

# Séparation aveugle de sources dépendantes

Marc CASTELLA<sup>1</sup>, Pierre COMON<sup>2</sup>

<sup>1</sup> GET/INT, UMR-CNRS 5157  
9 rue Charles Fourier, 91011 Évry Cedex, France

<sup>2</sup>I3S, UNSA, CNRS UMR 6070  
BP.121, Sophia-Antipolis cedex, France  
marc.castella@int-evry.fr, pcomon@i3s.unice.fr

**Résumé** – Cette communication considère le problème de la séparation aveugle de sources pour un modèle de mélange linéaire instantané. La nouveauté réside dans la considération de sources non statistiquement indépendantes. Nous introduisons trois modèles de sources dépendantes et montrons que leurs cumulants ont des propriétés intéressantes.

Le comportement d'algorithmes conçus pour des sources indépendantes est étudié dans le cas de sources dépendantes: selon le cas, la séparation peut réussir, ou des indéterminations supplémentaires peuvent survenir. A cette occasion, une potentielle confusion est rapidement discutée concernant l'indépendance par paires. Enfin, il est intéressant de noter que les cas réels et complexes se distinguent.

**Abstract** – This paper deals with the problem of blind source separation in the case of a linear instantaneous mixture. The novelty consists in the fact that non independent sources are considered. We introduce three models of dependent sources and show that their cumulants have interesting properties.

The behaviour of algorithms is investigated, when these are designed for independent sources, but applied to non independent sources: depending on the context, the separation may either be successful, or some additional indeterminacies may occur. A possible confusion is shortly discussed about pairwise independence. Finally, it is interesting to note that the real-valued and complex-valued cases differ.

## 1 Introduction

Le problème dit de séparation aveugle de sources a donné lieu à de nombreuses et fructueuses recherches au cours des dernières années, tant au plan méthodologique qu'applicatif. Dans cette problématique, les signaux observés proviennent d'un mélange linéaire de signaux sources. La (quasi-) unanimité des travaux passés suppose l'indépendance mutuelle des différentes sources, ce qui a mené aux techniques dites d'analyse en composantes indépendantes (ou ICA) [7, 2]. En nous inspirant de [4, 5], nous nous attachons à relâcher partiellement l'hypothèse d'indépendance mutuelle. Parmi les rares travaux liés à cette hypothèse, on trouve [1, 8, 6].

## 2 Modèle et notations

Nous considérons un ensemble de  $N$  signaux sources  $(s_i(n))_{n \in \mathbb{Z}, i = 1, \dots, N}$ . Dans ce résumé, la dépendance temporelle n'est pas utilisée et ne sera pas indiquée explicitement. Les sources sont mélangées, produisant un vecteur d'observations  $\mathbf{x} = (\mathbf{x}(n))_{n \in \mathbb{Z}}$  de dimension  $P$  qui suit le modèle :

$$\mathbf{x} = \mathbf{A}\mathbf{s}$$

où  $\mathbf{s} = (s_1, \dots, s_N)^\top$ ,  $\mathbf{x} = (x_1, \dots, x_P)^\top$  et  $\mathbf{A}$  est une matrice  $P \times N$  appelée matrice de mélange. Nous supposons  $\mathbf{A}$  inversible à gauche.

La séparation de sources consiste à rechercher une matrice  $\mathbf{B}$  de taille  $N \times P$ , dite séparante, telle que la sortie

$\mathbf{y} = \mathbf{B}\mathbf{x}$  corresponde aux sources originales. Lorsque seules les observations sont utilisées pour atteindre ce but, le problème est appelé séparation *aveugle* de sources. En notant  $\mathbf{G} \triangleq \mathbf{B}\mathbf{A}$  la matrice globale de taille  $N \times N$ , le problème de séparation est résolu lorsque  $\mathbf{G}$  se résume à une matrice triviale, c.-à-d. le produit d'une matrice diagonale et d'une permutation : ce sont là les ambiguïtés connues de la séparation aveugle.

