Under-determined Sparse Blind Source Separation of Nonnegative and Partially Overlapped Data

Yuanchang Sun* and Jack Xin*

Abstract

We study the solvability of sparse blind separation of n non-negative sources from m linear mixtures in the under-determined regime $m < n$. The geometric properties of the mixture matrix and the sparseness structure of the source matrix are closely related to the identification of the mixing matrix. We first illustrate and establish necessary and sufficient conditions for the unique separation for the case of m mixtures and $m + 1$ sources, and develop a novel algorithm based on data geometry, source sparseness, and l_1 minimization. Then we extend the results to any order $m \times n$, $3 \leq m < n$ based on the degree of degeneracy of the columns of the mixing matrix. Numerical results substantiate the proposed solvability conditions, and show satisfactory performance of our approach.

Key Words: under-determined, non-negative sources, blind separation, sparseness, uniqueness, geometric method, l_1 minimization, clustering.

AMS Subject Classifications: 94A12, 65H10, 65K10, 90C05.

*Department of Mathematics, University of California at Irvine, Irvine, CA 92697, USA.

1 Introduction

The goal of this paper is to study blind source separation (BSS) problem of non-negative data when fewer mixture signals than sources are available. Such a case is referred to as under-determined. The under-determined blind source separation (uBSS) presents additional challenge than determined or over-determined BSS in that the mixing matrix is non-invertible. For simplicity, we consider the linear BSS model:

$$X = A\,S, \tag{1.1}$$

where $X \in \mathbb{R}^{m\times p}$ is the mixture matrix containing known mixed signals as its rows, $S \in \mathbb{R}^{n\times p}$ is the unknown source matrix, $A \in \mathbb{R}^{m\times n}$ is the unknown mixing matrix. All matrices are non-negative. The dimensions of the matrices are expressed in terms of three numbers: (1) p is the number of available samples, (2) m is the number of mixture signals, and (3) n is the number of source signals. Both X and S are sampled functions of an acquisition variable which may be time, frequency, position, or wavenumber depending on the measurement device. The mathematical problem is to estimate non-negative A and S from X, which is also known as non-negative matrix factorization (NMF).

BSS has found numerous applications in areas from engineering to neuroscience [5, 9, 10, 18, 21, 22, 23, 25], and a number of methods have been proposed based on *a priori* knowledge of source signals such as spatio-temporal decorrelation, statistical independence, sparseness, etc. For instance, independent component analysis (ICA [9, 10, 11, 12, 21, 22, 23]) recovers statistically independent source signals and mixing matrix A. The statistical independence requires uncorrelated source signals, and this condition however does not always hold in real world problems. Hence ICA methods practically look for approximately independent components. Recently there have been several studies of ICA and its applications in computer tomography, biomedical image processing, where non-negative constraints are imposed for the mixing matrix A and/or estimated source signals S [7, 16, 27, 28, 31]. The present work is motivated by the Nuclear Magnetic Resonance (NMR) spectroscopy data which should not be assumed to satisfy statistical independence, especially when the molecules responsible for each source may share common structural features [30]. Besides, the properly phased absorption-mode NMR spectral signals from a single-pulse experiment are positive. Therefore ICA-based methods would not work for this class of data. Although the NMF introduced by Lee and Seung in [20] does not assume the statistical independence of the source components, the NMF algorithms in general converge to different solutions on each run due to the *non-convexity* of the problem. For NMR data, a better working assumption is the partial source sparseness condition proposed by Naanaa and Nuzillard (NN) in [24]. The source signals are only required to be non-overlapping at some acquisition locations (see NNA in section 2). Such a local sparseness condition leads to a dramatic mathematical simplification of a general non-convex NMF problem. Geometrically speaking, the problem of finding mixing matrix A reduces to the identification of a minimal cone containing the columns of mixture matrix X. Linear programming is used to identify the cone in NN's approach, while authors of [17] proposed a geometric algorithm called extreme vector algorithm (EVA) to find the spanning edges of the cone. The working condition for EVA is

called extreme data property, which essentially is the same as NN's source condition. In fact, NN's sparseness assumption and the geometric construction of columns of A were known in the 1990's [1, 37] in the context of blind hyper-spectral unmixing. The analogue of NN's assumption is called pixel purity assumption. The resulting geometric (cone) method is the so called N-findr [37], and is now a benchmark in hyperspectral unmixing. NN's method can be viewed as an application of N-findr [37] to NMR data. However, the NN's approach and EVA method are designed for solving the determined or over-determined case $m \geq n$. For non-negative uBSS, new methods need to be developed. First, one may ask: Is the NN source sparseness assumption good enough for non-negative uBSS ?

