MULTIDIMENSIONAL INDEPENDENT COMPONENT ANALYSIS.

Jean-François Cardoso

CNRS/ENST, 46 rue Barrault, 75634 Paris Cedex 13

cardoso@sig.enst.fr   http://sig.enst.fr/∼cardoso/stuff.html

ABSTRACT

This discussion paper proposes to generalize the notion of Independent Component Analysis (ICA) to the notion of *Multidimensional* Independent Component Analysis (MICA). We start from the ICA or blind source separation (BSS) model and show that it can be uniquely identified provided it is properly parameterized in terms of one-dimensional subspaces. From this standpoint, the BSS/ICA model is generalized to multidimensional components. We discuss how ICA standard algorithms can be adapted to MICA decomposition. The relevance of these ideas is illustrated by a MICA decomposition of ECG signals.

1. BLIND SOURCE SEPARATION

We start by considering the blind source separation (BSS) problem in the simplest model: an $n \times 1$ vector of observations $\mathbf{x}$ is modeled as

$$\mathbf{x} = A\mathbf{s}, \quad \text{with} \quad A = [\mathbf{a}_1, \ldots, \mathbf{a}_n] \tag{1}$$

where $\mathbf{s}$ is a $n \times 1$ vector with *statistically independent* components and matrix A is an $n \times n$ invertible matrix. The entries of $\mathbf{s} = [s_1, \ldots, s_n]^\dagger$ are referred to as the 'source signals'. For the sake of simplicity, the discussion is restricted throughout to the case of zero-mean real signals.

The source separation problem may be stated as: *Identify mixing matrix A and/or estimate the source signals based only on observations of $\mathbf{x}$ and assuming only i) statistical independence of the 'sources' and ii) linear independence of the columns of A.* The strength of this model is that the two independence assumptions stated above are physically plausible in several instances and are strong enough to provide some kind of identifiability (see below), thus alleviating the need of any further modeling of the source distributions or of the mixing matrix.

Indeterminacies. The source separation problem as stated above is clearly undetermined: if nothing is known a priori neither about the amplitude of a particular source s_p nor about the amplitude of the corresponding column of A, then a scalar factor can be exchanged between s_p and $\mathbf{a}_p$ without changing the product: $\mathbf{a}_p s_p = (\alpha \mathbf{a}_p)\ (\alpha^{-1} s_p)$ for any real $\alpha \neq 0$. Also we note that the ordering of the source signals is immaterial and is nothing but a notational device. Thus, in complete ignorance of the source distributions, source signals can be recovered at best up to a permutation, scales and signs.

These indeterminations are well known and have been discussed at length in the blind source separation literature. They could be crudely expressed as: 'under the working assumptions, the mixing matrix A *does not exist*.' We now discuss what quantities can be truly determined from the distribution of $\mathbf{x}$ in the BSS model.

Proceedings of ICASSP '98. Seattle.

2. INDEPENDENT COMPONENT ANALYSIS

In the literature, the terms 'blind source separation' and 'independent component analysis' [1] are often used indifferently: they refer to the same model (1) with the same assumptions, pursue the same objectives and are addressed with the same algorithms. This is a source of confusion and a 'waste of terminology'... It is also unfortunate because the term 'analysis' refers to the idea of decomposition into smaller, simpler elements and very often this decomposition is into a sum of terms, calling for an additive model rather than a multiplicative as described by eq. (1).

ICA as an additive model. We claim that an interesting reformulation of the basic ICA model is obtained by defining the 'components' as $\mathbf{x}_p \stackrel{\text{def}}{=} \mathbf{a}_p\ s_p$ for $1 \leq p \leq n$ so that model (1) can be rewritten as an additve, component-based, model:

$$\mathbf{x} = \sum_{p=1}^{n} \mathbf{x}_p \tag{2}$$

This (admittedly trivial) rewriting of the original model calls for a change of standpoint. While model (1) is a multiplicative model reading: *the observed vector is the the product of a mixing matrix A by a source vector* $\mathbf{s}$, model (2) is an additive model reading: *the observed vector is a sum of n one-dimensional independent* ***vectors*** $\mathbf{x}_1, \ldots, \mathbf{x}_n$.

