec 20000 itérations et un pas de $5 \cdot 10^{-4}$. Il est clair que l'algorithme que nous proposons est lent, mais il faut relativiser ce comportement en tenant compte de la complexité du problème d'une part, et du caractère sous-optimal de la procédure qui est le point à améliorer si l'on souhaite de meilleures performances. On peut s'orienter vers des algorithmes plus rapides, de type moindres carrés récursifs par exemple, pour minimiser $\|y(t) - \sum_{k=1,N} A_k y(t-k)\|^2$.

5 Conclusion.

Séparer un mélange convolutif où le nombre des capteurs est supérieur à celui des sources peut se faire en trois étapes -

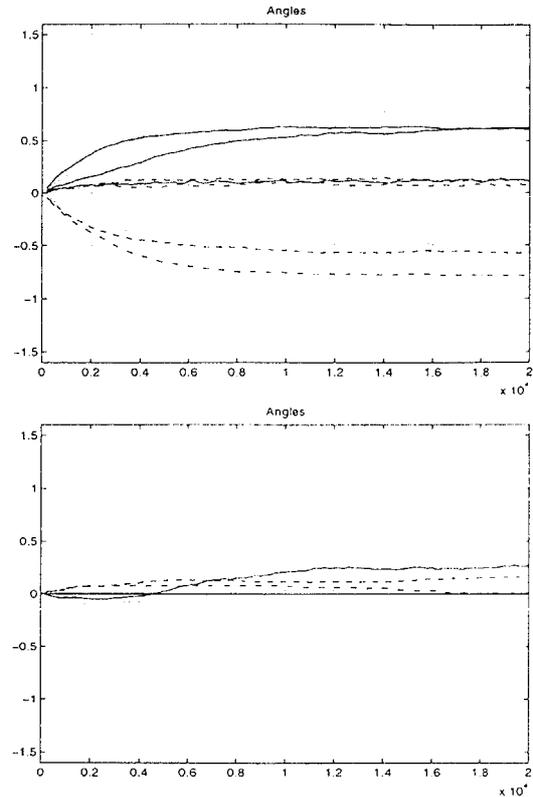


Figure 6: angles des 1ère et 3ème sections.

prédiction, séparation instantanée, synthèse - et presque exclusivement à l'ordre 2. En adaptatif, la difficulté réside dans le choix d'une paramétrisation du prédicteur qui garantisse à chaque instant la stabilité du filtre de synthèse. Ceci peut être réalisé via un treillis décomposé normalisé. Les paramètres de la structure séparatrice peuvent être identifiés de façon adaptative par une cascade de LMS minimisant des critères scalaires identiques. La cascade est alors convergente mais sous-optimale. Elle serait à améliorer par l'utilisation d'un algorithme RLS décomposé.

Références

- [1] N. Delfosse. *Séparation Aveugle Adaptative de Mélanges de Signaux Indépendants*. Thèse de doctorat de l'ENST. en préparation.
- [2] N. Delfosse and P. Loubaton. Adaptive blind separation of independent sources: a deflation approach. *Signal Processing*, 45(1), 1995. to appear in december, 1995.
- [3] P. Delsarte, Y. Genin, and Y. Kamp. Schur parametrization of positive definite block-toeplitz systems. *SIAM J. on Appl. Math.*, 36(1):933-945, February 1979.
- [4] Y. Inouye. Modeling of multichannel time series and extrapolation of matrix-valued autocorrelation sequences. *IEEE Tr. on ASSP*. 31(1), February 1983.
- [5] M.S. Jazaerli. *Contribution à la Décomposition des Algorithmes de Moindres Carrés Vectoriels Normalisés - Application au Traitement d'Antenne*. Thèse de l'Université de Rennes I, 1987.