

Séparation adaptative de sources indépendantes par une approche de déflation

Nathalie Delfosse
Telecom Paris, Dept. Signal
46 rue Barrault, 75634 Paris Cedex 13

Philippe Loubaton
Telecom Paris, Dept. Signal
46 rue Barrault, 75634 Paris Cedex 13

Résumé

Dans cet article, on présente un nouveau schéma adaptatif de séparation de sources utilisant les statistiques d'ordre 2 et 4. Contrairement aux algorithmes existants, sa convergence est assurée grâce au recours à une procédure de déflation. Ses performances asymptotiques sont évaluées analytiquement, et quelques résultats de simulations sont présentés.

1 Introduction.

Dans de nombreuses applications mettant en jeu des traitements multicapteurs, le signal q -dimensionnel $y(n)$ (supposé réel par souci de simplicité) dont les composantes représentent les divers signaux capteurs s'écrit sous la forme

$$y(n) = \sum_{k=1}^p h_k w_k(n) = H w(n) \quad (1.1)$$

où les $(w_k)_{k=1,p}$ sont des signaux stationnaires indépendants entre eux, et où les $(h_k)_{k=1,p}$ sont des vecteurs inconnus, linéairement indépendants. Le problème, désormais classique, de la séparation aveugle des signaux $(w_k)_{k=1,p}$, consiste à extraire les $(w_k)_{k=1,p}$, ou de façon équivalente, à identifier l'inverse à gauche de la matrice H , à partir des observations $y(n)$. Il est bien connu que la seule utilisation de $S_y = E(y(n)y^T(n))$ ne peut permettre de résoudre cette question, car la connaissance de S_y ne permet de déterminer H qu'à une matrice orthogonale près. Cependant, si pour chaque k , les variables $(w_k(n))_{n \in \mathbb{Z}}$ sont corrélées, il est alors possible de lever cette indétermination en utilisant les valeurs prises par la fonction d'autocorrélation $E(y(n+m)y^T(m))$ de y pour $n \geq 1$ (voir par exemple [5], [14]). Mais il existe un certain nombre de situations où les signaux sources peuvent être des suites indépendantes identiquement distribuées (i.i.d.). Dans ce cas, sous l'hypothèse que les $(w_k)_{k=1,p}$ sont non gaussiens, l'utilisation de statistiques d'ordre supérieur à 2 de $y(n)$ permet de séparer les sources. Puisque cela n'introduit aucune restriction dans ce cadre, nous allons supposer à présent que $p = q$, i.e. que la matrice H est carrée.

La plupart des méthodes basées sur les statistiques d'ordre supérieur ont été développées dans un cadre non adaptatif, et il semble que le problème spécifique de la séparation aveugle adaptative ne soit pas encore résolu d'une façon satisfaisante: en effet, aucun des algorithmes proposés dans la littérature n'est assuré de converger vers l'une des valeurs souhaitées. Ainsi, les algorithmes de type Jutten-Hérault possèdent des attracteurs indésirables (voir [13], [9]), de même que l'implantation adaptative de la méthode proposée récemment par Pham et al ([10]). Cardoso et Laheld ([3]), après blanchiment des données, minimisent de façon adaptative sur l'ensemble des matrices orthogonales le contraste $\mathcal{M}(G) = \sum_{k=1}^p (E(Gy(n))_k)^4$, minimal en H^T (à une permutation près). Mais, comme pour les contrastes introduits par Comon dans [2], il n'est pas établi qu'il n'existe pas d'extrema parasites. On ne peut donc pas assurer la convergence de ce type d'algorithme. Le but de cet article est de présenter des schémas de séparation dont la convergence est établie.

Abstract

In this paper, a new adaptive algorithm for the blind source separation problem is introduced. In contrast with the existing algorithms, its convergence is guaranteed by using a deflation approach. Its asymptotical performance is analytically calculated and some simulation results are presented.