There have been several studies on the uBSS of speech signals [7, 33, 34]. However, few results are available for the uBSS of non-negative and partially overlapped data (e.g. NMR signals). In [19], the authors extract three source spectra from two measured mixed spectra in NMR spectroscopy. Their method first recovers the mixing matrix A by clustering the mixture data in the wavelet domain, then solve for S via a linear programming. The source signals in [19] are assumed to be nowhere overlapping. Moreover, this method is limited to two mixtures. In this paper, we consider non-negative signals with overlap and study uBSS for arbitrary number of mixtures. We are particularly concerned with the conditions for unique solvability of A and S up to scaling and permutation. Motivated by NN's sparse condition, we further explore the geometric structure of column vectors of the mixture matrix. It turns out that an additional sparseness condition on the sources (besides that of NN's) and a delicate one degenerate column condition of the mixing matrix A are needed for the unique separation in uBSS. Geometrically, NN method can only recover the spanning edges of a minimal cone containing the column vectors of X, which is not enough in general to extract all column vectors of the mixing matrix A in uBSS. Our additional conditions allow the recovery of the remaining columns of A as special interior points of the cone which lie at intersections of certain hyperplanes. Counter examples are illustrated if these additional conditions fail. Under the additional conditions on the sparseness of S and the degree of degeneracy of A, we present a new algorithm which first retrieves A by combining NN and the geometric property of the mixtures, then recoveries S by solving an l_1 minimization problem.

The paper is organized as follows. In section 2, we review the essentials of NN's approach and its local sparseness assumption, then show by counter examples that extra conditions are needed for unique recovery in uBSS. In section 3, we introduce the extra sparseness condition on the sources to accomplish unique recovery up to scaling and permutation. The geometric study of mixture matrix suggests that this condition is in fact optimal. In section 4, we develop a new method of uBSS. We propose a novel algorithm to identify A based on the data geometry, then solve for S using l_1 minimization to ensure a sparse representation. In section 5, numerical experiments are performed to verify the optimal sparseness condition and test the effectiveness of our method. Various examples including real world data are prepared to show the reliability of the method. Additionally, clustering methods are discussed to recognize the hyperplanes in the geometric structure of the noisy data. In section 6, we generalize the method from treating mixing matrix of order $m \times (m+1)$ to any order $m \times n$, $3 \leq m < n$. Concluding remarks are in section 7.

This work was partially supported by NSF-ADT grant DMS-0911277. The authors thank Professor Stanley Osher for his interest and suggestions, and Mr. Jie Feng for helpful discussions.

2 Source Sparseness and Examples

In [24], Naanaa and Nuzillard (NN) presented an efficient sparse BSS method and its mathematical analysis for non-negative and partially overlapped signals in the (over)-determined cases of model (1.1) where $m \geq n$. The mixing matrix A is full rank [24]. In simple terms, NN's key sparseness assumption (NNA) on source signals is that each source has a stand alone peak at some location of acquisition variable where the other sources are identically zero. More precisely, the source matrix $S \geq 0$ is assumed to satisfy the following condition:

Assumption (NNA). *For each* $i \in \{1, 2, \ldots, n\}$ *there exists an* $j_i \in \{1, 2, \ldots, p\}$ *such that* $s_{i,j_i} > 0$ *and* $s_{k,j_i} = 0$ $(k = 1, \ldots, i-1, i+1, \ldots, n)$.

If equation (1.1) is written in terms of columns as

$$X^j = \sum_{k=1}^{n} s_{k,j} A^k, \qquad j = 1, \ldots, p, \tag{2.1}$$

the NNA implies that $X^{j_i} = s_{i,j_i} A^i$, $i = 1, \ldots, n$ or $A^i = \frac{1}{s_{i,j_i}} X^{j_i}$. Hence equation (2.1) is rewritten as