This point of view is closer to the classic (and closely related) Principal Component Analysis (PCA) technique. PCA provides a *decomposition* of a second-order vector $\mathbf{x}$ as the sum of its projections onto the principal axis of its covariance matrix. The PCA components are geometrically orthogonal by construction and also statistically orthogonal (uncorrelated). The more ambitious ICA approach is to look for components which are not necessarily geometrically orthogonal but are statistically independent (that is 'more than statistically orthogonal' since independence is much stronger than mere uncorrelation).

A geometric parameterization. The new view of ICA introduced above is 'matrix-free'. In this section, we discuss how it can be parameterized and why the appropriate parameterization is uniquely determined contrarily to the matrix-based parameterization of model (1).

In the component model (2), the smallest subspace containing the *p*th component is referred to as the 'component (sub)space' for the *p*th component. This is indeed the one-dimensional linear subspace spanned by the *p*th column of A. The orthogonal projector onto this subspace is denoted Π_p and can be obtained as

$$\Pi_p \stackrel{\text{def}}{=} \frac{\mathbf{a}_p \mathbf{a}_p^\dagger}{\mathbf{a}_p^\dagger \mathbf{a}_p}, \qquad 1 \leq p \leq n. \tag{3}$$

The knowledge of the projectors Π_p for $p = 1, \ldots, n$ is of course sufficient for separating the components: it is easily verified that

$$\mathbf{x}_p = \tilde{\Pi}_p \mathbf{x} \quad \text{with} \quad \tilde{\Pi}_p \stackrel{\text{def}}{=} \Pi_p \left(\sum_{q=1}^{n} \Pi_q \right)^{\#} \tag{4}$$

because matrix $\tilde{\Pi}_p$ is the projector onto the p-th component space orthogonally to all the other components. In (4), superscript $\#$ denotes pseudo-inversion (we could use a regular inverse here but a pseudo-inverse is required below).

For the sake of consistency *i.e.* to get rid of the modeling in terms of mixing matrix, it is desirable to define the projector Π_p explicitly in terms of the component $\mathbf{x}_p$ rather than as a function of $\mathbf{a}_p$. This can be done for instance as $\Pi_p = (\mathrm{E}|\mathbf{x}_p|^2)^{-1}\mathrm{E}\{\mathbf{x}_p^\dagger \mathbf{x}_p\}$ if $\mathbf{x}_p$ has finite variance or by a similar trick otherwise.

By focusing on the spaces containing to each component rather than on the columns of A, we obtain the desired result of getting rid of the indeterminations of scale and sign. In some sense, we move from an *algebraic* description of a mixture in terms of a 'mixing matrix' to a *geometric* description in terms of 'component spaces'.

The global 'parameter of interest' in model (2) is not the mixing matrix but the set

$$\mathcal{P} = (\Pi_1, \ldots, \Pi_n)$$

of the orthogonal projection matrices onto each of the component spaces. The last step to removing indeterminacies in the ICA model is maybe only rhetorical: it is easily seen that *the* unordered *set* $\mathcal{P}$ *is* uniquely *determined whenever matrix* A *is determined up to scale, sign and column order*. Equivalently, if matrix A is identifiable up to the above-mentioned indeterminacies of scale, sign and order, then $\mathcal{P}$ is uniquely identifiable and, knowing $\mathcal{P}$, the (unordered) set of independent components $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ can be uniquely recovered via (4).

In summary, we have completed the reparameterization of the standard ICA model into a 'component model' which is geometric in spirit and free of indeterminacies. More importantly, this new perspective suggests an extension to a more general model of multidimensional independent components.

3. MULTIDIMENSIONAL INDEPENDENT COMPONENT ANALYSIS

The geometrical description of ICA discussed in the previous section offers a simple way to generalize ICA into multidimensional ICA (MICA).

Definition. *Let* $E_1, \ldots, E_c$ *be* c *linear subspaces of* $\mathbf{R}^n$. *They are said to be* linearly independent *if any vector* $\mathbf{x}$ *of* $E_1 \oplus \cdots \oplus E_c$ *admits of a* unique *decomposition as* $\mathbf{x} = \sum_{p=1}^{c} \mathbf{x}_p$ *with* $\mathbf{x}_p \in E_p$ *for* $1 \leq p \leq c$. In such a case, the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_c$ are called the *linear components* of $\mathbf{x}$ on the set $E_1, \ldots, E_c$.