Nous adaptons ici une idée des articles [7] et [8], consacrés au problème de la déconvolution des signaux vectoriels - "généralisation" du nôtre: les vecteurs constants $(h_k)_{k=1,p}$ sont remplacés par des fonctions de transfert. Par souci de simplicité, nous ne considérons que le cas où les cumulants d'ordre 4 des signaux sources sont strictement négatifs; nous indiquerons plus bas comment le cas général peut être traité. Comme dans nombre de méthodes de séparation, nous commençons par effectuer un blanchiment (adaptatif) des données, qui nous autorise à supposer que $E(y(n)y^T(n)) = I$, et que la matrice H est orthogonale. Notre méthode repose sur le fait qu'il est simple d'extraire l'un des signaux sources en adaptant à notre contexte l'approche développée par Shalvi et Weinstein ([12]) dans le cadre de la déconvolution aveugle. Plus précisément, soit \mathcal{K} la fonction définie sur l'ensemble des vecteurs lignes de norme 1 de R^p par

$$\mathcal{K}(g) = E(gy(n))^4/4 \quad (1.2)$$

Il est facile d'établir (voir [4]) que les arguments des minima locaux de \mathcal{K} sur la sphère unité \mathcal{S} de R^p sont les vecteurs $(h_k^T)_{k=1,p}$. Dans ces conditions, si $g(n)$ désigne le vecteur généré à l'instant n par l'algorithme du gradient stochastique associé à la minimisation de \mathcal{K} sur \mathcal{S} , la suite $g(n)$ est assurée de converger vers l'un des vecteurs $(h_k^T)_{k=1,p}$ (par exemple le vecteur h_p^T), et le signal $r(n) = g(n)y(n)$ représentera un estimateur du signal $w_p(n)$. Pour extraire les autres signaux sources, nous utilisons une approche de déflation, basée sur une paramétrisation des vecteurs $g(n)$ de norme 1, introduite par Regalia dans [11], et qui permet la génération à chaque instant n d'une matrice $(p-1) \times p$, notée $G'(n)$, pour laquelle $G(n) = \begin{pmatrix} G'(n) \\ g(n) \end{pmatrix}$ est une matrice orthogonale. Le point

fondamental est que $y^1(n) = G'(n)y(n)$ va converger vers un signal $(p-1)$ -dimensionnel, non corrélé avec $w_p(n)$ du fait que $g(n)G'^T(n) = 0$, et qui pourra s'écrire sous la forme $H^1 w^1(n)$, où $w^1(n) = (w_1(n), \dots, w_{p-1}(n))^T$, et où H^1 est une matrice orthogonale $(p-1) \times (p-1)$. Par conséquent, si $g^1(n)$ désigne une suite de vecteurs de norme 1 de R^{p-1} , obtenue en échangeant $y(n)$ avec $y^1(n)$ dans ce qui précède, il est possible d'établir que $g^1(n)y^1(n)$ convergera vers l'une des excitations $w_k(n)_{k=1,p-1}$. En itérant cette procédure, on peut extraire tous les signaux sources (voir figure 1).

Cette approche peut être étendue au cas où les signes des cumulants d'ordre 4 des signaux sources sont quelconques. Pour cela, il suffit de remplacer dans ce qui précède la fonction \mathcal{K} par la fonction \mathcal{J} définie sur \mathcal{S} par



$$\mathcal{J}(g) = [c_4(gy(n))]^2/4 = (E(gy(n))^4 - 3(E(gy(n))^2)^2)/4 \quad (1.3)$$

Il est établi dans [4] que les maxima locaux de \mathcal{J} sur \mathcal{S} sont les vecteurs $(h_k^T)_{k=1,p}$. Notons toutefois que \mathcal{J} ne coïncide pas avec l'espérance mathématique d'un critère, de sorte que la maximisation adaptative de \mathcal{J} ne peut s'effectuer par un gradient stochastique. Pour cela, il est possible d'utiliser un des algorithmes proposés sans justification dans [12]. La preuve de leur convergence pourra être trouvée dans [4].

2 Paramétrisation du vecteur $g(n)$.

Commençons par présenter la paramétrisation des vecteurs de norme 1 que nous utilisons pour minimiser \mathcal{K} sur \mathcal{S} . Elle repose sur le résultat suivant (voir [11]).

