

Critères de Contraste Déterministes pour la Séparation de Sources

Pierre COMON, Jérôme LEBRUN

Laboratoire I3S, Les Algorithmes/Euclide/B
2000 route des Lucioles, BP 121, 06903 Sophia-Antipolis Cedex, France.
comon@i3s.unice.fr, lebrun@i3s.unice.fr

Résumé – La séparation aveugle de sources est souvent menée sous l’hypothèse que les sources sont statistiquement indépendantes. Cependant, cette hypothèse n’est pas indispensable, et peut être remplacée par une hypothèse sur la distribution des sources, même si les sources sont identiquement distribuées. Les fonctions de contraste sont des critères d’optimisation qui jouissent de certaines propriétés d’identifiabilité. On définit ici une famille de fonctions de contraste adaptées aux sources discrètes. Parmi les avantages présentés par ces contrastes, on peut mentionner la possibilité de traiter des sources corrélées, ou déterministes.

Abstract – Blind Source Separation is often carried out under the assumption that sources are statistically mutually independent. However, this assumption is not mandatory, and can be replaced by some assumption on the source distribution, even if sources are identically distributed. Contrast functions are optimization criteria that enjoy identifiability properties. A family of contrast functions is defined herein, which is dedicated to discrete sources. Among others, these contrasts have the advantage to be applicable to correlated, or deterministic sources.

1 Position du problème

L’égalisation et l’identification aveugles ont fait l’objet d’un intérêt croissant depuis 1975. Un des avantages principaux des techniques aveugles, en télécommunications par exemple, est qu’en supprimant les séquences pilotes, on peut augmenter le débit. Mais il existe d’autres avantages provenant des limitations des techniques classiques. En effet, les techniques exploitant des séquences pilotes sont difficiles à mettre en oeuvre lorsque les canaux sont de réponse très longue ; la présence d’un résidu de porteuse peut aussi compliquer l’égalisation. Au lieu d’exploiter des relations entrée-sortie pour identifier le canal (approches utilisant des séquences pilotes), on utilise une ou plusieurs propriétés des entrées (approches dites aveugles).

1.1 Modèle d’observation.

On se limite dans cet article aux signaux linéairement modulés. L’enveloppe complexe d’un signal émis $s(t)$ prend la forme suivante en bande de base [20] : $s(t) = \sum_k g(t - kT) s[k]$. On remarquera que la distinction entre processus à temps discret et à temps continu est faite à l’aide de crochets et de parenthèses. Après propagation dans le canal et le filtre de réception, le signal reçu sur l’antenne s’écrit :

$$y(t) = \sum_k h(t - kT_s) s[k]$$

où h est la convolution entre le filtre d’émission, le canal, et le filtre de réception. Si le signal reçu est échantillonné à la cadence $1/T_s$, alors :

$$y[n] \stackrel{\text{def}}{=} y(nT_s) = \sum_k h[n - k] s[k] \quad (1)$$

avec $h[k] \stackrel{\text{def}}{=} T_s h(kT_s)$. Pour des systèmes à entrées et sorties multiples (MIMO), le signal émis $s[k]$ et le signal reçu $y[k]$

peuvent être considérés comme des processus multivariés, et sont notés en gras ; leur dimension sera notée P et K , respectivement. Le modèle (1) peut se réécrire

$$\mathbf{y}[n] = \sum_k \mathbf{H}[n - k] \mathbf{s}[k] \quad (2)$$

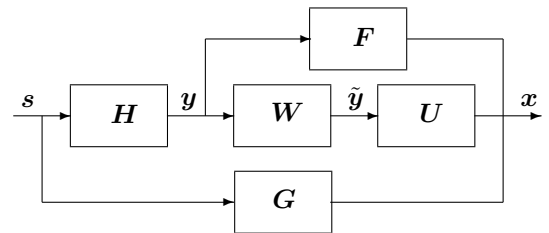
où la réponse impulsionnelle globale $\mathbf{H}[k]$ est maintenant une suite de matrices $K \times P$. Sa transformée en z sera notée

$$\check{\mathbf{H}}[z] \stackrel{\text{def}}{=} \sum_k \mathbf{H}[k] z^{-k}$$

Dans le contexte qui nous occupe, les entrées $s_j[k]$ sont appelées *sources*.