Dans cette contribution, les critères de séparation seront considérés comme des fonctions de  $\mathbf{G}$ . Les méthodes de séparation peuvent également procéder de façon itérative et extraire les sources une par une (par exemple dans le cas d'une déflation). Nous écrirons alors  $y = \mathbf{b}\mathbf{x} = \mathbf{g}\mathbf{s}$  où  $\mathbf{b}$  et  $\mathbf{g} \triangleq \mathbf{b}\mathbf{A}$  sont des lignes respectivement de  $\mathbf{B}$  et  $\mathbf{G}$ , et  $y$  sera l'unique sortie de la méthode d'extraction. Dans ce cas, les critères de séparation sont considérés comme des fonctions de  $\mathbf{g}$ . Enfin, on notera  $E\{\cdot\}$  l'espérance mathématique,  $\text{Cum}\{\cdot\}$  le cumulants d'un ensemble de variables aléatoires et  $\text{Cum}_4\{y\} \triangleq \text{Cum}\{y, y, y, y\}$ .

## 3 Mélange et sources à valeurs réelles

### 3.1 Séparation de trois sources dépendantes

Nous exhibons un exemple de sources dépendantes qui peuvent être séparées avec succès à l'aide des méthodes

classiques d'ICA fondées sur un blanchiment suivi d'une maximisation d'un contraste à base de cumulants d'ordre quatre. L'exemple comporte des sources binaires BPSK (binary phase shift keying) c.-à-d. qui prennent les valeurs  $+1$  et  $-1$  de façon équiprobable. Considérons un vecteur de sources défini par les hypothèses :

H.1  $\mathbf{s} \triangleq (s_1, s_2, s_3)^\top$  où :  $s_1$  est BPSK ;  $s_2$  est non gaussienne, indépendante de  $s_1$  et satisfait  $E\{s_2\} = E\{s_2^3\} = 0$  ; enfin  $s_3 = s_1 s_2$ .

On a alors le lemme suivant :

**Lemme 1** *Les sources  $s_1, s_2, s_3$  définies par H.1 sont dépendantes mutuellement. Elles sont toutefois décorrélées et leurs cumulants croisés d'ordre quatre sont tous nuls :*

$$\begin{aligned} \text{Cum}\{s_i, s_j\} &= 0 \text{ sauf si } i = j, \\ \text{Cum}\{s_i, s_j, s_k, s_l\} &= 0 \text{ sauf si } i = j = k = l. \end{aligned} \quad (1)$$

*Preuve:* D'après la définition de  $s_1, s_2$  et leur indépendance, on vérifie aisément que  $E\{s_1\} = E\{s_2\} = E\{s_3\} = 0$ . Pour de telles variables centrées, les cumulants s'expriment en fonction des moments :

$$\begin{aligned} \text{Cum}\{s_i, s_j\} &= E\{s_i s_j\} \\ \text{Cum}\{s_i, s_j, s_k, s_l\} &= E\{s_i s_j s_k s_l\} - E\{s_i s_j\} E\{s_k s_l\} \\ &\quad - E\{s_i s_k\} E\{s_j s_l\} \\ &\quad - E\{s_i s_l\} E\{s_j s_k\} \end{aligned} \quad (2)$$

En utilisant alors la définition de  $s_1, s_2$  et leur indépendance, il est aisé de vérifier tous les cas des équations (2) et (3). Par ailleurs, le cumulants d'ordre trois s'écrit :

$$\begin{aligned} \text{Cum}\{s_1, s_2, s_3\} &= E\{s_1 s_2 s_3\} \\ &= E\{s_1^2 s_2^2\} = E\{s_1^2\} E\{s_2^2\} > 0 \end{aligned}$$

ce qui prouve la dépendance mutuelle de  $s_1, s_2, s_3$ . ■  
A l'aide de ces propriétés, il est curieux mais intéressant de noter qu'à l'aide d'algorithmes d'ICA basés sur des cumulants d'ordre quatre, on peut facilement séparer de telles sources, bien que non indépendantes :

**Proposition 1** *Soit  $y = \mathbf{g}\mathbf{s}$  où le vecteur  $\mathbf{s}$  des sources est défini par H.1. La fonction :*

$$\mathbf{g} \mapsto |\text{Cum}_4\{y\}|$$

*définit un contraste MISO (multi input/single output), c.-à-d. que sa maximisation sur l'ensemble des vecteurs de norme unité ( $\|\mathbf{g}\|^2 = 1$ ) mène à une solution  $\mathbf{g}$  dont une et une seule composante est non nulle.*

*Preuve:* La proposition ci-dessus découle immédiatement du lemme 1 : en effet, la démonstration de la validité de tels contrastes est fondée sur les seules propriétés (1) vérifiées par des sources mutuellement indépendantes à l'ordre 4. ■

La proposition ci-dessus se généralise à des contrastes MIMO (multi input/multi output) tels que définis dans [3]. Si  $\mathbf{y} = \mathbf{G}\mathbf{s}$  où  $\mathbf{s}$  est toujours défini par H.1, la maximisation de

$$\mathbf{G} \mapsto \sum_{i=1}^N |\text{Cum}_4\{y_i\}|^2 \quad (4)$$

sur le groupe des matrices orthogonales mène à une matrice  $\mathbf{G}$  triviale (permutation, facteur multiplicatif).

