A GEOMETRIC APPROACH TO BLIND SEPARATION OF NONNEGATIVE AND DEPENDENT SOURCE SIGNALS

Wady Naanaa

Faculty of Sciences, University of Monastir 5000, Tunisia
phone: + (216) 21 568 516, fax: + (216) 73 500 278, email: naanaa.wady@planet.tn

ABSTRACT

Blind source separation (BSS) consists in processing a set of observed mixed signals to separate them into a set of original components. Most of the current blind separation methods assumes that the source signals are "as statistically independent as possible". In many real-world cases, however, source signals are considerably dependent. In order to cope with such signals, we proposed in [1] a geometric method that separates dependent signals provided that they are nonnegative and locally orthogonal.

This paper also presents a geometric method for separating nonnegative source signals which relies on an assumption weaker than local orthogonality. The separation problem relies on the identification of relevant facets of the data cone. After a rigorous proof of the proposed method, we give the details of the separation algorithm and report experiments carried out on signals from various origins, clearly showing the contribution of our method.

1. INTRODUCTION

Blind source separation (BSS) is a data processing task which is commonly know as the *cocktail-party* problem. A concrete description of the problem assumes we are in a cocktail-party where many groups of person are speaking. There are also many sensors (microphones) scattered in different spots. Each sensor performs a recording of the sound detectable in the spot where it is. The goal is to retrieve the speech given by each person from the various recordings.

There has been a considerable interest for separating source signals blindly because such a situation occurs in numerous fields such as analytical chemistry [2], communication [3], data mining [4], medical sciences [5]... More formally, solving a blind source separation problem consists in retrieving n unknown source signals from n of their mixtures, despite the lack of information about the mixing process. This can be expressed, in the instantaneous and noiseless case, by the following equation:

$$X = AS \tag{1}$$

where X is an $n \times m$ matrix holding in its rows the detected signals. A is an $n \times n$ mixing matrix whose entries are the unknown mixing coefficients and S is an $n \times m$ matrix holding in its rows the unknown source signals. A full identification of A and S is not possible because the sources can be permuted and scaled provided that the columns of A are transformed accordingly. More precisely, if P is a permutation matrix and Λ a nonsingular diagonal matrix then we have: $AS = (AP\Lambda)(\Lambda^{-1}P^{-1}S)$. The pairs (A,S) and $(AP\Lambda, \Lambda^{-1}P^{-1}S)$ are regarded as equivalent solutions in the sense of BSS problems. However, the model described by (1) is not viable in practice. Indeed, a more realistic model is the one involving additive noise:

$$X = AS + N \tag{2}$$

where N is an $n \times m$ matrix modeling the sensor noise. This of course yields a more difficult problem because the components of N are also unknown.

The blind separation of signals has been studied in the field of statistical signal processing where the concept of independent component analysis (ICA) is used to recover the source signals [6]. The mixtures and sources are therefore defined as sampled functions of n acquisition variables who may correspond to time slots, frequency slots, positions, wavenumber, etc... depending on the nature of the physical process under investigation. ICA-based methods focus on source signals that are statistically independent. But, in practice, these methods may be applied even if statistical independence does not hold. In such case, the goal is to find a set of source signals that are "as statistically independent as possible" given the observed signals. But the quality of the separation is degraded as dependence between sources increases. The separation process is achieved through the use of mathematical tools from second order statistics [7], fourth order statistics [6, 8] or information theory [9].

Estimated sources presenting as few dependence as possible, given the observed data, may not be satisfactory, especially when the true source signals are known to be correlated. This situation happens in many scientific fields, like in the medical or the chemical fields [2, 5]. The present work originates from analytical chemistry, more precisely, from nuclear magnetic resonance (NMR) spectroscopy of organic molecules in solution. There is no reason postulate the non-correlation of the NMR spectra of the molecules that form a mixture produced by a chemical reaction or extracted from a biological system. This is particularly true if the molecules share common structural features and therefore are difficult to separate by a chemical process. The NMR technique allows the practical realization of diffusion filters that modulates signal intensity according to the translational diffusion coefficient of the molecules [10]. Like in imaging techniques, the signals have positive values. Therefore, there is a need for blind separation algorithms that incorporate this constraint imposed by the physical origin of the signals.

In a previous paper, we proposed a geometric method designed for nonnegative and locally orthogonal source signals [1]. The blind source separation is therein expressed as the identification of the extreme directions of the cone containing the data. The geometric approach has been explored in many works ever since [18, 20, 19]. This paper presents an improvement of the method proposed in [1]. It addresses the weakening of the local orthogonality assumption. The latter is replaced by an assumption that can be informally described by a "good" distribution of the source points in the nonnegative orthant. We show that under the new hypothesis, the separation process can be expressed as the identification of relevant facets of the data cone.