1.2 Hypothèses et taxinomie.

Le but de l’égalisation aveugle est de délivrer une estimée des séquences d’entrée $s_j[k]$ à partir des seules séquences observées $y_i[n]$.



La séquence émise $s[n]$ traverse le canal $\check{\mathbf{H}}[z]$, puis est déconvoluée par l’égaliseur \mathbf{F} pour donner $x[n]$.

L’hypothèse la plus communément utilisée pour l’égalisation aveugle est l’*indépendance statistique* entre les symboles successifs :

Hypothèse H1 Les sources $s_j[k]$ sont toutes des séquences i.i.d.

Pour des systèmes MIMO, l'indépendance statistique entre sources est souvent utilisée :

Hypothèse H2 Les sources $s_j[k]$ sont mutuellement statistiquement indépendantes.

Ces hypothèses peuvent être généralement affaiblies à des propriétés d'indépendance ou de blancheur moins fortes, de par l'utilisation de moments/cumulants d'ordre fini [7]. Le cas où les sources sont des *processus linéaires* est couvert par l'hypothèse H1, à cause des indéterminations inhérentes au problème (cf. section 2.1) ; cette dernière n'est donc pas très restrictive. Le cas de sources gaussiennes soulève néanmoins un problème. Dans ce cas, toute l'information est contenue dans les moments d'ordre 2, ce qui n'est pas suffisant pour établir l'identifiabilité. Pour cette raison, on a besoins de l'hypothèse suivante, en plus des hypothèses H1 et H2 :

Hypothèse H3 Au plus une source est gaussienne.

C'est par exemple sous les hypothèses H2 et H3 que l'*Analyse en Composantes Indépendantes* (ACI) a été développée [2] [19] [15] [14] [5] ; voir aussi [8] [2] [18] [6].

Il existe bien d'autres approches du problème, basées sur des hypothèses différentes, qui ne seront pas abordées dans le présent article, et donnent lieu à des algorithmes théoriquement plus simples. En revanche, on s'intéresse ici plus particulièrement aux sources discrètes, cas pertinent en communications numériques. on admettra donc les hypothèses suivantes, à la place des hypothèses H1 à H3 :

Hypothèse H4 Les sources $s_j[n]$ appartiennent à un alphabet fini \mathcal{A} caractérisé par les d racines complexes distinctes d'un polynôme de degré d , $Q(z)$

Hypothèse H5 Les sources $s_j[n]$ sont "suffisamment excitantes".

Par *suffisamment excitantes*, on veut dire que suffisamment d'états différents du P -uplet \mathbf{s} apparaissent dans la matrice de données. Par exemple, pour des alphabets généraux, il suffit que toutes les séquences binaires soient présentes ; il y en a $C_d^2 2^P$, ce qui donne une borne minimale à la durée d'observation. Pour les alphabets PSK, la durée d'observation peut être réduite à $\lceil \frac{d}{2} \rceil 2^P$, où $\lceil \cdot \rceil$ désigne l'arrondi à l'entier supérieur. Cette condition peut sans doute être affaiblie.

Suivant les hypothèses que l'on adopte (indépendance, caractère discret, SISO/MIMO, $P > K$ ou non...), toute une variété de problèmes peut être étudiée [7]. A partir de maintenant, nous allons nous consacrer essentiellement aux hypothèses H4 et H5, sauf mention contraire, et nous limiter au cas où $P \leq K$ (mélanges sur-déterminés). Lorsque plus de sources que de capteurs sont présentes, le problème devient plus compliqué (mélanges sous-déterminés), et des outils spécifiques doivent être utilisés. En particulier, le mélange ne peut généralement pas être linéairement inversé ; les approches par contrastes n'ont pour l'instant pas pu être mises au point pour cette raison.