De nombreux algorithmes classiques de séparation de sources (ou ICA) blanchissent d'abord les données, ce qui impose dès lors à la matrice  $\mathbf{G}$  d'être orthogonale. C'est en particulier le cas de l'algorithme dans [3]. Ce dernier est de plus basé sur la fonction de contraste (4). Ceci justifie que l'algorithme puisse séparer des sources données par H.1. En réalité, il en est de même pour tout algorithme basé sur un blanchiment et associé à un critère de contraste basé sur l'annulation des cumulants croisés d'ordre quatre (par exemple JADE [2]).

## 3.2 Séparation de quatre sources BPSK dépendantes

### 3.2.1 Résultat préliminaire

Nous considérons le cas suivant :

H.2  $\mathbf{s} = (s_1, s_2, s_3, s_4)^\top$  où  $s_1, s_2$  et  $s_3$  sont BPSK mutuellement indépendantes et  $s_4 = s_1 s_2 s_3$ .

Ce cas a été considéré dans [4] et il a été montré que l'on a pour de telles sources :

$$\begin{cases} \text{Cum}\{s_i, s_i, s_i, s_i\} = -2, \forall i \in \{1, \dots, 4\} \\ \text{Cum}\{s_1, s_2, s_3, s_4\} = 1 \\ \text{Tous les autres cumulants croisés sont nuls.} \end{cases} \quad (5)$$

De plus, ces sources sont indépendantes par paires, mais pas mutuellement comme le prouve (5). Il est donc clair que ces deux propriétés ne sont pas équivalentes et il est intéressant de rappeler la propriété suivante qui est une conséquence directe du théorème de Darmois [3, p.294] :

**Propriété 1** *Soit  $\mathbf{s}$  un vecteur aléatoire dont les composantes sont mutuellement indépendantes et  $\mathbf{x} = \mathbf{G}\mathbf{s}$ . Alors l'indépendance mutuelle des entrées  $x_i$  est équivalente à leur indépendance par paires.*

Grâce à cette propriété, l'algorithme d'ICA proposé dans [3] fonctionne en se contentant d'imposer l'indépendance par paires. Il faut toutefois souligner que ceci ne vaut que dans la mesure où le vecteur de départ (c.-à-d.  $\mathbf{s}$ , les sources) est à composantes *mutuellement indépendantes*. En effet, la proposition suivante illustre le danger d'une telle confusion : elle explique que la séparation n'est pas toujours réalisée par l'algorithme donné dans [3] dans le cas d'un mélange dont les sources  $\mathbf{s}$  sont uniquement indépendantes par paires.

**Proposition 2** *Soit  $y = \mathbf{g}\mathbf{s}$  où le vecteur des sources est défini par H.2. Supposons que le vecteur  $(s_1, s_2, s_3)$  prenne ses  $2^3$  valeurs possibles. Si le signal  $y$  est tel que ses valeurs soient dans  $\{-1, +1\}$ , alors  $\mathbf{g} = (g_1, g_2, g_3, g_4)$  est l'une des solutions :*

$$\begin{cases} \exists i \in \{1, \dots, 4\} & g_i = \pm 1, \text{ et } : \forall j \neq i, g_j = 0 \\ \exists i \in \{1, \dots, 4\} & g_i = \pm 1/2, \text{ et } : \forall j \neq i, g_j = -g_i \end{cases} \quad (6)$$

*Preuve:* Si  $y = \mathbf{g}\mathbf{s}$ , en utilisant le fait que  $s_i^2 = 1$  pour  $i = 1, \dots, 4$ , on a dans le cas des sources données par H.2 :

$$y^2 = g_1^2 + g_2^2 + g_3^2 + g_4^2 + 2 \left[ (g_1g_2 + g_3g_4)s_1s_2 + (g_1g_3 + g_2g_4)s_1s_3 + (g_2g_3 + g_1g_4)s_2s_3 \right]$$