The paper is organized as follows. In section 2, we remind some notions and properties of convex geometry. Section 3 provides the theoretical basis of our method. The separation algorithm is described in section 4. Section 5 reports experiments carried out on real-world source signals providing a comparison between the JADE algorithm [8] and ours. Finally, section 6 is a brief conclusion.

 ISSN 2076-1465

2. ELEMENTS FROM CONVEX GEOMETRY

Let us first remind some concepts and properties from convex geometry, the theory on which the proposed method is based. For more on the definitions and proofs omitted in this section see [11].

To begin with, we introduce some notations. We write $A \sqsubseteq B$ to mean that each column of matrix A is collinear to, at least, one column of matrix B. If $A \sqsubseteq B$ and A has less columns than B then we write $A \sqsubset B$. By $A^{\setminus i}$, (resp. $A_{\setminus i}$), we denote the submatrix obtained from A by removing the i^{th} column (resp. row).

2.1 convex cones

A subset $\mathscr{K}$ of Euclidean space $\mathbb{R}^n$ is a *convex cone* iff $\alpha_1 x_1 + \alpha_2 x_2 \in \mathscr{K}$ for all $x_1, x_2 \in \mathscr{K}$ and $\alpha_1, \alpha_2 \geq 0$. The *dimension* of a convex cone is given by the dimension of the smallest affine set containing it.

Let X be a matrix of $\mathbb{R}^{n\times m}$. The subset of $\mathbb{R}^n$ defined by

$$\mathscr{X} = cone(X) = \{X\alpha \mid \alpha \succeq 0\} \tag{3}$$

is a convex cone which is termed *polyhedral* and X is said to be a *generating matrix* of $\mathscr{X}$ since every element of $\mathscr{X}$ is a nonnegative linear combination of X columns.

Let $\mathscr{K}$ be a convex cone. $\mathscr{K}$ is said to be *pointed* if and only if it does not contain an element x and its opposite $-x$, unless x is a zero vector.

2.2 Extreme directions

Let $\mathscr{K}$ be a pointed convex cone. A nonzero vector $x_\varepsilon \in \mathscr{K}$ is an *extreme direction* of $\mathscr{K}$ iff for all $\alpha_1, \alpha_2 \geq 0$ and for all $x_1, x_2 \in \mathscr{K} \setminus \{\alpha_\varepsilon x_\varepsilon \mid \alpha_\varepsilon \geq 0\}$, we have $x_\varepsilon \neq \alpha_1 x_1 + \alpha_2 x_2$.
Clearly, if x_ε is an extreme direction of $\mathscr{K}$ then all $\alpha_\varepsilon x_\varepsilon,\ \alpha_\varepsilon \geq 0$ are also extreme directions of $\mathscr{K}$ and the subset $\{\alpha_\varepsilon x_\varepsilon \mid \alpha_\varepsilon \geq 0\}$ is called an *extreme ray* of $\mathscr{K}$. In what follows, all nonzero vectors belonging to the same extreme ray will be considered as identical directions.

If a matrix X holds exactly the set of extreme directions of a pointed polyhedral convex cone $\mathscr{X}$ arranged columnar then X is said to be a *minimal* generating matrix of $\mathscr{X}$. Let $\mathscr{X} = cone(X)$ be a pointed polyhedral convex cone in $\mathbb{R}^n$. If X is fat full-rank (i.e. $rank(X) = n$), then $\mathscr{X}$ is termed *proper*. $\mathscr{X}$ is called *simplicial* iff it has exactly n extreme directions.

2.3 Dual cone

For any convex cone, its dual is a unique cone which is always convex. When a polyhedral convex cone $\mathscr{X} = cone(X)$ is pointed, its dual $\mathscr{X}^*$ can be expressed by means of a generating matrix of $\mathscr{X}$:

$$\mathscr{X}^* = \{y \in \mathbb{R}^n \mid X^T y \succeq 0\} \tag{4}$$

If the extreme directions number of a pointed, polyhedral and convex cone does not exceed the ambient space dimension then its dual can be defined by

$$\mathscr{X}^* = \{X^{\dagger T}\alpha \mid \alpha \succeq 0\} \tag{5}$$

were $X^\dagger$ denotes the pseudo-inverse of X. Equations (4) and (5) define the same cone. These are respectively the *face* description and the *vertex* description of $\mathscr{X}^*$. All cones that are pointed, polyhedral and convex admit both descriptions. However, for pointed, polyhedral convex cones whose extreme directions number is in excess of the ambient space dimension, conversion between face description and equivalent vertex description is not trivial [11]. In fact, this is a well studied problem in convex geometry which can be solved by various algorithms [12, 13].