Le caractère discret a déjà été utilisé pour la séparation de sources dans [12] [25] [23] [21] [22] [24] [17] [9] [16] [13],

mais pas dans le cadre de maximisation de contrastes. Pour les mélanges sous-déterminés, nous renvoyons aux références citées dans [5].

2 Contrastes

2.1 Filtres triviaux.

Le filtre séparateur, $\check{F}[z]$, s'il existe, vise à délivrer une sortie $\mathbf{x}[n]$, qui doit satisfaire aussi précisément que possible les hypothèses **Hi**. Mais il est clair qu'il existe certains filtres qui ne les affectent pas. Ces filtres sont appelés les *filtres triviaux*. Par exemple, sous les hypothèses H1 à H3, on peut montrer que

Proposition 1 Sous les hypothèses H1 à H3, les filtres triviaux sont de la forme $\check{T}[z] = \mathbf{P} \check{D}[z]$, où \mathbf{P} est une permutation et $\check{D}[z]$ un filtre diagonal. De plus, à cause de la propriété i.i.d. (hypothèse H1), les composantes de $\check{D}[z]$ doivent être de la forme $\check{D}_{pp}[z] = \lambda_p z^{d_p}$, où d_p est un entier.

Il est donc illusoire de vouloir estimer le couple $(\check{H}[z], \mathbf{s}[k])$. Il faut plutôt estimer un représentant de la classe d'équivalence des solutions : une fois une solution trouvée, la classe entière de solutions peut être générée par filtrage trivial.

2.2 Définition des fonctionnelles de contraste

Soit \mathcal{H} un ensemble de filtres, et $\mathcal{H} \cdot \mathcal{S}$ l'ensemble des processus obtenus par action des filtres de \mathcal{H} sur les processus source de \mathcal{S} . Notons \mathcal{T} le sous-ensemble de \mathcal{H} des filtres triviaux, défini dans la proposition 1. Un critère d'optimisation $\Upsilon(\mathbf{H}; \mathbf{x})$, sera dit contraste défini sur $\mathcal{H} \times \mathcal{H} \cdot \mathcal{S}$, s'il satisfait les trois propriétés suivantes [3] :

P1 Invariance : Le contraste ne doit pas changer à l'intérieur de l'ensemble des solutions acceptables, c.à.d :

$$\forall \mathbf{H} \in \mathcal{T}, \forall \mathbf{x} \in \mathcal{H} \cdot \mathcal{S}, \Upsilon(\mathbf{H}; \mathbf{x}) = \Upsilon(\mathbf{I}; \mathbf{x}).$$

P2 Domination : Si les sources sont déjà séparées, l'action de tout filtre doit faire diminuer le contraste :

$$\forall \mathbf{s} \in \mathcal{S}, \forall \mathbf{H} \in \mathcal{H}, \text{ then } \Upsilon(\mathbf{H}; \mathbf{s}) \leq \Upsilon(\mathbf{I}; \mathbf{s}).$$

P3 Discrimination : Le contraste maximal ne doit être atteint que pour des solutions liées entre elles par filtrage trivial :

$$\forall \mathbf{s} \in \mathcal{S}, \Upsilon(\mathbf{H}; \mathbf{s}) = \Upsilon(\mathbf{I}; \mathbf{s}) \Rightarrow \mathbf{H} \in \mathcal{T}.$$

Par exemple, le critère le plus naturel pour mesurer l'indépendance statistique entre P variables aléatoires z_p est la divergence de Kullback entre la distribution de probabilité conjointe et le produit des distributions marginales [2]. Si on adopte la divergence de Kullback-Leibler, on aboutit alors à l'Information Mutuelle (IM) [2]. L'IM est donc un critère de contraste possible. Cependant, il est difficile à utiliser parce que difficile à estimer, surtout en grande dimension (e.g. mélanges convolutifs), même si certains algorithmes itératifs ont été proposés sur cette base [1]. C'est pourquoi on préfère les contrastes statistiques basés sur les cumulants.