Comme  $(s_1, s_2, s_3)$  prend toutes les valeurs possibles de  $\{-1, 1\}^3$ , il résulte de  $y^2 = 1$  que les équations suivantes sont nécessairement vérifiées :

$$\begin{cases} g_1^2 + g_2^2 + g_3^2 + g_4^2 = 1 \\ g_1g_2 + g_3g_4 = g_1g_3 + g_2g_4 = g_2g_3 + g_1g_4 = 0 \end{cases} \quad (7)$$

On constatera d'abord que les valeurs de (6) satisfont (7). Par ailleurs, un système polynomial de  $N$  équations de degré  $d$  avec  $N$  inconnues, s'il admet un nombre fini de solutions (ce qui est le cas ici), en a au plus  $d^N$ . Ainsi, les 16 solutions données par (6) correspondent à l'ensemble des valeurs possibles pour le quadruplet  $(g_1, g_2, g_3, g_4)$ . ■

A l'aide du lien établi dans le paragraphe suivant avec les critères de contraste classiques, cette proposition permet d'expliquer pourquoi la maximisation par paires de l'algorithme [3] fonctionne pour ces sources et pour certaines matrices de mélange telles que celles étudiées dans [4]. En réalité, dans le cas général d'un mélange quelconque, cet algorithme ne fonctionne pas pour séparer des sources BPSK qui seraient uniquement indépendantes par paires.

### 3.2.2 Résultats de séparation pour des critères de contraste connus

Le critère du module constant (CM) est l'un des plus connus pour la séparation aveugle de sources. Dans le cas réel, il s'écrit :

$$J_{\text{CM}}(\mathbf{g}) \triangleq \mathbb{E} \left\{ (y^2 - 1)^2 \right\} \quad \text{avec : } y = \mathbf{g}\mathbf{s}$$

Ce critère admet zéro pour valeur minimale, qui est atteinte lorsque  $\mathbf{g}$  a une composante à  $\pm 1$  et les autres nulles. De plus, en cas d'annulation de  $J_{\text{CM}}(\mathbf{g})$  alors on a l'égalité  $y^2 - 1 = 0$  presque sûrement et la proposition 2 s'applique. Il résulte de cette discussion :

**Proposition 3** *Pour un vecteur de sources donné par H.2, la minimisation du critère du module constant par rapport à  $\mathbf{g}$  mène à l'une des solutions données par l'équation (6).*

De plus, un lien avec le contraste lié à l'auto-cumulant d'ordre quatre peut être établi en remarquant que, sous la condition  $\mathbb{E} \{ y^2 \} = 1$ , on a  $\text{Cum}_4\{y\} = \mathbb{E} \left\{ (y^2 - 1)^2 \right\} - 2$ . Ainsi, la minimisation du critère de module constant est équivalente à la maximisation sur l'ensemble des vecteurs ligne de norme unité ( $\|\mathbf{g}\|^2 = 1$ ) du critère

$$\mathbf{g} \mapsto -\text{Cum}_4\{y\}.$$

Etablir le résultat correspondant pour le critère  $\mathbf{g} \mapsto |\text{Cum}_4\{y\}|$  mérite une légère attention car le cumulants n'est pas de signe constant en raison de la non indépendance des sources. On a toutefois :

**Proposition 4** *Soit  $y = \mathbf{g}\mathbf{s}$  où  $\mathbf{s}$  est défini par H.2. Alors sous la contrainte  $\|\mathbf{g}\| = 1$ , on a :*

- (i) *La maximisation de  $\mathbf{g} \mapsto -\text{Cum}_4\{y\}$  donne l'une des solutions de (6).*
- (ii)  *$|\text{Cum}_4\{y\}| \leq 2$  et de plus l'égalité  $|\text{Cum}_4\{y\}| = 2$  n'est atteinte que si et seulement si  $\mathbf{g}$  est l'une des solutions de (6).*

*Preuve:* (i) résulte de la proposition 3 et des arguments concernant l'équivalence avec le critère du module constant. Par ailleurs, en utilisant la multilinéarité des cumulants et (5), il vient :

$$\text{Cum}_4\{y\} = -2(g_1^4 + g_2^4 + g_3^4 + g_4^4) + 24(g_1g_2g_3g_4) \quad (8)$$

L'étude de la fonction polynomiale ci-dessus permet alors de tirer les conclusions souhaitées. En effet, l'optimisation de (8) mène au Lagrangien suivant :

$$\mathcal{L} = -2 \sum_{i=1}^4 g_i^4 + 24 \prod_{i=1}^4 g_i - \lambda \left( \sum_{i=1}^4 g_i^2 - 1 \right) \quad (9)$$