Lemma 2.1 *Tout vecteur unitaire g de \mathbb{R}^p dont la dernière composante est strictement positive coïncide avec la dernière ligne d'une matrice orthogonale $G(\theta)$ donnée par*

$$G(\theta) = G_1(\theta_1) \cdots G_{p-1}(\theta_{p-1}) \quad (2.1)$$

où les $(\theta_k)_{k=1,p-1}$ sont des angles de $]-\pi/2, \pi/2[$ définis de façon unique, et les $G_k(\theta_k)$, les rotations de Givens

$$G_k(\theta_k) = \begin{pmatrix} I_{k-1} & & & \\ & \cos \theta_k & & -\sin \theta_k \\ & & I_{p-1-k} & \\ & \sin \theta_k & & \cos \theta_k \end{pmatrix} \quad (2.2)$$

Dans la suite, nous noterons $g(\theta)$ la dernière ligne de $G(\theta)$, et $G'(\theta)$ la matrice $(p-1) \times p$ constituée des $(p-1)$ premières lignes de $G(\theta)$. Avec l'hypothèse peu restrictive que les dernières composantes des vecteurs $(h_k)_{k=1,p}$ sont non nulles, le Lemme ci-dessus montre que la minimisation adaptative de \mathcal{K} sur \mathcal{S} équivaut à la détermination d'une suite de vecteurs angulaires $\theta(n)$ pour laquelle $g(\theta(n))$ converge vers un minimum de \mathcal{K} . Posons $r(n, \theta) = g(\theta)y(n)$. Alors, l'algorithme du gradient stochastique correspondant à la minimisation de \mathcal{K} sur \mathcal{S} coïncide avec celui associé à la minimisation de $E(r^2(n, \theta))$. Soit $\Gamma(\theta)$ la matrice diagonale dont les éléments $(\gamma_k(\theta))_{k=1,p-1}$ sont donnés par $\gamma_1(\theta) = 1$, et $\gamma_k(\theta) = \prod_{l=1,k-1} \cos \theta_l$ pour $k > 1$. En utilisant le fait que $\partial g(\theta)/\partial \theta = \Gamma(\theta)G'(\theta)$, l'algorithme du gradient stochastique s'écrit

$$\theta(n+1) = \theta(n) - \mu r^3(n) \Gamma(\theta(n)) y^1(n) \quad (2.3)$$

avec $r(n) = r(n, \theta(n))$ et $y^1(n) = G'(n, \theta(n))y(n)$, pour alléger les notations. Remarquons pour terminer que, pour permettre à l'algorithme (2.3) de s'adapter à d'éventuelles non stationnarités, son pas μ a été choisi constant. Contrairement au cas où on utilise des pas μ_n convergeant vers 0 en $1/n$, il est bien connu que (2.3) ne converge pas presque sûrement (p.s.), et que l'on ne dispose d'aucun résultat permettant d'analyser rigoureusement son comportement asymptotique. Cependant, l'approche développée par Kushner, basée sur la notion de convergence faible, suggère que si μ est choisi assez petit, (2.3) oscillera autour de l'un des points de convergence possible de l'algorithme à pas décroissant correspondant. Par souci de simplicité, nous continuerons à parler de convergence de l'algorithme (2.3) dans ce cas.

3 L'algorithme de déflation.

D'après ce qui précède, l'extraction de l'une des excitations $(w_k)_{k=1,p}$ peut être réalisée très simplement grâce à l'algorithme (2.3). Ainsi que nous l'avons indiqué dans l'introduction, nous proposons de séparer un autre signal source en appliquant au signal $(p-1)$ -dimensionnel $y^1(n)$

défini plus haut un algorithme similaire à (2.3). Plus précisément, soit $(\theta^1(n))_{n \in \mathbb{N}}$ la suite de vecteurs angulaires de dimension $(p-2)$ générée par l'algorithme