Properties 1 *For any convex cone $\mathscr{K}$ and its dual $\mathscr{K}^*$, we have:*
(i) $\mathscr{K}^{**} = \mathscr{K}$
(ii) *For any convex cone $\mathscr{K}'$, we have $\mathscr{K} \subseteq \mathscr{K}' \Rightarrow \mathscr{K}^* \supseteq \mathscr{K}'^*$.*
(iii) *$\mathscr{K}$ is polyhedral if and only if $\mathscr{K}^*$ is polyhedral.*
(iv) *$\mathscr{K}$ is proper if and only if $\mathscr{K}^*$ is proper.*
(v) *$\mathscr{K}$ is simplicial if and only if $\mathscr{K}^*$ is simplicial.*

2.4 Cone faces

A *face* of cone $\mathscr{K}$ is a cone $\mathscr{F} \subset \mathscr{K}$ such that for all $x \in \mathscr{F}$, if $x = x_1 + x_2$ with $x_1, x_2 \in \mathscr{K}$ then $x_1, x_2 \in \mathscr{F}$. If $\mathscr{K}$ has dimension n then a face of $\mathscr{K}$ having dimension $n-1$ is called *facet*.

Properties 2
(i) *The extreme directions of polyhedral proper cone $\mathscr{X}$ are respectively orthogonal to the facets of its dual $\mathscr{X}^*$; likewise, the extreme directions of polyhedral proper cone $\mathscr{X}^*$ are respectively orthogonal to the facets of $\mathscr{X}$.*
(ii) *Given a nonsingular matrix A, if cone(F) is a face of a polyhedral convex cone cone(S) then cone(AF) is a face of cone(AS).*

3. METHOD

For the purpose of establishing the theoretical background of the proposed method, we state the blind source separation problem by considering the noiseless case:

Problem 1 *Given a matrix $X \in \mathbb{R}^{n\times m}$ with $n \leq m$, find a nonsingular matrix $A \in \mathbb{R}^{n\times n}$ and a matrix $S \in \mathbb{R}^{n\times m}$ such that $X = AS$.*

In our case, we are concerned with the separation of mixtures obtained from nonnegative source signals. Hence, as a first working hypothesis, we assume that the source matrix S is nonnegative:

Hypothesis 1 $S \succeq 0$.

A consequence of hypothesis 1 is that the column vectors of X are constrained to be nonnegative linear combinations of the column vectors of A. Notice, however, that matrix X is not necessarily nonnegative because A may contain nonnegative coefficients. For this reason, nonnegative matrix factorization methods [17] cannot be applied to problem 1. With the aid of hypothesis 1, the following lemma expresses a week version of problem 1 in terms of a convex geometry problem.

Lemma 3 *Let A and X be two matrices with the same number of rows, then there exists $S \succeq 0$ such that $X = AS$ if and only if cone$(X) \subseteq$ cone(A).*

Proof:
($\Rightarrow$): Let x be in $cone(X)$. Then $x = X\alpha,\ \alpha \succeq 0$. It follows that $x = AS\alpha,\ \alpha \succeq 0$. But $S \succeq 0$, then $\alpha' = S\alpha \succeq 0$. We therefore obtain $x = A\alpha',\ \alpha' \succeq 0$, which is equivalent to $x \in cone(A)$.
($\Leftarrow$): For any column x^i of X, we trivially have $x^i \in cone(X)$. Since $cone(X) \subseteq cone(A)$ then $x^i \in cone(A)$ which is equivalent to $\exists\, s^i \succeq 0, x^i = As^i$. By applying this to every column of X, we obtain $X = AS,\ S \succeq 0$, where S is the matrix whose columns are the s^i's.

According to lemma 3, solving problem 1 amounts to finding a simplicial cone containing the data cone $\mathscr{X} = cone(X)$.

Our second assumption is that the n source signals are linearly independent.

Hypothesis 2 *The source matrix S is fat full-rank, i.e. $rank(S) = n$.*

The statement of problem 1 imposes that the mixing matrix A is nonsingular, then we deduce that data matrix X is also full-rank. And since S is nonnegative, we can easily show that $\mathscr{X}$ is necessarily pointed. Hence, $\mathscr{X}$ is a proper polyhedral cone. There is, in general, an infinity of simplicial cones containing a given proper cone. Indeed, suppose for instance that our data matrix is nonnegative then all simplicial cones containing the positive orthant, contain also the data cone and there is an infinity of such simplicial cones. So we must resort to supplementary hypothesis in order to limit the number of candidate solutions.