Definition 1. *A random* n*-dimensional vector* $\mathbf{x}$ *admits of a MICA decomposition* $\{\mathbf{x}_1, \ldots, \mathbf{x}_c\}$ *in* c *components if it exists* c *linearly independent 'component subspaces'* $E_1, \ldots, E_n$ *of* $\mathbf{R}^n$ *on which the linear components of* $\mathbf{x}$ *are* statistically *independent.*

To illustrate the definition, let us consider it in a matrix-vector style. Let $\mathbf{s}_1, \ldots, \mathbf{s}_c$ be c statistically independent random vectors with dimensions $n_1, \ldots, n_c$ and let $\mathbf{s} = [\mathbf{s}_1^\dagger, \ldots, \mathbf{s}_c^\dagger]^\dagger$. Let $A_1, \ldots, A_c$ be c matrices of sizes $n \times n_1, \ldots, n \times n_c$, such that $A = [A_1, \ldots, A_c]$ is full column rank. Then, vector $A\mathbf{s}$ admits of a MICA decomposition onto the spaces $E_1, \ldots, E_c$ where $E_p = \mathrm{Span}(A_p)$ for $1 \leq p \leq c$. The orthogonal projector onto E_p is $\Pi_1, \ldots, \Pi_c$:

$$\Pi_p = A_p (A_p^\dagger A_p)^{-1} A_p^\dagger, \qquad 1 \leq p \leq c \tag{5}$$

which is the multidimensional equivalent of (3) while eq. (4) holds 'as is' in the multidimensional case. The corresponding MICA components are $\mathbf{x}_p = A_p \mathbf{s}_p$ for $1 \leq p \leq c$ indeed. The very same components are also obtained as $\mathbf{x}_p = (A_p C_p)(C_p^{-1}\mathbf{s}_p)$ for any invertible $n_p \times n_p$ matrix C_p. Therefore, in the MICA setting, indeterminations appear more severe than in the ICA setting: the n_p-dimensional source vector $\mathbf{s}_p$ is determined only up to an invertible $n_p \times n_p$ matrix factor.

Minimal parameterization. If a matrix is determined up to right multiplication by an arbitrary invertible factor, only its column space is determined. Therefore, the appropriate parameterization of a MICA decomposition is in terms of the subspaces on which the components are obtained. This is the reason why we directly define above a MICA decomposition in terms of component subspaces. Algebraically, the component subspaces are in one-to-one correspondence with the orthogonal projectors onto them. Therefore a MICA decomposition may be defined by specifying an (unordered) set of linearly independent component subspaces $E_1, \ldots, E_c$ or, equivalently, an (unordered) set $\mathcal{P} = (\Pi_1, \ldots, \Pi_c)$ of orthogonal projectors onto these subspaces. Again, given the parameter $\mathcal{P}$, the components are uniquely determined from $\mathbf{x}$ via (4).

Before addressing some uniqueness issues, the case when the dimensions of the components is too large must be addressed. A given MICA decomposition $\{\mathbf{x}_1, \ldots, \mathbf{x}_c\}$ may be obtained by different sets of projectors (or subspaces) if $n > \sum_{p=1}^{c} \dim(E_p)$ *i.e.* if the random vector $\mathbf{x}$ does not 'fill in the whole space'. In this case, it exists a least one fixed (deterministic) vector $\mathbf{u}$ with unit norm such that $\mathbf{u}^\dagger \mathbf{x} = 0$ and any component subspace, say E_1, can be inflated to $\bar{E}_1 = E_1 \oplus \mathrm{Span}(\mathbf{u})$ (or, equivalently, Π_1 can be changed to $\bar{\Pi}_1 = \Pi_1 + \mathbf{u}\mathbf{u}^\dagger$). We stress that this operation does *not* affect the MICA decomposition $\{\mathbf{x}_1, \ldots, \mathbf{x}_c\}$ itself *i.e.* all the *components* $\mathbf{x}_1, \ldots, \mathbf{x}_c$ remain unchanged when E_1 is increased into $\bar{E}_1$. In order to fix this indetermination, we require that the parameter $\mathcal{P} = (\Pi_1, \ldots, \Pi_c)$ associated to a given MICA decomposition $\{\mathbf{x}_1, \ldots, \mathbf{x}_c\}$ be minimal in the sense that, for all p, subspace E_p should have the smallest dimensionality required to accommodate component $\mathbf{x}_p$. For instance, if $\mathbf{x}_p$ have finite second order moments, then Π_p should be the orthogonal projector onto the range of the covariance matrix of $\mathbf{x}$.