$$\theta^1(n+1) = \theta^1(n) - \mu r_1^3(n) \Gamma(\theta^1(n)) y^2(n) \quad (3.1)$$

où $r_1(n) = g(\theta^1(n))y^1(n)$ et $y^2(n) = G'(\theta^1(n))y^1(n)$. Établissons par des arguments un peu plus formels que ceux de l'introduction que $r(n)$ et $r^1(n)$ convergent vers deux signaux sources différents. Pour cela, supposons que le pas constant μ de (2.3) et (3.1) soit remplacé par une suite μ_n convergeant vers 0 en $1/n$. Soit θ_* la limite p.s. de (2.3), et supposons, sans que cela introduise la moindre restriction, que $g(\theta_*) = h_p^T$. Le signal $r(n, \theta_*)$ coïncide alors avec $w_p(n)$, tandis que $y^1(n, \theta_*)$ s'écrit $y^1(n, \theta_*) = G'(\theta_*)y(n) = G'(\theta_*)Hw(n)$. Or $h_p = g^T(\theta_*)$, la dernière colonne de $G'(\theta_*)H$ est donc réduite au vecteur nul du fait de l'orthogonalité structurelle de $g(\theta_*)$ avec les lignes de $G'(\theta_*)$. Par conséquent, $y^1(n, \theta_*)$ s'écrit $y^1(n, \theta_*) = H^1 w^1(n)$ où H^1 est une matrice orthogonale $(p-1) \times (p-1)$ et où $w^1(n) = (w_1(n), \dots, w_{p-1}(n))^T$. Si θ_* était connu, l'algorithme du gradient stochastique associé à la minimisation de $E[g(\theta^1)y^1(n, \theta_*)]^4$ permettrait d'extraire l'un des signaux sources $(w_k)_{k=1,p-1}$. Montrons que cet algorithme a le même comportement asymptotique que (3.1). Pour cela, posons $\phi^1(\theta^1, \theta, y(n)) = r_1^3(n, \theta^1, \theta) \Gamma(\theta^1) y^2(n, \theta^1, \theta)$. D'après la formule des accroissements finis,

$$\phi^1(\theta(n), \theta^1(n), y(n)) = \phi^1(\theta_*, \theta^1(n), y(n)) + \frac{\partial \phi^1}{\partial \theta}(\tilde{\theta}(n), \theta^1(n), y(n))(\theta(n) - \theta_*) \quad (3.2)$$

où $\tilde{\theta}(n)$ appartient au segment joignant $\theta(n)$ à θ_* , et où $\phi^1(\theta_*, \theta^1(n), y(n)) = r_1^3(n, \theta_*) \Gamma(\theta^1(n)) G'(\theta^1(n)) y^1(n, \theta_*)$ n'est rien d'autre que le terme de réactualisation correspondant à l'algorithme du gradient stochastique associé à la minimisation de $E[g(\theta^1)y^1(n, \theta_*)]^4$. Puisque $(\theta(n) - \theta_*)$ converge p.s. vers 0, (3.1) peut être réécrit sous la forme

$$\theta^1(n+1) = \theta^1(n) - \mu_n r_1^3(n, \theta_*) \Gamma(\theta^1(n)) G'(\theta^1(n)) y^1(n, \theta_*) + \mu_n \epsilon(n) \quad (3.3)$$

où, modulo quelques conditions techniques, $\epsilon(n)$ converge vers 0 p.s. En utilisant les résultats de [6], il apparaît que (3.1) a le même comportement asymptotique que (3.3) amputé du terme $\mu_n \epsilon(n)$. D'après ce qui précède, ceci implique que le signal $r^1(n)$ généré par (3.1) convergerait vers l'un des signaux sources $(w_k)_{k=1,p-1}$ si le pas de l'algorithme était décroissant. Il suffit d'utiliser l'argument invoqué en section 2 pour en déduire le comportement de (3.1) dans le cas plus réaliste où le pas est constant. Un raisonnement analogue justifierait la convergence des algorithmes d'extraction des signaux restants. Le schéma global de séparation est alors le suivant:

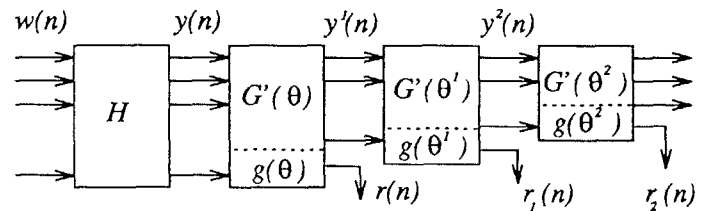


Figure 1: Séparation de sources par déflation

Rappelons enfin que dans notre analyse, nous avons supposé parfait le blanchiment des observations. Tel n'est certes pas le cas, et en pratique, il convient de générer de façon adaptative,

à partir de $y(n)$, un estimateur $\tilde{y}(n)$ du signal blanchi, puis de remplacer $y(n)$ par $\tilde{y}(n)$ dans les algorithmes qui précèdent. Il est facile de justifier que ceci n'affecte pas la convergence du schéma en utilisant des arguments du même type que ceux que nous venons d'employer.

4 Performances statistiques.

Nous allons fournir ici quelques résultats concernant les variances asymptotiques des différents estimateurs mis en œuvre. Pour cela, nous utilisons le résultat suivant: soit l'algorithme

$$\alpha(n+1) = \alpha(n) + \mu\psi(\alpha(n), z(n)) \quad (4.1)$$

où $z(n)$ est un processus stationnaire ergodique, et ψ une fonction suffisamment régulière. Supposons que (4.1) "converge" vers le point α_* . Alors, la matrice de covariance de l'erreur normalisée $1/\sqrt{\mu}(\alpha(n) - \alpha_*)$ tend quand $\mu \rightarrow 0$ vers l'unique solution P de l'équation de Lyapounov $AP + PA^T + Q = 0$ où A est la matrice dérivée au point α_* du champ moyen $\bar{\psi}(\alpha) = E(\psi(\alpha, z(n)))$ et Q est donnée par $Q = \sum_{n=-\infty}^{\infty} E(\psi(\alpha_*, z(n))\psi(\alpha_*, z(0))^T)$. Notons que la démonstration rigoureuse de ce résultat nécessite une hypothèse très forte sur (4.1): que son équation différentielle moyenne associée possède un *unique* attracteur. Nos algorithmes admettant plusieurs points de convergence possible, cette condition n'est pas vérifiée, mais on déduit le résultat précédent en utilisant formellement un résultat plus général d'approximation diffusion (voir [1], p.104).

Puisque nos algorithmes se mettent sous la forme (4.1), il suffit de calculer A et Q , pour obtenir P . Il est difficile d'effectuer les calculs analytiques permettant d'évaluer les variances de tous les étages de la procédure de déflation, en prenant en compte l'effet du blanchiment. Nous ne donnons donc que des résultats incomplets, mais qui permettent d'avoir une idée des grandes tendances. Notons enfin que dans nos calculs, nous avons fait l'hypothèse que les signaux sources sont des suites i.i.d.. Ceci est cohérent avec l'utilisation de méthodes basées sur les statistiques d'ordre supérieur de $y(n)$.

4.1 Cas d'un blanchiment exact.

Commençons par donner quelques résultats dans le cas où le blanchiment est supposé parfait. Le calcul des variances asymptotiques du premier étage de notre schéma (i.e. l'algorithme (2.3)) est particulièrement simple. En effet, supposons que (2.3) permette d'extraire le signal w_p . On montre alors (voir [4]) que la matrice P est donnée par $P = -E(w_p^6)/2c_4(w_p)I$, et que la variance de l'erreur de reconstruction normalisée (i.e. $1/\sqrt{\mu}(r(n) - w_p(n))$) est égale à $-\frac{E(w_p^6)}{2c_4(w_p)} \sum_{k=1}^{p-1} \gamma_k^2(\theta_*)$ (on rappelle que les cumulants des sources sont supposés négatifs). Ces expressions appellent quelques commentaires. La matrice P est un multiple de l'identité; les erreurs sur les diverses composantes de θ sont donc non corrélées entre elles, et de même variance. D'autre part, le terme $-E(w_p^6)/2c_4(w_p)$ indique clairement comment les performances se dégradent lorsque la loi de w_p tend vers une gaussienne: d'une part le cumulants d'ordre 4 tend vers 0, et d'autre part, le moment d'ordre 6 grandit. Enfin, il est remarquable de constater que $-E(w_p^6)/2c_4(w_p)$ est minimum lorsque w_p suit une loi Bernoulli. Par conséquent, les performances de notre schéma seront d'autant meilleures que les lois seront proches de lois de Bernoulli.