$$X^j = \sum_{i=1}^{n} \frac{s_{i,j}}{s_{i,j_i}} X^{j_i} \ , \tag{2.2}$$

which says that every column of X is a non-negative linear combination of the columns of $\hat{A}$. Here $\hat{A} = [X^{j_1}, \ldots, X^{j_n}]$ is the submatrix of X consisting of n columns each of which is collinear to a particular column of A. It should be noted that j_i $(i = 1, \ldots, n)$ are not known and have to be computed. Once all the j_i's are found, an estimation of the mixing matrix is obtained. The identification of $\hat{A}$'s columns is equivalent to identifying a convex cone of a finite collection of vectors [15]. The convex cone encloses the data columns in matrix X, and is the smallest of such cones. Such a minimal enclosing convex cone can be found by linear programming methods. For model (1.1), the following constrained equations are formulated for the identification of $\hat{A}$,

$$\sum_{j=1, j\neq k}^{p} X^j \lambda_j = X^k, \quad \lambda_j \geq 0, \quad k = 1, \ldots, p. \tag{2.3}$$

Then a column vector X^k will be a column of $\hat{A}$ if and only if the constrained equation (2.3) is inconsistent (has no solution X^j, $j \neq k$). However, if noises are present, the following optimization problems are suggested to estimate the mixing matrix

$$\text{minimize} \ \ \text{score} = \| \sum_{j=1, j\neq k}^{p} X^j \lambda_j - X^k \|_2 \ , k = 1, \ldots, p \tag{2.4}$$

$$\text{subject to} \ \ \lambda_j \geq 0 \ . \tag{2.5}$$

For each column, a score is associated to it. A column with a low score is unlikely to be a column of $\hat{A}$ because this column is roughly a non-negative linear combination of the other columns of X. On the other hand, a high score means that the corresponding column is far from being a non-negative linear combination of other columns of X. Practically, the n columns from X with highest scores are selected to form $\hat{A}$, the mixing matrix. The Moore-Penrose inverse $\hat{A}^+$ of $\hat{A}$ is then calculated and an estimate of S is achieved: $\hat{S} = \hat{A}^+ X$.

The NN method is very efficient in separating the NNA sources for determined and over-determined BSS problems. It is also robust in that major peaks could still be recovered when NNA is violated to certain extent. A recent study of the authors' investigated how to post-process with abundance of mixture data, and how to improve mixing matrix estimation with major peak based corrections [32]. Here we are interested in extending NN method to uBSS while maintaining the uniqueness of source recovery. The following two examples show that extra conditions are necessary.

Example 1: Let $(m, n) = (2, 3)$, and assume that the mixing matrix $A \in \mathbb{R}^{2\times 3}$ has pairwise linearly independent columns, and the remaining column is a non-negative linear combination of the other two. The source matrix satisfies NNA. The NN method can only detect the two columns which span the cone of column vectors of X, see A^1 and A^2 in Fig. 2.1. The remaining (third) column of A is contained in the cone, but it is impossible to identify it. Any interior vector in the cone could be a candidate.

Example 2: Let $(m, n) = (3, 4)$, and assume that none of A's four columns is a non-negative linear combination of the others. The source matrix satisfies NNA. Then the n columns of A form a convex cone enclosing the data points (columns of X). NN method recovers all the columns of A's by identifying the edges of the minimal bounding convex cone. Fig. 2.2 shows the four column vectors $A^1, \cdots, A^4$. By proper scaling of A^4, we arrange them on a plane. Any point contained in both triangle $A^1A^2A^3$ and $A^1A^2A^4$ admit two linear representations, indicating non-uniqueness of column vector of S.

The second example can be extended to $m \geq 3$ as well when columns of data matrix X admit multiple representations by column vectors of A. The examples suggest that solving uBSS requires extra sparse conditions on S in addition to NNA, thereby matrix S has a smaller number of freedom and a unique solution is more feasible. In the next section, we propose a maximum overlap condition on the NNA source signals to guarantee a unique separation of A and S.