La recherche des points d'annulation permet de vérifier que toutes les solutions satisfont  $|\text{Cum}_4\{y\}| \leq 2$ . La proposition découle alors facilement. ■

## 4 Mélange et sources à valeurs complexes

Nous considérons maintenant le cas d'une matrice de mélange et de sources à valeurs complexes. Nous nous basons sur le cas de sources QPSK, c.-à-d. qui prennent leurs valeurs de façon équiprobable dans l'ensemble  $\{e^{i\frac{\pi}{4}}, e^{-i\frac{\pi}{4}}, e^{i\frac{3\pi}{4}}, e^{-i\frac{3\pi}{4}}\}$ . De telles sources permettent d'illustrer la différence de comportement entre les cas à valeurs complexes et réelles. Introduisons le vecteur de sources suivant, dont les hypothèses sont proches de H.2 :

H.3  $\mathbf{s} = (s_1, s_2, s_3, s_4)^T$  où  $s_1, s_2$  et  $s_3$  sont QPSK mutuellement indépendantes et  $s_4 = s_1s_2s_3$ .

De telles sources présentent la particularité de vérifier :

**Lemme 2** *Les sources H.3 sont mutuellement dépendantes et vérifient :  $\text{Cum}\{s_1, s_2, s_3, s_4^*\} = 0$ . Elles sont décorrélatées au second ordre et leurs cumulants croisés circulaires (c.-à-d. avec autant de conjugués que de non conjugués) d'ordre quatre sont nuls :*

$$\begin{cases} \text{Cum}\{s_i, s_j^*\} = 0 \text{ et } \text{Cum}\{s_i, s_j\} = 0 \text{ sauf si } i = j, \\ \text{Cum}\{s_i, s_j, s_k^*, s_l^*\} = 0 \text{ sauf si } i = j = k = l. \end{cases}$$

Les propriétés précédemment démontrées dans le cas réel s'adaptent alors de la façon suivante :

**Proposition 5** *Soit  $y = \mathbf{g}\mathbf{s}$  où le vecteur des sources est défini par H.3. Supposons que le vecteur  $(s_1, s_2, s_3)$  prenne ses  $4^3$  valeurs possibles. Si le signal  $y$  est tel que ses valeurs vérifient  $|y|^2 = 1$ , alors  $\mathbf{g} = (g_1, g_2, g_3, g_4)$  est tel que :*

$$\exists i \in \{1, \dots, 4\} \quad |g_i| = 1, \text{ et } \forall j \neq i, g_j = 0 \quad (10)$$

*Preuve:* Si  $y = \mathbf{g}\mathbf{s}$ , en utilisant le fait que  $|s_i|^2 = 1$  pour  $i = 1, \dots, 4$ , il vient pour les sources H.3 :

$$|y|^2 = \sum_{i=1}^4 |g_i|^2 + \sum_{i \neq j} g_i g_j^* s_i s_j^*$$

Comme  $(s_1, s_2, s_3)$  prend toutes les valeurs de  $\{e^{i\frac{\pi}{4}}, e^{-i\frac{\pi}{4}}, e^{i\frac{3\pi}{4}}, e^{-i\frac{3\pi}{4}}\}$ , la condition  $|y|^2 = 1$  entraîne que les équations suivantes sont nécessairement vérifiées :

$$\begin{cases} |g_1|^2 + |g_2|^2 + |g_3|^2 + |g_4|^2 = 1 \\ g_1 g_2^* = g_1 g_3^* = g_1 g_4^* = g_2 g_3^* = g_2 g_4^* = g_3 g_4^* = 0 \end{cases}$$

La résolution de ce système en les variables  $|g_1|, |g_2|, |g_3|$  et  $|g_4|$ , donne les solutions indiquées à l'équation (10). ■

Il est particulièrement intéressant de noter que ce cas se distingue nettement du cas à valeurs réelles : en effet, les solutions indésirables de l'équation (6) n'existent plus dans ce cas en raison des symétries particulières des sources QPSK. Enfin, il est possible d'établir la propriété suivante concernant deux critères de séparation usuels :

**Proposition 6** Soit  $y = \mathbf{g}\mathbf{s}$  où les sources vérifient H.3. Un vecteur  $\mathbf{g}$  séparant (c.-à-d. qui vérifie (10)) est obtenu à l'issue de l'une ou l'autre des opérations :