Nous avons été en mesure d'effectuer les mêmes calculs pour le deuxième étage de la procédure, i.e. l'algorithme (3.1). Les calculs sont plus compliqués, et la matrice de covariance P_1 des angles $\theta^1(n)$ n'est plus diagonale. En fait, en supposant que (3.1) permette l'extraction de w_{p-1} , on peut

montrer que P_1 est la somme de $-E(w_{p-1}^6)/2c_4(w_{p-1})I$, et d'un terme supplémentaire dû au désajustement de $\theta(n)$ autour de θ_* . On peut en déduire l'expression de la variance de l'erreur de reconstruction normalisée. Elle fait apparaître le terme correspondant à (3.1) si $\theta(n)$ était égal à θ_* , et un terme supplémentaire T , dont l'expression générale dans le cas $p=3$ est $T = -E(w_p^6)/2c_4(w_p)(1+C(\theta_*, \theta_*^1))$ où $C(\theta_*, \theta_*^1)$ est un facteur, dépendant des points de convergence θ_* et θ_*^1 , et compris entre 0 et 4. Notons que T est toujours plus grand que $-E(w_p^6)/2c_4(w_p)$, de sorte que la déflation introduit un supplément de variance incompressible. D'autre part, la valeur maximale prise par T est égale à $-5E(w_p^6)/2c_4(w_p)$, qui peut être supérieure à la variance de (3.1) si $\theta(n)$ était figé à θ_* (i.e. $-E(w_{p-1}^6)/2c_4(w_{p-1})$). Ceci montre que la perte de performance due au désajustement sur $\theta(n)$ peut être prépondérante. On imagine aisément que cet effet sera d'autant plus sensible que le nombre de signaux à extraire sera grand.

4.2 Influence du blanchiment.

L'algorithme de blanchiment que nous avons considéré est une décomposition de Cholevsky adaptative, qui peut être effectuée en résolvant $(p-1)$ problèmes de prédiction linéaire du type