3 Maximum Overlap Condition

Consider $m \geq 2$ mixtures and $n > m$ sources. We propose to strengthen NNA by the $(m-1)$-tuplewise maximum overlap condition (MOC) on the source signals:

Assumption (MOC-NNA)**.** *For each column of the source matrix S, there are at most $m-1$ nonzero entries. Furthermore, for each $i \in \{1, 2, \ldots, n\}$ there exists an $j_i \in \{1, 2, \ldots, p\}$ such that $s_{i,j_i} > 0$ and $s_{k,j_i} = 0$ $(k = 1, \ldots, i-1, i+1, \ldots, n)$.*

MOC puts a maximum overlap of $m-1$ entries or minimum sparseness condition on the columns of S. Simply said, this condition requires not only that each source has

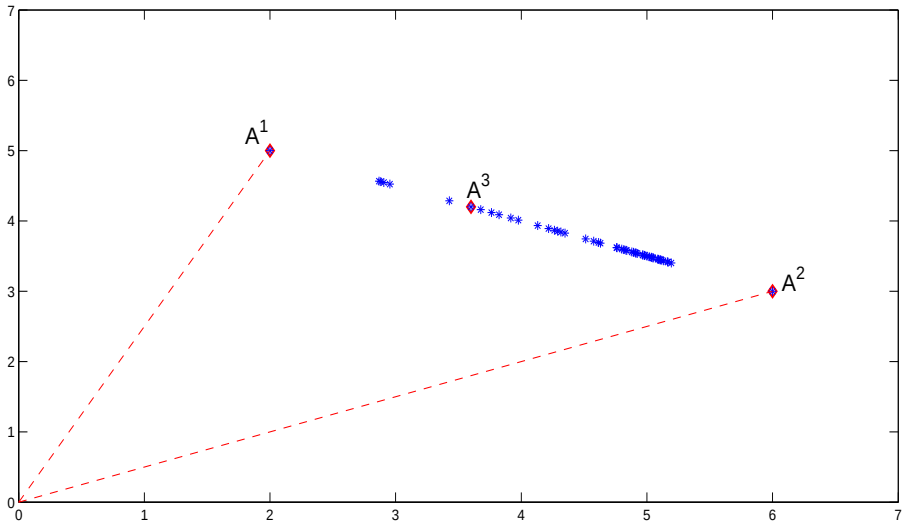

Figure 2.1: Red diamonds are the column vectors of the true mixing matrix A; blue stars represent the column vectors of X. NN's approach can uniquely identify A^1, A^2 which span the minimal cone on the data points. The third column A^3 can not be uniquely extracted from the data points. In fact, any blue star could be an estimation of A^3, thus the resulting source recovery is non-unique.

a stand alone peak at some location of acquisition variable where the other sources are identically zero, but also that there are at most $m-1$ active sources at each location of acquisition variable. Below we study several under-determined cases, and demonstrate uniqueness of recovery.

3.1 Two Mixtures

Consider 2 mixtures from $n > 2$ sources. MOC means that there is at most one active source in each location. Assume that the columns of the mixing matrix A are pairwise linearly independent, then the two dimensional column vectors of X can be visualized in Fig. 3.1. Estimation of column vectors of the mixing matrix A is easily achieved by clustering (e.g. K-means) [6], since there is only one nonzero entry in each column of S. MOC-NNA implies that

$$X^j = A^k s_{k,j} \ , \tag{3.1}$$

$s_{k,j}$ is the nonzero entry in the j-th column of S. Therefore, the source matrix S can be uniquely determined once A is recovered. The column vectors of X are aligned with n different lines shown in the left plot of Fig. 3.1. Moreover, the normalized n points are obtained when column vectors are projected to the unit circle as shown in the right plot of Fig. 3.1, and these n points can be the estimation of A's columns. The number of sources n is read off from the number of clusters on the unit circle, and may not need to be known in advance.

