IEEE Workshop SPAWC'03, June 15-18, 2003, Rome, Italy

CONTRAST FUNCTIONS FOR DETERMINISTIC BLIND SOURCE SEPARATION

Pierre COMON and Jérôme LEBRUN

I3S Laboratory, Algorithmes-Euclide-B, 2000 route des Lucioles
BP 121, F-06903 Sophia-Antipolis cedex, France
`name@unice.fr www.i3s.unice.fr/~name`

ABSTRACT

Blind Source Separation (BSS) is often carried out under the assumption that sources are statistically mutually independent, at least in the sense of cumulants of given order. However, this assumption is not mandatory, and can be replaced by some assumption on the source distribution, even if all sources are identically distributed. Contrast functions are optimization criteria that satisfy some identifiability conditions. In this paper, one defines a distance to any discrete constellation, and proves that this family of criteria indeed defines contrast functions. The advantage of such criteria is that they are deterministic, and do not involve the estimation of sample statistics, such as moments or cumulants, hence a potentially shorter convergence time.

1. THE MIMO BLIND DECONVOLUTION PROBLEM

Blind equalization or identification schemes have been the subject of growing interest since 1975. One of the main advantages of blind techniques is that, by deleting pilot sequences, one can increase the transmission rate. But there are other advantages, which stem from limitations of classical approaches. In fact, techniques based on pilot sequences are difficult to use when channel responses are long, or fast varying, compared to the length of the pilot sequence. The presence of a carrier residual can also make the equalization task more difficult.

Instead of basing the identification or equalization schemes on input-output measurements (data-aided approaches), some properties about the inputs are exploited (blind approaches), as is now explained.

1.1. Modeling.

Limiting our discussion to linear modulations, the complex envelope of a transmitted signal $s(t)$ takes the form in baseband [20]: $s(t) = \sum_k g(t-kT)\, s[k]$. Note the distinction between discrete-time and continuous time processes via brackets and parentheses: $s[k] = s(kT_s)$, where T_s is the symbol period. After propagation through the channel and the receive filter, the signal received on the antenna may be written as: $y(t) = \sum_k h(t-kT_s)\, s[k]$ where h is the convolution of the transmit filter, the channel, and the receive filter. If the received signal is sampled at the rate $1/T_s$, it can be modeled as:

$$y[n] = \sum_k h[n-k]\, s[k] \tag{1}$$

with $h[k] \stackrel{\text{def}}{=} T_s\, h(kT_s)$. For Multiple Input Multiple Output (MIMO) systems, the transmitted signal $\boldsymbol{s}[k]$ and the received signal $\boldsymbol{y}[k]$ may be considered as vector-valued discrete-time processes; their dimension is denoted by P and K, respectively. Note the boldface that emphasizes the multi-dimensional character of the processes. Model (1) can then be rewritten as:

$$\boldsymbol{y}[n] = \sum_k \boldsymbol{H}[n-k]\, \boldsymbol{s}[k] \tag{2}$$

where the global channel impulse response $\boldsymbol{H}[k]$ is now a sequence of $K \times P$ matrices. Its $z-$transform, with a slight abuse of notation, is denoted as

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

In the present context, inputs $s_j[k]$ are referred to as *sources*.

1.2. Symbol rate mismatch.

The case where source symbol rates are different or unknown is not addressed in this paper, although it is an interesting issue, in surveillance or interception contexts for instance. However, one can still say that the output appears as a convolution, but not as a *discrete convolution* anymore. In fact, if the sample rate is $1/T'$ at the receiver, we have: $\boldsymbol{y}[n] = \sum_k \boldsymbol{H}(nT' - kT_s)\, \boldsymbol{s}[k]$

1.3. Carrier offset.

Another important issue is that of carrier mismatch. If the carrier frequency at the receiver is slightly different from the modulating carrier, then there is a *carrier residual*, which one can merely represent in baseband by a multiplicative exponential. For a SISO channel, this can be written as:

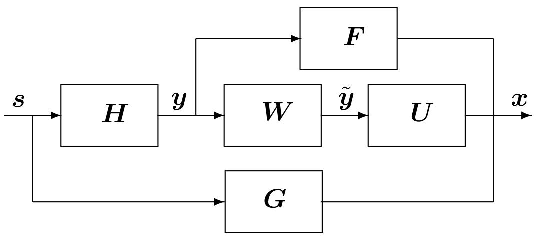
$$y[n] = \sum_k h[n-k]\, s[k]\, e^{\jmath\, k\, \delta} \tag{3}$$

where δ is a small real number, and the dotless $\jmath$ denotes $\sqrt{-1}$.