As a third hypothesis, we assume that the source points are spread in such a way that every facet of the nonnegative orthant

contains a facet of $\mathscr{S} = cone(S)$, the source cone. This can be equivalently expressed by imposing that S contains $n-1$ linearly independent column vectors that are orthogonal to every vector e_i of the standard basis of $\mathbb{R}^n$.

Hypothesis 3 *S contains $n-1$ linearly independent column vectors orthogonal to each e_i, $i : 1,\ldots,n$.*

Let us designate by $S_{i=0}$ the matrix composed of the columns of S which are orthogonal to e_i and by $\mathscr{S}_{i=0}$ the polyhedral convex cones generated by $S_{i=0}$.

Lemma 4 *$\mathscr{S}_{i=0}$ is a facet of $\mathscr{S}$, for $i : 1,\ldots,n$.*

Proof: Since $S_{i=0}$ is a submatrix of S whose rank is $n-1$ (hypothesis 3) whereas S has rank n, we have $\mathscr{S}_{i=0} \subset \mathscr{S}$. Now, we show that cones $\mathscr{S}_{i=0}, i : 1,\ldots,n$ are faces of $\mathscr{S}$. Let us prove that for every $s \in \mathscr{S}_{i=0}$ such that $s = s' + s''$ with $s', s'' \in \mathscr{S}$, we have $s', s'' \in \mathscr{S}_{i=0}$.

Since $s \in \mathscr{S}_{i=0}$ then $s_i = 0$. Moreover, $s' \in \mathscr{S}$, which implies that $s'_i \geq 0$. Therefore, in order to have $s = s' + s''$, s'_i must be zero. On the other hand, $s' \in \mathscr{S}$ then by distinguishing the columns of $S_{i=0}$ from the other columns of S, (which are denoted $S_{i\neq 0}$), we can write $s' = S_{i=0}\alpha + S_{i\neq 0}\bar{\alpha}$, with $\alpha, \bar{\alpha} \succeq 0$. More precisely, $s'_i = (S_{i=0})_i\alpha + (S_{i\neq 0})_i\bar{\alpha} = 0$ and since $(S_{i=0})_i$ is a zero vector and $(S_{i\neq 0})_i$ is a nonnegative and nonzero vector, $\bar{\alpha}$ must be zero. It follows that $s' = S_{i=0}\alpha,\ \alpha \succeq 0$, which is equivalent to $s' \in \mathscr{S}_{i=0}$. We can proceed in the same manner to show that s'' is also in $\mathscr{S}_{i=0}$. Thus, $\mathscr{S}_{i=0}$ is a face of $\mathscr{S}$. Finally, $S_{i=0}$ has rank $n-1$, then $\mathscr{S}_{i=0}$ is a facet of $\mathscr{S}$.

Lemma 5 *Each facet of $\mathscr{A}$ contains a facet of $\mathscr{X}$.*

Proof: According to lemma 4, $\mathscr{S}_{i=0} = cone(S_{i=0})$ is a facet of $\mathscr{S} = cone(S)$. It follows, by property 2-(ii), that $\mathscr{X}_{i=0} = cone(AS_{i=0})$ is a facet of $\mathscr{X} = cone(AS)$. But, the i^{th} row of $S_{i=0}$ is a zero vector. Then $AS_{i=0} = A^{\setminus i}(S_{i=0})_{\setminus i}$. On the other hand, by lemma 3, $cone(A^{\setminus i}(S_{i=0})_{\setminus i}) \subseteq cone(A^{\setminus i})$. Thus, $\mathscr{X}_{i=0} \subseteq cone(A^{\setminus i})$. But, $cone(A^{\setminus i}) \subset \mathscr{A}$ trivially verifies the face definition, moreover, the $n-1$ columns of $A^{\setminus i}$ are linearly independent. Hence, $cone(A^{\setminus i})$ is a facet of $\mathscr{A}$. The result follows.

Lemma 5 tells us that simplicial cone $\mathscr{A}$ may be determined through its facets which must contain facets of $\mathscr{X}$. These facets are in turn identified by considering the dual cones.