- minimisation du critère de module constant :

$$\mathbf{g} \mapsto E \left\{ \left| |y|^2 - 1 \right|^2 \right\}$$

- maximisation sous la contrainte  $\|\mathbf{g}\| = 1$  du critère :

$$\mathbf{g} \mapsto |\text{Cum} \{y, y^*, y, y^*\}|$$

*Preuve:* Les mêmes arguments que ceux de la proposition 3 permettent d'affirmer que la minimisation du critère du module constant permet de séparer les sources complexes données par H.3. La deuxième partie provient du même argument que celui donné dans la proposition 1 : dans le cas de sources mutuellement indépendantes, seules les propriétés du lemme 2 sont utilisées pour la démonstration de la validité du critère de contraste en question. ■

On pourra remarquer que les arguments de la proposition ci-dessus s'appliquent également pour des algorithmes et des critères de séparation MIMO. En particulier, l'algorithme par paires [3] ou bien l'algorithme JADE [2] permettent de séparer de telles sources. Une illustration typique est fournie par la figure 1.

## Références

- [1] J.-F. Cardoso. Multidimensional independent component analysis. In *Proc. ICASSP '98. Seattle*, 1998.
- [2] J.-F. Cardoso and A. Souloumiac. Blind beamforming for non gaussian signals. In *IEEE- Proceedings-F*, volume 140, pages 362–370, 1993.
- [3] P. Comon. Independent component analysis, a new concept. *Signal Processing*, 36(3) :287–314, April 1994.
- [4] P. Comon. Blind identification and source separation in  $2 \times 3$  under-determined mixtures. *IEEE Trans. Signal Processing*, 52(1) :11–22, January 2004.

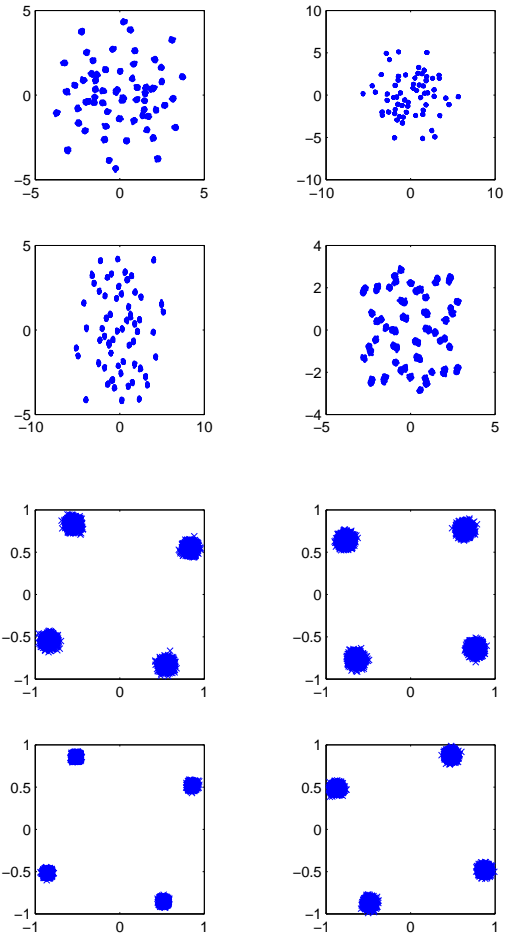


FIG. 1 – Résultat typique d'une séparation par l'algorithme JADE d'un mélange de sources définies par H.3. Constellation des observations (haut) et des sources reconstruites (bas).

- [5] P. Comon and O. Grellier. Non-linear inversion of underdetermined mixtures. In *Proc. of ICA '99*, pages 461–465, Aussois, France, January 1999.
- [6] A. Hyvärinen and S. Shimizu. A quasi-stochastic gradient algorithm for variance-dependent component analysis. In *Proc. International Conference on Artificial Neural Networks (ICANN2006)*, pages 211–220, Athens, Greece, 2006.
- [7] C. Jutten and J. Herault. Blind separation of sources, part i : An adaptative algorithm based on neuromimetic architecture. *Signal Processing*, 24(1) :1–10, 1991.
- [8] T.-H. Li. Finite-alphabet information and multivariate blind deconvolution and identification of linear systems. *IEEE Trans. on Information Theory*, 49(1) :330–337, January 2003.