In the SISO case, carrier residual and blind equalization commute. This is not enjoyed by MIMO channels; we must first equalize the channel, and then carrier residuals can be mitigated individually. The problem of carrier residual mitigation, if not treated jointly with blind equalization, can be seen as a SISO problem. See [4] and references therein.

1.4. Assumptions and taxonomy.

The goal of blind equalization is to yield an estimate of input sequences $s_j[k]$ from the sole observation of output sequences $y_i[n]$.



The transmitted sequence $\boldsymbol{s}[n]$ propagates through a channel $\check{\boldsymbol{H}}[z]$, is then whitened by a filter $\check{\boldsymbol{W}}[z]$, and is eventually deconvolved by a paraunitary equalizer $\check{\boldsymbol{U}}[z]$ to yield output $\boldsymbol{x}[n]$.

In order to blindly equalize convolutive models, the most widely used assumption is the *statistical independence* between successive symbols.

Hypothesis H1 *Sources $s_j[k]$ are all i.i.d. sequences.*

For MIMO models, the independence assumption between sources is often utilized:

Hypothesis H2 *Sources $s_j[k]$ are mutually statistically independent.*

These hypotheses can generally be deflated to less strong whiteness/independence properties, because moments of finite orders are used [7]. Let us stress that the case where sources are *linear processes* can also be treated in a similar manner as i.i.d. sources; Hypothesis H1 is thus not very restrictive. A particular case however raises problems, namely that of Gaussian sources. In that case, all the information is contained in moments up to order 2, which is not sufficient to establish identifiability. For this reason, it is necessary to resort to a third hypothesis, along with hypotheses H1 and H2:

Hypothesis H3 *At most one source is Gaussian.*

On the other hand, there exist other frameworks in which hypotheses are different. For instance, if sources have different spectra, or if they are non stationary, or cyclostationary, then they can be separated with the help of appropriate techniques (cf. section 1.5). These three cases are not addressed in the present paper, and yield quite different (easier) theoretical problems. Nevertheless, the special framework of *discrete sources*, which is relevant in the context of digital communications, is our main concern. Therefore, we assume the following hypotheses, instead of hypotheses H1 to H3:

Hypothesis H4 *The sources $s_j[n]$ belong to a known finite alphabet $\mathcal{A}$ characterized by the d distinct complex roots of a polynomial of degree d, $Q(z)$*

Hypothesis H5 *Sources $s_j[n]$ are sufficiently exciting*

By *sufficiently exciting*, it is meant that all d^P possible states of the $P-$uplet $\boldsymbol{s}$ appear in the data matrix. In particular, the observation length must be at least d^P, and the source data matrix is full row rank P.

Depending on hypotheses (independence, discrete character, SISO/MIMO, $P > K$ or not...), a whole variety of problems can be stated [7]. From now on, and unless otherwise specified, we shall concentrate only on Hypotheses H4 and H5, and on the case where $P \leq K$ (over-determined mixtures). When more sources than sensors are present, the problem becomes more complicated (under-determined mixtures), and specific tools generally need to be utilized.

Example Assume the model is MIMO static. Then $\boldsymbol{y}[n] = \boldsymbol{H}\ \boldsymbol{s}[n]$, where $\boldsymbol{y}[n]$ and $\boldsymbol{s}[n]$ are realizations of random variables. In that case, hypothesis H1 is not mandatory anymore. The estimation of the pair $(\boldsymbol{H}, \boldsymbol{s}[n])$ from the sole observations $\boldsymbol{y}[n]$ under hypotheses H2 and H3 is now called *Independent Component Analysis* (ICA) [2] [19] [15] [14] [5].