Invariance. MICA decompositions are invariant in the following sense: if $\mathbf{x} \in \mathbf{R}^n$ admits of a MICA decomposition in c components $\{\mathbf{x}_1, \ldots, \mathbf{x}_c\}$ and M is an $m \times n$ matrix with full column rank then the random vector $\mathbf{y} = M\mathbf{x} \in \mathbf{R}^m$ admits of a MICA decomposition into c components $\{\mathbf{y}_1 = M\mathbf{x}_1, \ldots, \mathbf{y}_c = M\mathbf{x}_c\}$. This is an obvious property: the components *do* change under the linear transform; however they undergo the same transform as the original variable $\mathbf{x}$ (the term 'covariance' is therefore more appropriate than 'invariance' but often used with a different meaning).

Note in the passing that principal component analysis is *not* invariant in the above sense because it leads components which are always orthogonal (PCA is at best invariant under orthogonal transformations).

Canonical MICA. Some care is required to uniquely define a MICA decomposition: definition 1 is not sufficient in this respect as discussed in the next two items.

i) Maximality of the decomposition. Assume (for instance) that a random vector $\mathbf{x}$ admits of a MICA decomposition in three components $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$. Then, according to the above definition of MICA, it also admits of another (coarser) decomposition in *two* components $\{\mathbf{x}_1 + \mathbf{x}_2, \mathbf{x}_3\}$. This decomposition is 'weaker' than the decomposition in three components. Actually the coarsest MICA decomposition compatible with definition 1 is just to take $\mathbf{x} = \mathbf{x}$ *i.e.* $\mathbf{x}$ has one component: itself!) Therefore, in order to avoid trivialities, a MICA decomposition should be requested to break a vector into as many independent components as possible.

ii) The Gaussian component. There is a difficulty if the result of breaking down a random vector into the largest possible number of independent components brings up more than one Gaussian component. Actually if a given component, say $\mathbf{x}_1$, of a MICA decomposition is normally distributed *and* has dimension $n_1 > 1$, it can always be decomposed into n_1 one-dimensional normally distributed sub-components, for instance by projecting $\mathbf{x}_1$ onto the first n_1 principal axis of its covariance matrix. However such a procedure would not yield an invariant decomposition because such sub-components are orthogonal by construction (see the above remark about the lack of invariance of PCA). One may think of other ways of decomposing a Gaussian vector into independent components but it does not seem possible to define a decomposition that would be invariant. In consequence, a special treatment is reserved to Gaussian components: they should *not* be split into independent subcomponents but rather all the Gaussian components (if any) should be kept pieced together as a unique component.

We may now define a canonical MICA decomposition of a random vector.

Definition 2 *The canonical MICA decomposition (if it exists) of a vector $\mathbf{x}$ is the unique MICA decomposition of $\mathbf{x}$ into $\mathbf{x} = \sum_{p=1}^{c} \mathbf{x}_p$ such that i) there is at most one Gaussian component and ii) no non-Gaussian component can be further decomposed into independent components.*

4. ICA FOR MICA

An ICA algorithm could be used for estimating a MICA decomposition by proceeding in two steps: **i)** run an ICA algorithm to obtain estimates of monodimensional source signals or an estimate of a mixing matrix **ii)** determine which source signals actually are independent and which should be grouped together as parts of a multidimensional component because they turn out *not* to be independent (see sec. 5 for an example) or because they are parts of the Gaussian component.

Letting aside the possibly non trivial task of performing step **ii**, it is important to determine how ICA algorithms designed to extract one-dimensional components behave when processing a mixture of multidimensional independent components *i.e.* to which extent they are able to perform step **i**. This is briefly examined below.