$$\text{Min } E \left(y_k(n) + \sum_{l=1}^{k-1} b_{k,l} y_l(n) \right)^2$$

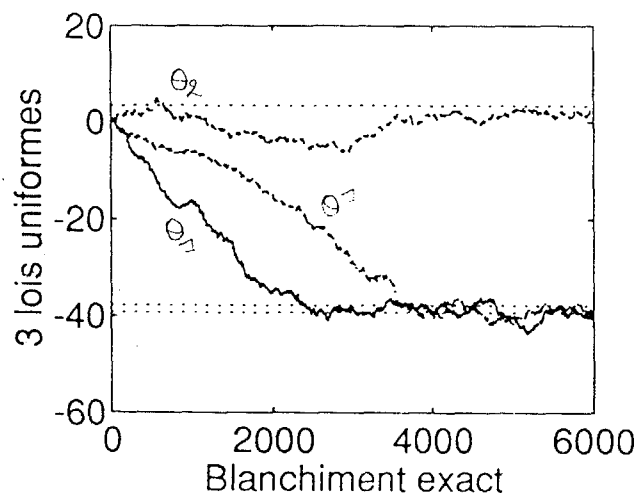
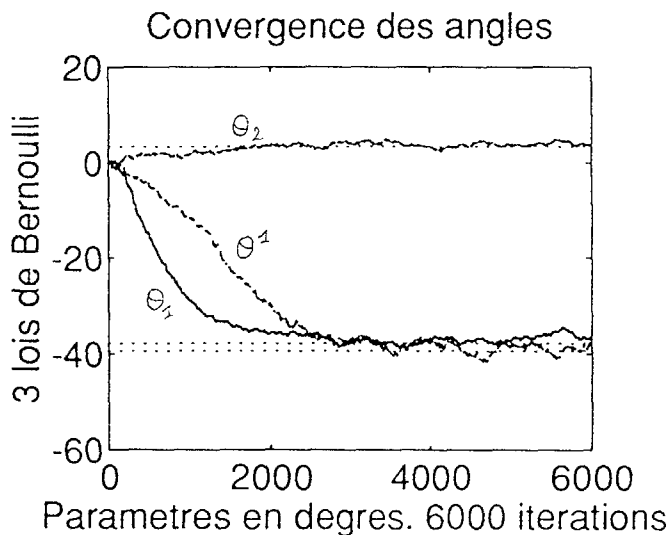
Le schéma global constitué de l'estimation du facteur de Cholevsky et des algorithmes de séparation peut encore se mettre sous la forme (4.1), de sorte que le calcul de variance peut encore s'effectuer comme plus haut. Nous avons calculé explicitement la variance des angles du premier étage (2.3) dans le cas où $p=3$. Puisque les formules explicites sont assez peu informatives, contentons nous de donner quelques indications qualitatives. Le point le plus important est que les variances sont fort peu dépendantes du conditionnement de la matrice de covariance S_y . En effet, l'éventuel supplément de variance dû au blanchiment imparfait nous a paru relativement marginal dans toutes les situations où nous avons l'avons calculé. Par conséquent, notre analyse nous conduit à conclure que le blanchiment imparfait n'a que peu d'effet sur les variances asymptotiques des algorithmes de séparation. Bien entendu, il a une influence, éventuellement néfaste en cas de mauvais conditionnement de S_y , sur la vitesse de convergence du schéma de séparation. Cependant, nous verrons dans les exemples présentés plus loin que, même pour un conditionnement très défavorable, le ralentissement de la convergence du schéma de séparation n'est pas catastrophique.

5 Quelques résultats de simulations.

Pour illustrer ce qui précède, nous allons présenter quelques résultats de simulation. Le nombre de sources p est fixé à 3, de sorte que le schéma de séparation se réduit aux deux algorithmes imbriqués (2.3) et (3.1). (2.3) adapte 2 angles θ_1 et θ_2 , tandis que (3.1) se réduit à l'adaptation d'un angle unique θ^1 . Afin d'illustrer l'influence de la loi des entrées sur les performances, nous avons représenté sur les figures ci-dessous l'évolution des angles dans le cas de 3 lois de Bernoulli et de trois lois uniformes. Dans les deux simulations, le blanchiment est parfait, et le pas des algorithmes a été choisi à $\mu = 0.001$. On constate très clairement que les variances (mais aussi les vitesses de convergence) se dégradent lors du passage loi de Bernoulli, loi uniforme. Ceci est cohérent avec nos évaluations de performances, puisque le terme clé $-E(w^6)/2c_4(w)$ passe de 0.25 à 1.6. Enfin, comme on pouvait s'y attendre, la convergence de l'algorithme (3.1) est plus lente que celle de (2.3) dans les deux cas.