1.5. Bibliographical comments.

For more references on the use of the discrete character of sources in MIMO over-determined mixtures, either static or convolutive, see [12] [25] [23] [21] [22] [24] [17] [9] [16] [13]. Under-determined mixtures are addressed in [5] and references therein. For a general account on blind techniques, see [14] [15] [8]. Some useful results on contrasts can be found in [2] [18] [6].

2. CONTRASTS

When noise is present in model (2), the estimation of inputs can be carried out according to a Maximum Likelihood (ML) or a Maximum A Posteriori (MAP) procedure if the noise has a known distribution. If this is not the case, noise must be considered as a nuisance. Contrast criteria are dedicated to this kind of situation.

2.1. Trivial filters.

The separating linear filter, $\check{\boldsymbol{F}}[z]$, if it exists, aims at delivering an output, $\boldsymbol{x}[n]$, which should satisfy as closely as possible hypotheses H4 and H5. But it is clear that there exist some filters that do not affect them. These are called the trivial filters. For instance, it has been proved that under hypotheses H1 to H3, that

Proposition 1 *Under hypotheses H1 to H3, trivial filters are of the form $\check{\boldsymbol{T}}[z] = \boldsymbol{P}\,\check{\boldsymbol{D}}[z]$, where $\boldsymbol{P}$ is a permutation, and $\check{\boldsymbol{D}}[z]$ a diagonal filter. In addition, because of the i.i.d. property of hypothesis H1, entries of $\check{\boldsymbol{D}}[z]$ must be of the form $\check{D}_{pp}[z] = \lambda_p\, z^{d_p}$, where d_p is an integer.*

Consequently, it is hopeless to estimate the pair $(\check{\boldsymbol{H}}[z], \boldsymbol{s}[k])$. One should rather try to estimate one representative of the equivalence class of solutions. Once one solution is found, all the others can be generated by trivial filtering.

2.2. Definition of contrast functionals

Let $\mathcal{H}$ be a set of filters, and denote $\mathcal{H}\cdot\mathcal{S}$ the set of processes obtained by operation of filters of $\mathcal{H}$ on processes of $\mathcal{S}$. Denote $\mathcal{T}$ the subset of $\mathcal{H}$ of trivial filters, defined in proposition 1. An optimization criterion, $\Upsilon(\boldsymbol{H};\boldsymbol{x})$, is referred to as a contrast, defined on $\mathcal{H}\times\mathcal{H}\cdot\mathcal{S}$, if it satisfies the three properties below [3]:

P1 *Invariance:* The contrast should not change within the set of acceptable solutions, which means that $\forall \boldsymbol{H} \in \mathcal{T}, \forall \boldsymbol{x} \in \mathcal{H}\cdot\mathcal{S},\ \Upsilon(\boldsymbol{H};\boldsymbol{x}) = \Upsilon(\boldsymbol{I};\boldsymbol{x})$.

P2 *Domination:* If sources are already separated, any filter should decrease the contrast. In other words, $\forall \boldsymbol{s} \in \mathcal{S}, \forall \boldsymbol{H} \in \mathcal{H}$, then $\Upsilon(\boldsymbol{H};\boldsymbol{s}) \leq \Upsilon(\boldsymbol{I};\boldsymbol{s})$.

P3 *Discrimination:* The maximum contrast should be reached only for filters linked to each other via trivial filters: $\forall \boldsymbol{s} \in \mathcal{S}, \Upsilon(\boldsymbol{H};\boldsymbol{s}) = \Upsilon(\boldsymbol{I};\boldsymbol{s}) \Rightarrow \boldsymbol{H} \in \mathcal{T}$.

The most natural criterion to measure the statistical mutual independence between P variables z_p is the divergence between the joint probability density and the product of the marginal ones [2]. If we assume the Kullback-Leibler divergence, we end up with the Mutual Information (MI) [2]. The MI is thus a first possible contrast function. However, its practical use is rather difficult, especially in large dimension (e.g. convolutive mixtures), even if some iterative algorithms have been devised [1]. Therefore, contrasts based on cumulants have been often preferred.

Now under hypotheses H4 and H5, it is quite natural to define the optimization criterion:

$$\Upsilon_P(\boldsymbol{x}) = \sum_n \sum_i |Q(x_i(n))|^2 . \tag{4}$$

2.3. Finite alphabets

Under mild hypotheses, criterion (4) turns out to be a contrast. The interest of exploiting the discrete character lies not only in a more accurate characterization of the desired output (than just non Gaussian or CM), but also in the fact that some other assumptions can be dropped. In this section, hypotheses H4 and H5 are solely used.