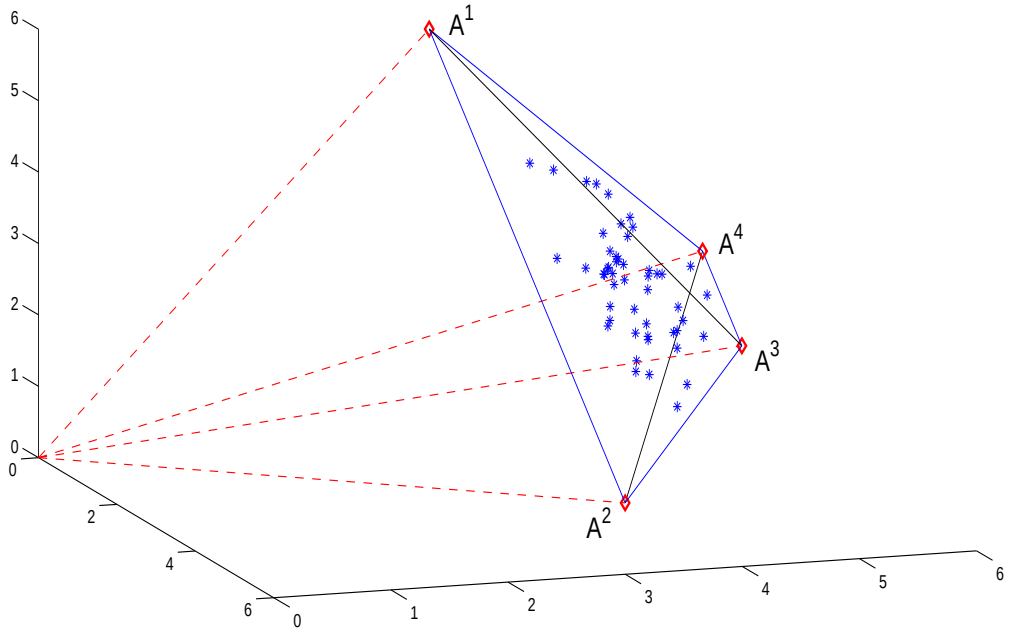

Figure 2.2: $A^1, \ldots A^4$ are the column vectors of A, which can be recovered by NN method. Blue stars contained in the cone are X's column vectors; they are linear combinations of $A^1, \ldots A^4$. However, their representations under basis A are not unique. For the points contained in both triangle $A^1A^2A^3$ and $A^1A^2A^4$, they have different representations, in other words, the source matrix S has different solutions.

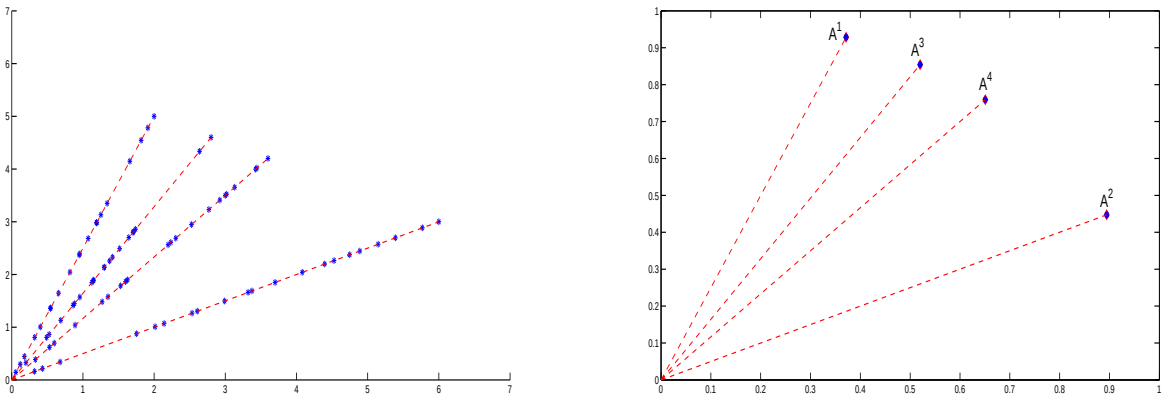

Figure 3.1: Two mixtures from four sources. The blue stars are column vectors of X; they are on four lines in the left plot. After normalization to unit l_2-norm, the column vectors of X cluster to four different vectors in the right plot.