2.3 Alphabets finis

En revanche, sous les hypothèses H4 et H5, il serait assez naturel de définir le critère d'optimisation suivant :

$$\Upsilon_P(\mathbf{x}) = - \sum_n \sum_i |Q(x_i(n))|^2. \quad (3)$$

Sous des hypothèses pas trop fortes, il s'avère que le critère (3) est un contraste, comme nous allons le montrer. L'intérêt d'exploiter le caractère discret réside non seulement dans le fait que les caractéristiques de la sortie attendue sont plus finement exploitées, mais aussi dans le fait que d'autres hypothèses peuvent alors être évacuées. Par exemple dans [9], les auteurs exploitent le caractère discret, mais conservent l'hypothèse d'indépendance des sources. Dans cette section, seules les hypothèses H4 et H5 sont utilisées.

Les sources peuvent notamment être corrélées et non stationnaires. En effet, l'approche proposée est entièrement algébrique et déterministe, de sorte qu'aucun outil statistique n'est requis.

Définition 2 Soient \mathcal{A} un alphabet fini dans le plan complexe, de cardinal d , et \mathcal{G} l'ensemble de nombres complexes γ tels que $\gamma \mathcal{A} \subset \mathcal{A}$.

Un premier résultat immédiat est que [6] :

Lemme 3 comme \mathcal{A} est fini, alors \mathcal{G} ne contient que des nombres de module unité.

Lemme 4 les filtres triviaux associés aux hypothèses H4 et H5 sont de la forme $P\check{D}[z]$, où les composantes de $\check{D}[z]$ s'écrivent $\check{D}_{pp}[z] = \gamma z^n$, avec $\gamma \in \mathcal{G}$ et $n \in \mathbb{Z}$.

Théorème 5 Soient \mathcal{S} l'ensemble des processus prenant leurs valeurs dans l'alphabet \mathcal{A} , et \mathcal{H} l'ensemble des filtres $P \times P$ à réponse finie inversibles (il est non vide pour $P > 1$). Alors le critère (3) est un contraste sous les hypothèses H4 et H5.

Preuve. Pour des raisons de place, la démonstration est condensée; nous renvoyons à [6] pour une version plus détaillée. Supposons que, pour un certain $\mathbf{c} \in \mathbb{C}^P$, nous avons $\mathbf{x}^T \mathbf{c} \in \mathcal{A}$, $\forall \mathbf{x} \in \mathcal{A}^P$. Notons $\mathcal{A} = \{a_1, \dots, a_d\}$, $\mathbf{c}^T = [c_1, \dots, c_P]$, et $\mathbf{1}$ le vecteur formé de P uns.

Tout d'abord, d'après le lemme 3, $\sum_i c_i \in \mathcal{G}$ et il existe un nombre $\gamma \in \mathcal{G}$ tel que $\gamma \sum_i c_i = 1$. En posant $c'_i = \gamma c_i$, il est suffisant de considérer le cas $\mathcal{G} = \{1\}$.

- Comme $\mathbf{x}^T \mathbf{c} \in \mathcal{A}$ pour tout vecteur \mathbf{x} contenant des éléments de \mathcal{A} , cela reste vrai en particulier pour $\mathbf{x}^T = a_p \mathbf{1}^T$. Cela implique que $\forall a_p \in \mathcal{A}$, $a_p \sum_i c_i \in \mathcal{A}$. Autrement dit, on a toujours

$$\sum_i c_i \in \mathcal{G}. \quad (4)$$

- Ensuite, pour tout couple de nombre complexes (a, b) , on définit la matrice $P \times P$ suivante : $\mathbf{B} = (a-b)\mathbf{I} + b\mathbf{1}\mathbf{1}^T$. Cette matrice a un déterminant égal à $(a-b)^{P-1}(a+(P-1)b)$. Pour $a+(P-1)b \neq 0$, son inverse s'écrit $\mathbf{B}^{-1} = [a-b]^{-1}[\mathbf{I} - b\mathbf{1}\mathbf{1}^T/[a+(P-1)b]]$.