Lemma 6 *Each extreme direction of $\mathscr{A}^*$ is an extreme direction of $\mathscr{X}^*$.*

Proof: First, notice that $\mathscr{A}^*$ and $\mathscr{X}^*$ are both proper polyhedral cones of $\mathbb{R}^n$ since $\mathscr{A}$ and $\mathscr{X}$ are so. Let a^*_ε be a nonzero extreme direction of $\mathscr{A}^*$. According to property 2-(i), a^*_ε is orthogonal to one of the facets of $\mathscr{A}$. The latter facet contains, according to lemma 5, a facet of $\mathscr{X}$ which will be denoted by $\mathscr{F}$. Then a^*_ε is also orthogonal to $\mathscr{F}$. Again according to property 2-(i), $\mathscr{F}$ is orthogonal to one of the extreme directions of $\mathscr{X}^*$ which will be denoted by x^*_ε. Since $\mathscr{F}$ has dimension $n-1$, a^*_ε and x^*_ε are necessarily collinear. Moreover, a^*_ε and x^*_ε belong to the same ray. Indeed, since $\mathscr{X} \subseteq \mathscr{A}$ and then $\mathscr{X}^* \supseteq \mathscr{A}^*$, both x^*_ε and a^*_ε are in $\mathscr{X}^*$. Then x^*_ε and a^*_ε cannot have opposite directions otherwise $\mathscr{X}^*$ will not be pointed. Thus, a^*_ε is also an extreme direction of $\mathscr{X}^*$.

A simplicial cone $\mathscr{A} = cone(A)$ can be determined via its dual $\mathscr{A}^* = cone(A^*)$. Indeed, by (5), we have

$$A = A^{*-T} \tag{6}$$

Lemma 6 provides a necessary condition for a given direction a^* to be an extreme direction of $\mathscr{A}^*$: a^* must be an extreme direction of $\mathscr{X}^*$. This condition can be written as $A^* \sqsubseteq X^*$ where A^* and X^* are respectively composed of the extreme directions of $\mathscr{A}^*$ and $\mathscr{X}^*$ arranged columnar. On the other hand, we know that a simplicial cone is completely identified by the set of its extreme directions. We begin therefore by computing the extreme directions of $\mathscr{X}^*$ from which we can identify those of $\mathscr{A}^*$. So far, we only have a face description of $\mathscr{X}^*$ given by (4), since X is the only known matrix. Then, we have to obtain a vertex description of $\mathscr{X}^*$ consisting of a matrix X^* whose columns are the extreme directions of $\mathscr{X}^*$. This can be obtained by means of one of the existing algorithms allowing the conversion from a face description to an equivalent vertex description [12, 13].

As suggested by lemma 6, for simplicial cone $\mathscr{A}^*$ to be determined, one has to select n columns among those of X^*. There are, in general, many such combinations since X^* may contain more than n columns. It follows that, hypothesis 3 does not guarantee the uniqueness of the solution to problem 1. Nevertheless, it constrains the number of solutions to be finite. More precisely, we have $\binom{m^*}{n}$ solutions which simultaneously verify hypothesis 1, 2 and 3, where m^* is the number of extreme directions of $\mathscr{X}^*$.

Theorem 7 *The rows of S are among those of $X^{*T}X$.*

Proof: By (1) and (6), we have $S = A^{*T}X$. On the other hand, by lemma 6, each column of A^* is a column of X^* and then, each row of A^{*T} is a row of X^{*T}, we deduce that the rows of S are among the rows of $X^{*T}X$.

At this stage, having found a set of candidate source signals stored in the rows of $X^{*T}X$, we need a supplementary criterion in order to select those corresponding to the true source signals. We propose to select a set of n sources which are "as orthogonal as possible" given the data. Since we are concerned with nonnegative source signals, this requirement results in rather sparse source matrices (where only a few of the components are significantly active). Sparseness is a property which is often verified by nonnegative signals as outlined in [14]. An other advantage of maximizing orthogonality is that, in accordance with hypothesis 2, non full-rank source matrices should be avoided. Hence, we need to define degrees of orthogonality. To this end, we propose to use the Gram determinant as a measure of orthogonality:

$$\Gamma(S) = det(SS^T) \tag{7}$$

where the rows $s_i, i : 1,\ldots,n$ of S are assumed to have unit norm. The properties of the Gram determinant (see [15]) imply that $0 \leq \Gamma(S) \leq \prod_{i=1}^{n} \|s_i\|^2 = 1$. The signals stored in the rows of S are as orthogonal as the associated Gram determinant is close to one. The maximal value is reached when the rows of S are pairwise orthogonal. Conversely, a value of $\Gamma(S)$ close to zero reveals nearly linearly dependent source signals. Let $\mathfrak{X}$ denotes the set of all $n \times m$ matrices obtained by selecting any n rows from $X^{*T}X$ and scaling them to get unit row vectors. Then, as an estimate of the source matrix, we propose the matrix that maximizes the expression:

Hypothesis 4 $\tilde{S} = \arg\max_{S\in\mathfrak{X}} \Gamma(S)$.