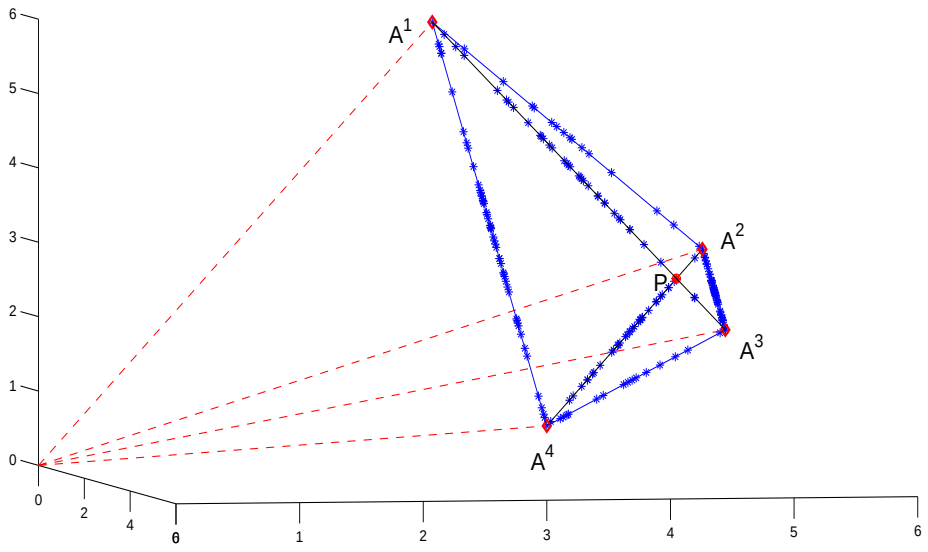

Figure 3.2: An example of 3 mixtures and 4 sources. A's column vectors define a convex quadrilateral, and X's column vectors (blue stars) lie on the edges and diagonals. Column vector P (red) is in the intersection of two diagonals, so $P = a_1A^1 + a_3A^3$ or $P = a_2A^2 + a_4A^4$. The corresponding column of S can not be uniquely determined. Matrix A violates one column degeneracy condition.

3.2 More Than Two Mixtures

If $m \geq 3$, MOC-NNA alone does not ensure a unique solution, suitable conditions on the mixing matrix A are also required. Fig. 3.2 shows the case of three mixtures and four sources, and we see that column vectors of X (blue stars) all lie on the lines defined by $A^1, \ldots, A^4$. For the solution of A and S, NN method can be used to recover A whose columns are the four vertexes of the quadrilateral. However, the source matrix S has no unique solution if there are data points at the intersection of the diagonals (point P in the plot). The column vectors of S are uniquely determined at all points but P. The column vector of X corresponding to P has two distinct representations: linear combination of either A^1, A^3 or A^2, A^4. High dimensional hyperplanes are difficult to visualize, but similar conclusion can be drawn.

Furthermore, Fig. 3.3 shows that uBSS has no unique solution for the case of m mixture and $n > m + 1$ sources in general. This example has three mixtures and five sources, and the sources satisfy MOC-NNA. The mixing matrix is assumed to have one degenerate column which is a non-negative linear combination of others (A^5 in the figure). The column vectors of the mixture matrix X (blue stars) are located in different lines, and each line is defined by two columns of A. Apparently the mixtures at P_1, P_2 and P_3 have different representations, e.g., P_1 can be a linear combination of A^1, A^3, or A^2, A^5. The solution for S is non-unique.

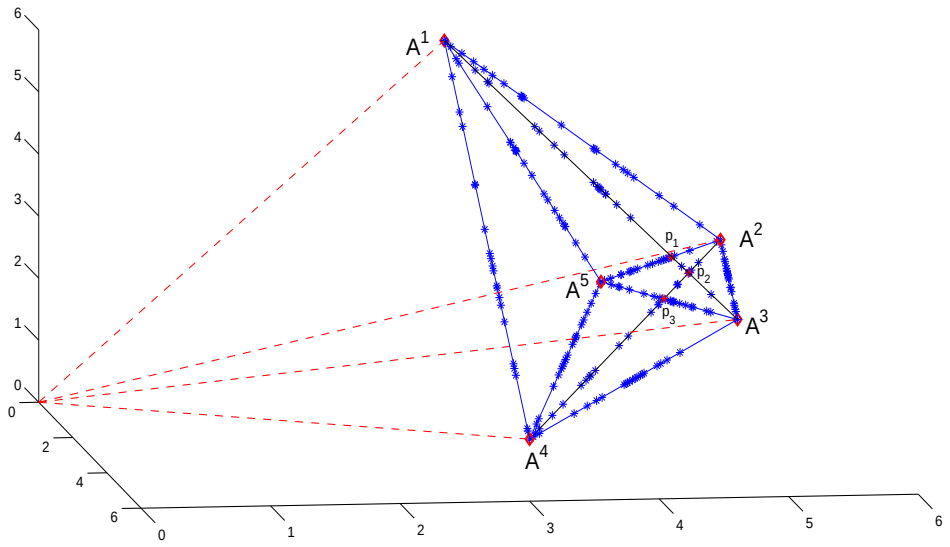

Figure 3.3: The geometry of three mixtures from five sources. Red diamond stands for the five columns of A; Blue stars are the column vectors of X.