- En conséquence, pour tout couple de symboles distincts de \mathcal{A} , (a, b) , il existe un vecteur α contenant P symboles de \mathcal{A} tel que $\mathbf{B}\mathbf{c} = \alpha$. D'après ce qui précède, on obtient en particulier que

$$\sum_i c_i = [a+(P-1)b]^{-1} \sum_i \alpha_i \quad (5)$$

Cas de symboles réels, avec $\mathcal{G} = \{1\}$. Notons $x_m = \min\{x, x \in \mathcal{A}\}$ et $x_M = \max\{x, x \in \mathcal{A}\}$. D'après (4),

$$\forall a, b \in \mathcal{A}, \exists \alpha_i \in \mathcal{A} / \alpha_i = c_i a + (1 - c_i) b$$

Donc c_i doit être réel aussi. On peut montrer facilement [6] que $c_i \in \{-1, 0, 1\}$, $\forall i$, $1 \leq i \leq P$, sinon \mathcal{A} serait infini. Si $c_i \in \{0, 1\}$, $\forall i$, alors, encore d'après $\sum_i c_i = 1$, il n'existe qu'une composante non nulle dans \mathbf{c} , et \mathbf{c} est finalement trivial. Supposons donc que $\exists c_i = -1$, $1 \leq i \leq P$. Mais alors $\beta = b + c_i(a-b) \in \mathcal{A}$ pour tout couple $(a, b) \in \mathcal{A}^2$; en particulier pour $c_i = -1$, $a = x_M$ et $b = x_m$, le symbole $\beta = x_m - (x_M - x_m)$ devrait appartenir à \mathcal{A} . Et $x_M > x_m \Leftrightarrow \beta < x_m$, ce qui contredit la définition de x_m .

Cas de symboles complexes avec $\mathcal{G} = \{1\}$. Si $d = 2$, le problème se ramène à un cas particulier d'alphabet réel (moyennant une rotation), cas déjà traité. Supposons donc $d > 2$, et choisissons un symbole b sur l'enveloppe convexe de \mathcal{A} . Comme $d > 2$, b a toujours deux voisins distincts sur l'enveloppe convexe. Choisissons un des deux voisins, noté a , de façon à aussi avoir $a + (P-1)b \neq 0$. Le résultat (5) s'applique alors. Comme $\sum_i c_i \in \mathcal{G}$, (5) implique que

$$\frac{1}{P} [a + (P-1)b] = \frac{1}{P} \sum_i \alpha_i \quad (6)$$

Montrons d'abord par l'absurde que α ne peut pas être proportionnel à $\mathbf{1}$. Supposons donc $\alpha = a_o \mathbf{1}$ pour un certain $a_o \in \mathcal{A}$. Alors d'après (6), $a_o = \frac{1}{P} a + \frac{P-1}{P} b$. Donc le symbole a_o est aussi sur l'enveloppe convexe de \mathcal{A} , et est plus près de b que a ne l'était. Ceci contredit l'hypothèse selon laquelle a était un des deux plus proches voisins de b .

Supposons maintenant que α contient au moins deux symboles distincts. S'il n'y a que deux symboles distincts, notés a et b , alors on a nécessairement une fois a et $(P-1)$ fois b (ou le contraire), et \mathbf{c} est trivial, comme nous l'avons déjà montré. Si tous les symboles de α sont réels, alors on est ramené au premier cas étudié précédemment. Supposons donc que le vecteur α contient un troisième symbole x distinct de a et b , sans être une combinaison linéaire de a et b . D'après (5), il doit alors exister au moins un autre symbole x' de l'autre côté de la ligne engendrée par $\{a, b\}$. Mais alors l'un d'entre eux se trouverait à l'extérieur de l'enveloppe convexe de \mathcal{A} . Cela contredit le fait que x et x' sont tous deux dans \mathcal{A} . \square

2.4 Contrastes PSK

Le cas le plus simple d'alphabet discret est probablement celui défini par l'équation polynomiale $P(z) = z^q - 1$. Ce sont des sources dites PSK- q .