Many ICA algorithms are based, explicitly or not, on an estimating function $H : \mathbf{R}^n \mapsto \mathbf{R}^{n\times n}$ satisfying $\mathrm{E}H(\mathbf{s}) = 0$ when the entries of $\mathbf{s}$ (possibly scaled) are independent [2]. On-line algorithms update a separating matrix B in such a way that stationary points are characterized by $\mathrm{E}H(B\mathbf{x}) = 0$ while off-line algorithms based on T samples are (possibly only asymptotically) equivalent to estimating a mixing matrix A as a solution of $1/T \sum_{t=1}^{T} H(A^{-1}\mathbf{x}(t)) = 0$. Typically, function H has one of the two forms

$$H_\psi(\mathbf{y}) = \psi(\mathbf{y})\mathbf{y}^\dagger - I \qquad (6)$$

$$H^\circ_\psi(\mathbf{y}) = \mathbf{y}\mathbf{y}^\dagger - I + \psi(\mathbf{y})\mathbf{y}^\dagger - \mathbf{y}\psi(\mathbf{y})^\dagger \qquad (7)$$

where $\psi : \mathbf{R}^n \mapsto \mathbf{R}^n$ is an entry-wise function: $[\psi(\mathbf{y})]_i = \psi_i(y_i)$. The general form of defs. (6) and (7) can be derived from a maximum likelihood approach in which function ψ is minus the derivative of the log-density of $\mathbf{s}$. Assuming independent entries in $\mathbf{s}$ yields the entry-wise form of ψ, which is specific of the BSS/ICA model. [1]

Assume that $\mathbf{x}$ follows the MICA model (2) with c components of dimensions $n_1, \ldots, n_c$, assuming for simplicity $n_1 + \cdots + n_c = n$. The question is: is there at least one matrix B such that $\mathrm{E}H(B\mathbf{x}) = 0$ and such that basis for $E_1, \ldots, E_c$ can be be obtained by selecting appropriate columns of B^{-1}? It is not difficult to show that under simple assumptions, this is indeed the case, at least for estimating functions of the form (6) or (7). In this case, an ICA algorithm which finds this stationary point has successfully completed step **i** in presence of multidimensional independent components. In other words, some ICA algorithms do have stationary points which are solution of the MICA problem in the two-step approach outlined above. The next open question is to determine in a general setting the conditions for stability of these stationary points.

5. ILLUSTRATION: FETAL ECG

We illustrate the MICA decomposition on the case of fetal ECG. We use a data set [3] of $T = 2500$ ECG points sampled at 500 Hz with 3 electrodes located on the abdomen of a pregnant woman. The first second of signal is displayed in the first column of fig. 1. As step **i)**, we run the JADE algorithm [4], yielding (in about 500 K flops) this estimate of a 3×3 mixing matrix:

$$\hat{A} = [\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \hat{\mathbf{a}}_3] \approx \left| \begin{array}{ccc} 3.481 & -4.508 & 7.506 \\ 1.162 & -9.527 & -15.399 \\ -2.377 & 1.514 & 11.267 \end{array} \right| . \qquad (8)$$

Applying $\hat{A}^{-1}$ to the observations yields an estimate of 3 source signals which are displayed in the second column. Step **ii)** is here trivial to the human eye: the cardiac rhythms in the second column reveal that the algorithm has extracted one source signal coming from the fetus heart and two 'source signals' coming from the mother heart. Therefore, this data set seems well modeled by a MICA decomposition into one monodimensional (fetus) component and one bi-dimensional component (mother). Since the fetus signal clearly appears first in the second column, the orthogonal projection matrix on the fetus subspace is estimated by $\Pi_f = |\hat{\mathbf{a}}_1|^{-2}\hat{\mathbf{a}}_1\hat{\mathbf{a}}_1^\dagger$ and the orthogonal projection matrix Π_m on to the mother subspace is estimated as the orthogonal projection matrix onto the column space of $[\hat{\mathbf{a}}_2, \hat{\mathbf{a}}_3]$:

$$\Pi_m \approx \left| \begin{array}{ccc} 0.77 & 0.17 & 0.38 \\ 0.17 & 0.87 & -0.28 \\ 0.38 & -0.28 & 0.35 \end{array} \right| , \Pi_b \approx \left| \begin{array}{ccc} 0.63 & 0.21 & -0.43 \\ 0.21 & 0.07 & -0.14 \\ -0.43 & -0.14 & 0.30 \end{array} \right| .$$

[1] An extension of the ML approach to the MICA problem would be to use c non-linear functions $\psi_p : \mathbf{R}^{n_p} \mapsto \mathbf{R}^{n_p}$ and define $\psi(\mathbf{y}) = [\psi_1(\mathbf{y}_1)^\dagger, \ldots, \psi_c(\mathbf{y}_c)^\dagger]^\dagger$ where $\mathbf{y} = [\mathbf{y}_1^\dagger, \ldots, \mathbf{y}_c^\dagger]^\dagger$ is a split of vector $\mathbf{y}$ into c parts of size $n_1, n_2, \ldots, n_c$.