For instance, sources can be correlated and non stationary. In fact, the approach proposed is entirely algebraic and deterministic, so that no statistical tool is required.

Definition 2 *Let $\mathcal{A}$ be a finite alphabet defined by $Q(x) = 0$, where Q is a polynomial of degree d with $d > 1$ distinct roots, and let $\mathcal{G}$ be the set of complex numbers γ such that $\gamma\,\mathcal{A} \subset \mathcal{A}$.*

As a first obvious result, we have [6]:

Lemma 3 *If $\mathcal{A}$ is finite, then $\mathcal{G}$ contains only unit modulus numbers.*

Lemma 4 *Trivial filters associated with hypotheses H4 and H5 are of the form $\boldsymbol{P}\,\check{\boldsymbol{D}}[z]$, where the entries of $\check{\boldsymbol{D}}[z]$ can be written as $\check{D}_{pp}[z] = \gamma\, z^n$, with $\gamma \in \mathcal{G}$ and $n \in \mathbb{Z}$.*

Theorem 5 *Let $\mathcal{S}$ be the set of processes taking their values in alphabet $\mathcal{A}$, and $\mathcal{H}$ the set of $P \times P$ invertible FIR filters. Then criterion (4) is a contrast under hypotheses 4 and 5.*

Proof. Some details are skipped for reasons of space; see [6] for a longer version. Assume that, for some $\boldsymbol{c} \in \mathbb{C}^P$, we have that $\boldsymbol{x}^{\mathsf{T}}\boldsymbol{c} \in \mathcal{A}, \forall \boldsymbol{x} \in \mathcal{A}^P$. Denote $\mathcal{A} = \{a_1, \ldots a_d\}$, $\boldsymbol{c}^{\mathsf{T}} = [c_1, \ldots c_P]$, and $\mathbf{1}$ the vector formed of P ones.

First of all, from 3, $\sum_i c_i \in \mathcal{G}$ and there exists a number $\gamma \in \mathcal{G}$ such that $\gamma \sum_i c_i = 1$. By defining $c_i' = \gamma\, c_i$, it is sufficient to consider the case $\mathcal{G} = \{1\}$.

• Because $\boldsymbol{x}^{\mathsf{T}}\boldsymbol{c} \in \mathcal{A}$ for any vector $\boldsymbol{x}$ containing elements of $\mathcal{A}$, it must be true in particular for $\boldsymbol{x}^{\mathsf{T}} = a_p\,\mathbf{1}^{\mathsf{T}}$. This implies that $\forall a_p \in \mathcal{A},\ a_p \sum_i c_i \in \mathcal{A}$. In other words, we must always have

$$\sum_i c_i \in \mathcal{G}. \tag{5}$$

• Next, for any pair of distinct complex numbers a and b, define the $P \times P$ matrix $\boldsymbol{B} = (a-b)\,\boldsymbol{I} + b\,\mathbf{1}\mathbf{1}^{\mathsf{T}}$. This matrix has a determinant equal to $(a-b)^{P-1}(a+(P-1)b)$. For $a + (P-1)b \neq 0$, its inverse takes the form $bB^{-1} = [a-b]^{-1}\,[\boldsymbol{I} - b\,\mathbf{1}\mathbf{1}^{\mathsf{T}}/[a+(P-1)b]]$.

• As a result, for any pair of distinct symbols of $\mathcal{A}$, a and b, there exists a vector $\boldsymbol{\alpha}$ containing P symbols of $\mathcal{A}$ such that $\boldsymbol{B}\boldsymbol{c} = \boldsymbol{\alpha}$. From above, we have in particular

$$\sum_i c_i = [a + (P-1)b]^{-1} \sum_i \alpha_i \tag{6}$$

Case of real symbols, with $\mathcal{G} = \{1\}$. Denote $x_m = \min\{x,\, x \in \mathcal{A}\}$ and $x_M = \max\{x,\, x \in \mathcal{A}\}$. From (5),