Once, we have determined an estimate of source matrix $\tilde{S}$, we can deduce an estimate of the mixing matrix according to (1).

4. ALGORITHM

The DEDS algorithm (for Dual Extreme Direction-based Separation) (see Function 1) is an implementation of the blind source separation method described in section 3. It is composed of two main steps. The first step consists in computing X^*, a matrix that generates the dual of the data cone. This step can be achieved by the $O(mn^2m^*)$ algorithm proposed in [13] where m^* denotes the number of extreme directions of $\mathscr{X}^*$. It can also be performed by the *double description method* [12] for which there is an implementation running in $O(m^{\lfloor n/2 \rfloor})$ steps.

The second step consists in extracting an estimate of the source matrix by selecting n rows from those of the matrix product $X^{*T}X$. Note that by (4), $X^{*T}X$ is a nonnegative matrix. According to hypothesis 4, the n selected rows must be associated with the maximum value of function Γ. Finding a globally optimal solution is computationally a difficult task (we suspect that it is NP-hard). For efficiency reasons, we used a polynomial greedy algorithm which has experimentally proved to give satisfactory sub-optimal solution in short times. This algorithm uses the orthogonality measure described above which is, nevertheless, expressed in a more appropriate form. This is intended to avoid a repeated computation of determinants, a task that may slow down the separation process.

Using the Gram-Schmidt orthogonalization, we can easily prove that

$$\Gamma(S) = det(SS^T) = \prod_{i=1}^{n} \|r_i\|^2 \tag{8}$$

where the r_i's are iteratively defined as the difference between s_i and its orthogonal projection on the range of $r_1, r_2, \ldots, r_{i-1}$:

$$r_i = s_i - \sum_{k=1}^{i-1} \frac{s_i r_k^T}{\|r_k\|^2} r_k \tag{9}$$

The DEDS algorithm calls function OptGramDet which begins by scaling the rows of $X^{*T}X$ to obtain unit norm vectors. Denote by Y the resulting matrix. At iteration i, the algorithm determines the row s_i of Y which is associated with the maximum $\|r_i\|$ according to (9). At a first glance, s_1 can be any of the rows of Y since all these vectors have unit norm and by (9), we have $r_1 = s_1$. Nonetheless, the choice of s_1 has a great impact on the quality of the overall solution. This is because a bad choice of s_1 may lead to a low $\Gamma(S)$ value at the end of the optimization process. The clue to a good chose is to determine the direction (a row of Y) which is the more orthogonal to the others. Such a direction can be obtained by calculating first the mean vector $\bar{y} = \frac{1}{m}\sum_{i=1}^{m} y_i$. Then, by projecting all the y_i's on $\bar{y}$ and choosing as $\tilde{s}_1$ the y_i associated with the highest residual norm:

$$\tilde{s}_1 = \arg\max_i \|y_i - \frac{y_i \bar{y}^T}{\|\bar{y}\|^2}\bar{y}\| \tag{10}$$

Once, we have in hand the first vector $\tilde{s}_1$, we apply the process described above to get the $n-1$ remaining vectors and to form an estimate of the source matrix $\tilde{S} = (\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_n)$. The complexity of the selection process requires, at most, $O(n^2m^2)$ steps. Indeed, it includes n steps of m linear projections of vectors in $\mathbb{R}^m$ on a basis containing, at most, $n-1$ vectors. Finally, an estimate of the mixing matrix is obtained via (1).

Function 1 DEDS$(X) \rightarrow (A,S)$

1. $n \longleftarrow \text{rank}(X)$
2. $X^* \longleftarrow \text{Dual}(X)$
3. $S \longleftarrow \text{OptGramDet}(X^{*T}X, n)$
4. $A \longleftarrow XS^{\dagger}$

5. RESULTS

We compared the blind source separation results obtained by the DEDS algorithm presented in this paper with those obtained by the JADE algorithm [8]. The latter has been chosen because it has proved to be able to effectively separate a wide variety of signals without resorting to parameter tuning. We have carried out experiments involving real source signals from two origins:

- Nuclear magnetic resonance (NMR) spectra resulting from the analysis of chemical organic compounds [2]. We experimented on four source signals, each of which is composed of 7000 samples.
- Four of the EEG signals available on the ICALAB site [16] sampled over 512 frequency slots after conversion to the frequency domain.

The source signals were subsequently mixed according to the model described by (2). The mixing matrices were generated at random. Their elements are normally distributed with zero mean and unit variance. To show the impact of the mixing process on the performances of both algorithms, we experimented with three groups of mixing matrices. In each group, the determinants of the mixing matrices were varied in a specified interval. We used the following intervals: 0.5 ± 0.1, 0.05 ± 0.01 and 0.005 ± 0.001.