Théorème 6 Soit \mathcal{S} l'ensemble des processus PSK- q , et \mathcal{H} l'ensemble des filtres inversibles $P \times P$ à réponse impulsionnelle finie. Alors

$$\Upsilon(\mathbf{x}) \stackrel{\text{def}}{=} - \sum_i \sum_n |x_i[n]^q - 1|^2$$

est un contraste sous la condition H5 que la source s soit suffisamment excitante.

Dans le cas PSK, les filtres triviaux sont de la forme $P\check{D}[z]$, où $\check{D}_{pp}[z]$ sont des rotations dans le plan complexe d'un angle multiple de $2\pi/q$, combinées avec des retards purs, et P des permutations.

3 Conclusion

La famille de contrastes que nous avons définie permet d'exécuter une égalisation aveugle MIMO lorsque les sources sont discrètes, d'une manière déterministe. Comme on n'estime pas explicitement de moments des observations, on peut s'attendre à une convergence rapide de l'algorithme de séparation, même pour des sources corrélées. On peut aussi s'attendre à de meilleures performances de séparation que les algorithmes CMA ou KMA lorsque les sources sont discrètes mais pas distribuées sur le cercle (e.g. QAM16). Ceci doit toutefois être confirmé par des simulations informatiques.

Il reste également plusieurs pistes à poursuivre. La première concerne le calcul de solutions algébriques, de plus en plus attrayantes dans un contexte de communications TDMA (algorithmes par blocs, ou en ligne). La deuxième concerne la prise en compte de l'estimation conjointe de la période symbole, problème important lorsque les techniques aveugles sont utilisées dans un but d'interception. Dans ce cas, si la cadence d'échantillonnage est de $1/T'$ à la réception, l'observation s'écrit comme la convolution suivante $y[n] = \sum_k \mathbf{H}(nT' - kT_s) s[k]$, qui n'est plus une *convolution discrète*. La troisième piste d'intérêt est l'estimation conjointe du résidu de porteuse, qui peut être représenté par une exponentielle multiplicative en bande de base. Par exemple pour un canal SISO :

$$y[n] = \sum_k h[n - k] s[k] e^{j k \delta} \quad (7)$$

où δ est un réel petit, et j la racine de -1 . Si dans le cas SISO, l'égalisation aveugle et le résidu de porteuse commutent, il n'en est pas de même dans le cas MIMO. On doit donc d'abord égaliser le canal, et ensuite éliminer le résidu de porteuse, ou bien faire les deux en même temps.