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

So c_i must be real too. It can be easily shown [6] that $c_i \in \{-1, 0, 1\}$, $\forall i$, $1 \leq i \leq P$, otherwise $\mathcal{A}$ would be infinite. If $c_i \in \{0, 1\}$, $\forall i$, then again from $\sum_i c_i = 1$, there is a single nonzero entry in $\boldsymbol{c}$, and $\boldsymbol{c}$ is eventually trivial. So assume $\exists c_i = -1$, $1 \leq i \leq P$. But then $\beta = b + c_i(a - b) \in \mathcal{A}$ for any pair $(a, b) \in \mathcal{A}^2$; in particular for $c_i = -1$, $a = x_M$ and $b = x_m$, the symbol $\beta = x_m - (x_M - x_m)$ should belong to $\mathcal{A}$. And $x_M > x_m \Leftrightarrow \beta < x_m$, which contradicts the definition of x_m.

Complex case with $\mathcal{G} = \{1\}$. If $d = 2$, the problem is equivalent to a particular case of real alphabet, already addressed. So assume $d > 2$, and choose a symbol b on the convex hull of $\mathcal{A}$. Since $d > 2$, b always has two distinct neighbors on the convex hull. So choose one of the two neighbors on the convex hull, denoted a, in order to also have $a + (P - 1)b \neq 0$. Result (6) then applies. Since $\sum_i c_i \in \mathcal{G}$, (6) yields

$$\frac{1}{P}\left[a + (P-1)b\right] = \frac{1}{P}\sum_i \alpha_i \tag{7}$$

Let us prove first that $\boldsymbol{\alpha}$ cannot be proportional to $\mathbf{1}$. Assume $\boldsymbol{\alpha} = a_o \mathbf{1}$ for some $a_o \in \mathcal{A}$. Then from (7), $a_o = \frac{1}{P}a + \frac{P-1}{P}b$. Therefore symbol a_o is also on the convex hull of $\mathcal{A}$, and is closest to b than a was. This contradicts the fact that a was one of its two neighbors of b.

So assume now that vector $\boldsymbol{\alpha}$ contains at least two distinct symbols. If these symbols are a and b, then we necessarily have 1 time a and $(P-1)$ times b, and $\boldsymbol{c}$ is trivial, as already seen. If all symbols of $\boldsymbol{\alpha}$ are real, this case is equivalent to stage 1, and has been already treated. Thus assume there is in $\boldsymbol{\alpha}$ a third symbol x distinct from a and b, being not a real linear combination of a and b. From (6), there must be at least another symbol x' on the other side of the line spanned by $\{a, b\}$. But then one of them lies outside the convex hull of $\mathcal{A}$. This contradicts the fact that both x and x' are in $\mathcal{A}$. □

2.4. PSK contrasts

The simplest case of discrete alphabet is defined by the polynomial equation $P(z) = z^q - 1$, for which we have

Lemma 6 *Define* $\boldsymbol{F}$ *as the matrix of the length*$-q$ *Fourier transform:* $F_{k\ell} = e^{\jmath 2\pi(k-1)(\ell-1)/q}$. *Then for any permutation* $\boldsymbol{P}$, *there exists a diagonal matrix* $\boldsymbol{\Delta}$ *containing only* q*th roots of unity such that*

$$\boldsymbol{F}\,\boldsymbol{P} = \boldsymbol{\Delta}\,\boldsymbol{F}$$

Theorem 7 *Let* $\mathcal{S}$ *be the set of PSK*$-q$ *processes, and* $\mathcal{H}$ *the set of* $P \times P$ *invertible FIR filters. Then*

$$\Upsilon(\boldsymbol{x}) \stackrel{\text{def}}{=} -\sum_i \sum_n \left| x_{i,n}^q - 1 \right|^2$$

is a contrast if the matrix $(s_i[n])$ *is full rank.*

The PSK case is interesting because it allows to derive a much simpler proof.