The additive noise is assumed to be white and Gaussian with uncorrelated samples having a variance which is assumed to be uniform over the rows of matrix N. The signal-to-noise ratio (SNR) was varied from 2 to 20 dB by a step of 2 dB. Each point appearing in the graphics of Figures 1 to 3 corresponds to the mean value of the performance index obtained on 1000 problem instances.

As a measure of algorithm performances, we used the Amari performance index [9] which is a real number varying in $[0,1]$. A signal separation is all the more correct as the value of the Amari performance index is close to zero.

Figures 1 to 3 depict the variation of the Amari performance index obtained in the various experiments. Each figure corresponds to one of the instance groups described above. In each graphic, we fixed the quality of the mixing process and the source signals from which the data are constructed and varied the signal-to-noise ratio.

By having a close look at the maximum and minimum points of the various curves, we can see that the quality of the mixing process is the most influential parameter. Indeed, the performances of both algorithms significantly degrade as the determinants of the mixing matrices decrease. On the other hand, the relevance of the signal-to-noise ratio is obvious with the exception of an unexplained behavior of JADE observed on the EEG instances of Figure 3 where we can see a weak rise of the performance index after having reached a minimum at 12 dB.

Now we turn to the comparison between the two algorithms. The quality of the mixing process has proved to be the parameter which affects the performance ratio of the two algorithms more. Indeed, on the two first groups of instances, i.e., those corresponding to the good and the medium mixing processes, DEDS is almost always superior to JADE. However, when we consider the third group of instances ($\Delta(A) \approx 0.005$), the trend is partially inverted in the experiment involving the NMR signals. This suggests that JADE is less sensitive to a degradation of the mixing process. Nonetheless, a compensation phenomenon seems to take place as we consider widely dependent source signals (EEG signals). And we are tempted to say that the performance deterioration that can be sustained by DEDS due to a bad mixing process can be recovered by a poor performance of JADE due to statistical dependence of signals. On the other hand, it is rather difficult to draw any general conclusion concerning the effect of the signal-to-noise ratio on the performance ratio of the two algorithms. On the first and third instance groups, the effect of the signal-to-noise ratio seems to be hidden by the effect of the mixing process quality. Conversely, on the second instance group, we can see (Figure 2) a rise in the performance ratio in favor of DEDS as the signal-to-noise ratio increases.

6. CONCLUSION

This paper presented a new blind source separation method which can be applied to nonnegative and statistically dependent source signals. The blind separation problem was expressed as the identification of relevant facets of the data cone. This task is achieved by an effective algorithm which first computes the dual of the data cone then selects, by means of a greedy process, a promising subset of source signals from a set of source candidates.

The application of the DEDS algorithm to simulated BSS instances involving real-world source signals showed that its perfor-

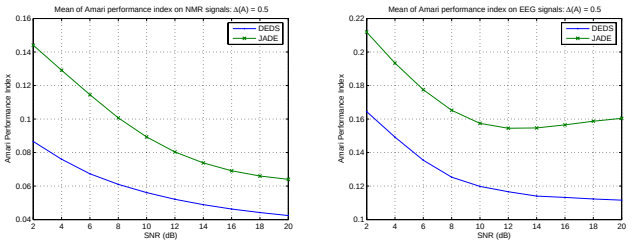

Figure 1: Mean of Amari performance index obtained with DEDS and JADE algorithms on randomly mixed real life signals. The random mixing matrices used in this experiment have their determinant varying in the interval 0.5 ± 0.1.

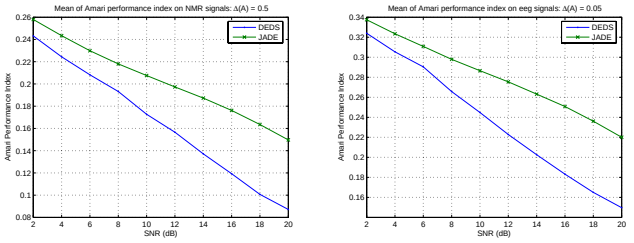

Figure 2: Mean of Amari performance index obtained with DEDS and JADE algorithms on randomly mixed real life signals. The random mixing matrices used in this experiment have their determinant varying in the interval 0.05 ± 0.01.

mance is highly competitive with one of the most popular blind separation algorithms: JADE. Further progress can be carried on, notably by weakening the assumption concerning the distribution of the source points over the nonnegative orthant. It is also conceivable to extend the method to cope with source signals with real values (not only nonnegative).