Using eq. (4) (or other equivalent but simpler expressions) yields

$$\tilde{\Pi}_m \approx \left| \begin{array}{ccc} 0.47 & 0.39 & 0.88 \\ -0.18 & 1.13 & 0.29 \\ 0.36 & -0.27 & 0.40 \end{array} \right|, \tilde{\Pi}_b \approx \left| \begin{array}{ccc} 0.53 & -0.39 & -0.88 \\ 0.18 & -0.13 & -0.29 \\ -0.36 & 0.27 & 0.60 \end{array} \right|$$

which allow to reconstruct a mother-only bidimensional component $\mathbf{x}_m = \tilde{\Pi}_m \mathbf{x}$ whose 3 entries are displayed in the third column and a fetus-only monodimensional component $\mathbf{x}_f = \tilde{\Pi}_f \mathbf{x}$ whose 3 entries are displayed in the last column. The same scale is used for each row of the figure in columns 1, 3 and 4 since vectors $\mathbf{x}$, $\mathbf{x}_m$ and $\mathbf{x}_f$ all live in the same space; the scale in the second column which represents vector $\mathbf{s}$ is arbitrary: it is *conventionally* determined by the JADE algorithm in such a way that the three entries of $\mathbf{s}$ have unit variance. Let us discuss these results.

1) A first important point is that fetal ECG extraction is possible by a BSS technique. This is important but only marginally related to our discussion. Similar successes have already been reported [5, 6] (better separation results are obtained with JADE when using the whole original data set in which 8 sensor outputs are available; the results presented herein are based only on the first 3 sensors due to lack of space and also because we only mean to illustrate the concept of MICA).

2) It is not possible to directly compare estimate (8) obtained by JADE to estimates of A obtained by other methods because, as already argued, matrix A has no real existence in the BSS/ICA model; only the projectors Π_m and Π_f (or functions of them) can be estimated unambiguously from the data in the MICA context. Similarly, the 'mother source signals' (rows 2 and 3 of the second column on the figure) have *a priori* no significance. Only the reconstructed mother signal $\mathbf{x}_m$ (third column of the figure) can be significantly compared to another $\mathbf{x}_m$ estimated with another ICA algorithm.

6. CONCLUSIONS

By reconsidering the notion of ICA, a more general perspective can be envisioned: multidimensional independent component analysis (MICA). It is based on a geometric parameterization which is free of the indeterminacies of matrix-based modeling. MICA relies on the idea of vector-valued component rather than on scalar 'source signals'. A canonical MICA decomposition is proposed as an invariant decomposition which can be empirically computed by post-processing the results of an ICA decomposition. This calls for further research on developing tools for detecting the existence of independent components in the ICA/MICA context.

Reproducible research. Inspired by the principle of Reproducible Research, we make the Matlab code used to produce the ECG figure freely available upon request or at
http://sig.enst.fr/∼cardoso/RRicassp98.html.

Acknowledgments. This work was initiated during a visit to Pr Amari at the Riken institute. We also gratefully acknowledge the people of the SISTA group at K.U. Leuven, the Netherlands for making the ECG data set available.

7. REFERENCES

[1] P. Comon, "Independent component analysis, a new concept ?," *Signal Processing, Elsevier*, vol. 36, pp. 287–314, Apr. 1994. Special issue on Higher-Order Statistics.

[2] J.-F. Cardoso, "Estimating equations for source separation," in *Proc. ICASSP'97*, pp. 3449–52, 1997.

[3] De Moor B.L.R. (ed.), "Daisy: Database for the identification of systems." http://www.esat.kuleuven.ac.be/sista/daisy, Oct. 1997.

[4] J.-F. Cardoso and A. Souloumiac, "Blind beamforming for non Gaussian signals," *IEE Proceedings-F*, vol. 140, pp. 362–370, Dec. 1993.

[5] L. D. Lathauwer, B. D. Moor, and J. Vandewalle, "Fetal electrocardiogram extraction by source subspace separation," in *Proc. HOS'95*, (Aiguablava, Spain), pp. 134–8, June 1995.

[6] E. Bacharakis, A. Nandi, and V. Zarzoso, "Foetal ECG extraction using blind source separation method," in *Proc. EUSIPCO*, (Trieste), 1996.