Faute de place, nous ne donnons pas de courbes concernant l'influence du blanchiment, et nous nous contentons de synthétiser les résultats qui sont présentés de façon plus détaillée dans [4]. L'algorithme de blanchiment utilisé est le plus simple possible, à savoir une décomposition de Cholevsky adaptative, dans laquelle les algorithmes de prédiction linéaire sont implantés par l'algorithme LMS. Dans le cas $p = 3$ que nous avons considéré, le nombre de prédictions se réduit à deux. C'est essentiellement sur la vitesse de convergence de l'algorithme minimisant $E(y_3(n) + b_1 y_1(n) + b_2 y_2(n))^2$ que le mauvais conditionnement de S_y peut avoir une influence. Il est en effet bien connu que son comportement dépend de la distribution des valeurs propres de la matrice de covariance de $(y_1(n), y_2(n))^T$. Nous avons donc considéré successivement les cas où le rapport des valeurs propres de cette matrice est 4, 40, 400, et nous avons appliqué le schéma global au cas de 3 signaux sources de lois uniformes. Les divers algorithmes intervenant dans le blanchiment ont été affectés d'un pas de 0.01, tandis que le pas de (2.3) et (3.1) a été choisi à 0.001. On ne constate aucune différence notable en terme de vitesse de convergence entre le cas où le blanchiment est parfait et les cas où le conditionnement de la matrice de covariance de $(y_1(n), y_2(n))^T$ est égal à 4 et 40 (convergence atteinte après environ 2400 échantillons). Par contre, on converge à 3200 échantillons et 4800 échantillons dans les cas où le conditionnement vaut 400 et 4000 respectivement.



6 Conclusion

Nous avons présenté dans cet article un nouveau schéma adaptatif de séparation de sources, basé sur une approche de déflation. Contrairement aux algorithmes existants, il est assuré de converger. Ses performances asymptotiques ont été évaluées analytiquement, et il est apparu qu'elles dépendaient des rapports des moments d'ordre 6 aux cumulants d'ordre 4 des signaux d'entrée. Enfin, on a montré que l'étape préalable de blanchiment adaptatif n'avait que peu d'influence sur les performances des algorithmes.

References

- [1] A. Benveniste, M. Métivier, P. Priouret, "Algorithmes Adaptatifs et Approximations Stochastiques", Masson, 1987.
- [2] P. Comon, "Independent component analysis", Proc. of the Int. Workshop on Higher Order Statistics, J.L. Lacoume Ed., Elsevier, Chamrousse, July 10-12 1991, pp. 29-39.
- [3] B. Laheld, J.F. Cardoso, "Séparation adaptative de sources. Implantation complexe sans contraintes", Quatorzième Colloque GRETSI.
- [4] N. Delfosse, Ph. Loubaton, "Adaptative separation of independent sources: a deflation approach", en préparation.
- [5] L. Féty, "Méthodes de traitement d'antenne adaptées aux radiocommunications", Thèse ENST, 1988.
- [6] H.J. Kushner, D.S. Clark, "Stochastic approximation methods for constrained and unconstrained systems", Applied Mathematical Sciences 26, Springer Verlag, 1978.
- [7] Ph. Loubaton, Ph. Regalia, "Blind deconvolution of multivariate signals by using adaptive FIR lossless filters, Proc. EUSIPCO-92, Bruxelles, August 1992, pp. 1061-1064.
- [8] Ph. Loubaton, Ph. Regalia, "Blind deconvolution of multivariate signals: a deflation approach", Proc. ICC 93, pp.1160-1164, Genève 1993.
- [9] E. Moreau, O. Macchi, "Two novel architectures for the self adaptive separation of signals", Proc. ICC 93, pp. 1154-1159, Genève 1993.
- [10] D.T. Pham, Ph. Garat, C. Jutten, "Separation of mixture of independent sources through a maximum likelihood approach", Proc. EUSIPCO-92, Bruxelles, August 1992, pp. 771-774.
- [11] Ph. Regalia, "An adaptive unit norm filter with applications to signal analysis and Karhunen-Loeve transformation", IEEE. Trans. Circuits and Systems, vol 37, n°5, pp. 646-649, May 1990.
- [12] O. Shalvi, E. Weinstein, "New criteria for blind deconvolution of non-minimum-phase systems (channels)", IEEE Trans. Information Theory, vol.36, n°2, pp. 312-321, March 1990.
- [13] E. Sorouchyari, "Blind separation of sources, Part III : stability analysis", Signal Processing, vol.24, pp.21-29, 1991.
- [14] L. Tong, V.C. Soon, Y.F. Huang, R. Liu, "A new blind identification algorithm", Proc. ISCAS 1990, New Orleans, May 1-3, 1990.