REFERENCES

[1] W. Naanaa and J.-M. Nuzillard, Blind Source Separation of Positive and Partially Correlated Data. Signal Processing 85(9) (2005) 1711–1722.

[2] D. Nuzillard and J.-M. Nuzillard, Application of blind source separation to 1-D and 2-D NMR spectroscopy, IEEE signal processing letters, 5 (1998) 209–211.

[3] W. Fu, X. Yang, N. Liu, X. Zeng, Blind source separation algorithm for communication complex signals in communication reconnaissance. Front. Electr. Electron. Eng. China, 3(3) (2008) 338–342.

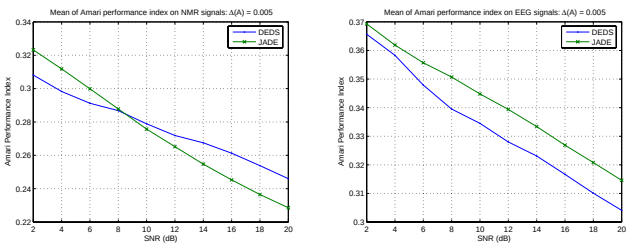

Figure 3: Mean of Amari performance index obtained with DEDS and JADE algorithms on randomly mixed real life signals. The random mixing matrices used in this experiment have their determinant varying in the interval 0.005 ± 0.001.

[4] V.P. Pauca, F. Shahnaz, M.W. Berry, R. J. Plemmons, Text mining using nonnegative matrix factorizations, in: Proc. of the 4^{th} SIAM Internationnal conference on data mining (SDM'04), (Lake Buena Vista, Fla., USA, 2004) 452–456, .

[5] F.J. Theis, Blind signals separation into groups of dependent signals using joint block diagonalization, ISCAS, 6 (2005) 5878–5881.

[6] P. Comon, Independent component analysis, A new concept?, Signal processing, 36 (1994) 287–314.

[7] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, E. Moulines, A blind source separation technique using second-order statistics, IEEE, Transaction on signal processing, 45 (1997) 434–444.

[8] J.-F. Cardoso and A. Souloumiac, Blind beamforming for non Gaussian signals, IEE Proceedings-F, 140(6) (1993) 362–370.

[9] S. Amari, A. Cichocki, H. H. Yang, A new learning algorithm for blind signal separation, in: Proc. Advances in Neural Information Processing Systems, 8 (1996) 757–763.

[10] W. S. Price, K. Hayamizu, H. Ide, Y. Arata, Strategies for Diagnosing and alleviating artifactual attenuation associated with large gradient pulses in PGSE NMR diffusion measurements, Journal of Magnetic Resonance, 139 (1999) 205–212.

[11] J. Dattorro, Convex Optimization & Euclidean Distance Geometry, (Meboo Publishing USA, 2005).

[12] T. Motzkin, H. Raiffa, G. Thompson, R. J. Thrall, The Double Description Method, in: Annals of Math Studies, Vol. 8, (Princeton University Press, 1953) 51–73.

[13] D. Avis and K. Fukuda, A pivoting algorithm for convex hulls and vertex enumeration of arrangements and polyhedra, Discrete Comput. Geom. 8 (1992) 295–313.

[14] D. J. Field, What is the goal of sensory coding? Neural Computation, 6 (1994) 559–601.

[15] S. S. Dragomir and B. Mond, On a property of Gram's determinant. Extracta Mathematicae, 11(2) (1996) 282–287.

[16] A. Cichocki, A. Amari, K. Siwek, T. Tanaka, *et al.*, ICALAB Toolboxes, `http://www.bsp.brain.riken.jp/ICALAB`.

[17] C.-J. Lin, Projected gradient methods for nonnegative matrix factorization, Neural Computation, 19(10) (2007) 2756–2779.

[18] T.-H. Chan, W.-K. Ma, C.-Y. Chi, Yue Wang, A Convex Analysis Framework for Blind Separation of Non-Negative Sources, IEEE Transactions on Signal Processing, vol. 56, pp. 5120–5134, 2008.

[19] T.-H. Chan, C.-Y. Chi, Y.-M. Huang, and W.-K. Ma, A convex analysis based minimum-volume enclosing simplex algorithm for hyperspectral unmixing, IEEE Trans. Signal Processing, vol. 57, no. 11, pp. 4418–4432, Nov. 2009.

[20] J. Li and J. Bioucas-Dias, Minimum volume simplex analysis: A fast algorithm to unmix hyperspectral data, In Proc. IEEE International Geoscience and Remote Sensing Symposium, vol. 4, Boston, MA, Aug. 8-12, 2008, 2369-2371.