Références

- [1] M. BABAIE-ZADEH, C. JUTTEN, K. NAYEBI, "Blind separation of post-nonlinear mixtures", in *Int. Conf. Indep. Comp. Ana. (ICA'01)*, San Diego, Dec. 2001, pp. 138–143.
- [2] P. COMON, "Independent Component Analysis, a new concept?", *Signal Processing, Elsevier*, vol. 36, no. 3, pp. 287–314, Apr. 1994, Special issue on Higher-Order Statistics.
- [3] P. COMON, "Contrasts for multichannel blind deconvolution", *Signal Processing Letters*, vol. 3, no. 7, pp. 209–211, July 1996.
- [4] P. COMON, "Independent component analysis, contrasts and convolutive mixtures", in *Second IMA Conference on Mathematics in Communications*, Lancaster, UK, Dec. 16-18, 2002, pp. 10–17, invited.
- [5] P. COMON, "Tensor decompositions", in *Mathematics in Signal Processing V*, J. G. McWhirter, I. K. Proudler, Eds., pp. 1–24. Clarendon Press, Oxford, UK, 2002.
- [6] P. COMON, "Contrasts, Independent Component Analysis, and blind deconvolution", submitted to *Int. Journal Adapt. Control Sig. Proc.*, 2003, I3S Research Report RR-2003-06.
- [7] P. COMON, P. CHEVALIER, "Source separation : Models, concepts, algorithms and performance", in *Unsupervised Adaptive Filtering, Vol. I, Blind Source Separation*, S. Haykin, Ed., Series on Adaptive and learning systems for communications signal processing and control, pp. 191–236. Wiley, 2000.
- [8] Z. DING, Y. LI, *Blind Equalization and Identification*, Dekker, New York, 2001.
- [9] E. GASSIAT, F. GAMBOA, "Source separation when the input sources are discrete or have constant modulus", *IEEE Trans. Sig. Proc.*, vol. 45, no. 12, pp. 3062–3072, Dec. 1997.
- [10] O. GRELLIER, P. COMON, "Blind separation and equalization of a channel with MSK inputs", in *SPIE Conference*, San Diego, July 19-24 1998, pp. 26–34, invited session.
- [11] O. GRELLIER, P. COMON, "Analytical blind discrete source separation", in *Eusipco*, Tampere, Finland, 5-8 sept. 2000.
- [12] O. GRELLIER, P. COMON, B. MOURRAIN, P. TREBUCHET, "Analytical blind channel identification", *IEEE Trans. Signal Processing*, vol. 50, no. 9, pp. 2196–2207, Sept. 2002.
- [13] F. GUSTAFSSON, B. WAHLBERG, "Blind equalization by direct examination of input sequences", *IEEE Transactions on Computers*, vol. 43, no. 7, pp. 2213–2222, July 1995.
- [14] S. HAYKIN, *Unsupervised Adaptive Filtering*, vol. 1, Wiley, 2000, series in Adaptive and Learning Systems for Communications, Signal Processing, and Control.
- [15] A. HYVÄRINEN, J. KARHUNEN, E. OJA, *Independent Component Analysis*, Wiley, 2001.
- [16] T. H. LI, "Analysis of a non-parametric blind equalizer for discrete valued signals", *IEEE Trans. on Sig. Proc.*, vol. 47, no. 4, pp. 925–935, Apr 1999.
- [17] T. H. LI, K. MBAREK, "A blind equalizer for nonstationary discrete-valued signals", *IEEE Trans. Sig. Proc.*, vol. 45, no. 1, pp. 247–254, Jan. 1997, Special issue on communications.
- [18] J.-C. PESQUET, E. MOREAU, "Cumulant based independence measures for linear mixtures", *IEEE Trans. Information Theory*, pp. 1947–1956, March 2001.
- [19] D. T. PHAM, "Blind separation of instantaneous mixture of sources via an independent component analysis", *IEEE Trans. Sig. Proc.*, vol. 44, no. 11, pp. 2768–2779, Nov. 1996.
- [20] J. G. PROAKIS, *Digital Communications*, McGraw-Hill, 1995, 3rd edition.
- [21] A. L. SWINDLEHURST, S. DAAS, J. YANG, "Analysis of a decision directed beamformer", *IEEE Trans. Sig. Proc.*, vol. 43, no. 12, pp. 2920–2927, Dec. 1995.
- [22] S. TALWAR, M. VIBERG, A. PAULRAJ, "Blind estimation of multiple co-channel digital signals arriving at an antenna array : Part I, algorithms", *IEEE Trans. Sig. Proc.*, pp. 1184–1197, May 1996.
- [23] A. J. van der VEEN, S. TALWAR, A. PAULRAJ, "Blind estimation of multiple digital signals transmitted over FIR channels", *IEEE Sig. Proc. Letters*, vol. 2, no. 5, pp. 99–102, May 1995.
- [24] T. WIGREN, "Avoiding ill-convergence of finite dimensional blind adaptation schemes excited by discrete symbol sequences", *Signal Processing, Elsevier*, vol. 62, no. 2, pp. 121–162, Oct. 1997.
- [25] D. YELLIN, B. PORAT, "Blind identification of FIR systems excited by discrete-alphabet inputs", *IEEE Trans. Sig. Proc.*, vol. 41, no. 3, pp. 1331–1339, 1993.