Proof. We prove the theorem for static mixtures: $\boldsymbol{x}[n] = \boldsymbol{A}\,\boldsymbol{s}[n]$. First, the contrast is obviously null for $x \in \mathcal{S}$, and always negative. Thus $\Upsilon(\boldsymbol{x}) \leq \Upsilon(\boldsymbol{s})$. If equality holds, this means that $x_i[n]^q = 1$ for all (i, n). We have a system of (polynomial) equations in unknowns A_{ip}. The matrix of source signals is formed of qth roots of unity. If it is full rank, so that there exist a sub-matrix $P \times P$, $\boldsymbol{S}$, related to the length$-q$ Fourier transform matrix, $\boldsymbol{F}$, by $\boldsymbol{S} = \boldsymbol{P}_1 \boldsymbol{\Delta}_1 \boldsymbol{F}$, where $\boldsymbol{P}_1$ is a permutation and $\boldsymbol{\Delta}_1$ is a $P \times q$ diagonal matrix containing qth roots of unity. Let $\boldsymbol{X} = \boldsymbol{A}\,\boldsymbol{S}$. Then $\boldsymbol{X}$ is invertible. But its entries satisfy $X_{ij}^q = 1$. Thus, there exist a permutation and a diagonal matrix of qth roots of unity such that $\boldsymbol{X} = \boldsymbol{P}_2 \boldsymbol{\Delta}_2 \boldsymbol{F}$. This implies that $\boldsymbol{A} = \boldsymbol{P}_2 \boldsymbol{\Delta}_2 \boldsymbol{\Delta}_1^- \boldsymbol{P}_1{}^{\mathsf{T}}$, where $\boldsymbol{\Delta}^-$ denotes the pseudo-inverse of $\boldsymbol{\Delta}$. From lemma 6, the latter result can be put in the form $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Delta}$. We have proved that $\boldsymbol{A}$ is trivial. □

In this case, trivial filters are of the form $\boldsymbol{P}\,\check{\boldsymbol{D}}[z]$, where $\check{D}_{pp}[z]$ are rotations in the complex plane of an angle multiple of $2\pi/q$ combined with a pure delay, and $\boldsymbol{P}$ are permutations.

2.5. Numerical algorithms

It is easy to run gradient ascents to find the maxima of $\Upsilon(\boldsymbol{x})$ defined in theorem 7 or 5. A typical iteration to estimate a vector of equalizer taps is for instance:

$$\boldsymbol{v} = \boldsymbol{f}(k) + \eta\,\boldsymbol{g}(k);\ \boldsymbol{f}(k+1) = \boldsymbol{v}/||\boldsymbol{v}|| \tag{8}$$

where $\boldsymbol{g}(k)$ denotes the gradient of the optimization criterion $J(\boldsymbol{f})$ calculated at $\boldsymbol{f}(k)$, and η the step size. Standard gradient implementations, especially with a fixed step, perform poorly because of the shape of the criterion, which contains many saddle points. The way the step size is adjusted (e.g. quasi-Newton) does not improve anything with this respect: if the algorithm is initialized near a saddle point, the iterations can stay a long time in its neighborhood, and suddenly burst out far away from the attraction basin, and take again a long time to come back. Yet, a significant improvement can be brought to this.

In fact, if the criterion $J(\boldsymbol{f})$, implicitly defined by $\Upsilon(\boldsymbol{x})$, is a rational function in the f_i's, then so is $J(\boldsymbol{f}(k)+\eta\,\boldsymbol{g}(k))$. As a consequence, all its stationary points can be explicitly computed, as roots of a polynomial in a single variable, and the *absolute* minimum/maximum easily found. This kind of algorithm may often give the possibility to leave the attraction basin of a local minimum, if any. The algorithm can obviously be implemented either off-line or on-line.

In the present case, there is even another alternative since algebraic solutions can be computed, as reported in [11] [10], among others.

Other approaches, that are not based on contrast maximization, exist in the literature, including [25] [21] [23] [22] [24] [17].

3. CONCLUDING REMARKS

The family of optimization criteria we have defined allow to carry out a MIMO channel Blind Equalization when inputs are discrete, in a deterministic manner. Since no moments are computed, a faster convergence may be expected. In addition, these criteria enjoy contrast properties, insuring the minimal identifiability conditions.

Algebraic block solutions become more and more attractive, especially in TDMA transmissions, because of the increased computational power. Therefore, off-line as well as on-line iterative algorithms are proposed. Computer experiments are currently being performed.

4. REFERENCES

[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.