3.3 $n = m + 1 > m \geq 3$

Let us consider $m \geq 3$ mixtures and $m + 1$ sources satisfying MOC-NNA condition. We further assume a degenerate mixing system: mixing matrix $A \in \mathbb{R}^{m\times(m+1)}$ has rank m, and one of its columns is a positive linear combination of other linearly independent columns. We shall call this assumption one column degenerate condition (OCDC). The main result is:

Theorem. *Up to scaling and permutation, the uBSS problem ($m \geq 3$ sources, $m+1$ mixtures) attains a unique factorization $A\,S$, provided that A satisfies OCDC, and S satisfies MOC-NNA.*

For example, Fig. 3.5 shows the configuration of three mixtures and four sources. The four red diamonds are the column vectors of A. The column vector A^4 is located inside the triangle $A^1A^2A^3$. Column vectors of X (blue stars) lie on different lines determined by columns of A. The three vertexes of the triangle can be identified using NN's approach. The three lines inside the triangle meet at one point which is A^4. The plot suggests that each column vector of X has a unique representation under the basis A, achieving unique recovery of the sources S. The proof of the theorem is as follows.

Proof. We consider m mixtures and $m + 1$ sources. Each column of mixture matrix X can be treated as a point in m-dimensional space, while each column of mixing matrix A can be treated as a vector extending from the origin to a corresponding point in the positive orthant. Without loss of generality, we assume that the last column A^{m+1} of A is degenerate. The other columns $A^1, \ldots, A^m$ are assumed to be linearly independent as in the OCDC condition. Hence

$$A^{m+1} = \sum_{i=1}^{m} \lambda_i A^i, \quad \sum_{i=1}^{m} \lambda_i = 1, \quad \lambda_i > 0\,. \tag{3.2}$$

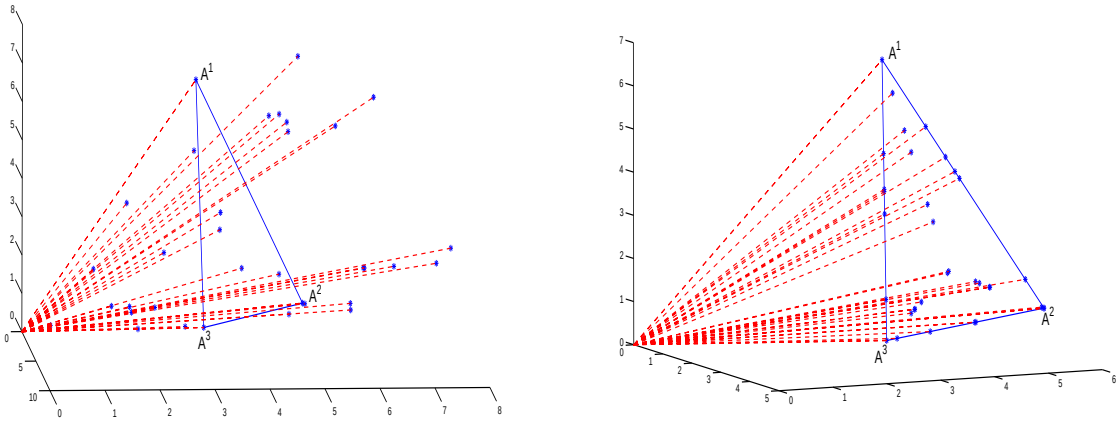

Figure 3.4: A cloud of non-negative data (left) rescaled to lie on a plane determined by A^1, A^2, A^3 (right).

A^{m+1} is contained in the convex cone spanned by $A^1, \ldots, A^m$ which can be identified from X's columns by NN's approach, and we shall call the cone Γ_A. Clearly Γ_A possesses scale invariance property, i.e., any multiple of a vector that belongs to Γ_A will also belongs to Γ_A. Next we will find the degenerate column A^{m+1} (or $\alpha A^{m+1}, \alpha > 0$) which is contained in X. Because of the scale invariance property, if the columns of X are rescaled to lie on an $(m-1)$-hyperplane, the minimal cone containing the rescaled data vectors actually is the cone Γ_A. Thus it is possible to restrict our attention to data on that hyperplane. This property is illustrated in Fig. 3.4, graphically the hyperplane cutting the cone forms a polyhedron. For the selection of the hyperplane, we consider the one determined by the m points $A^1, \ldots, A^m$ (the plane containing the triangle $A^1A^2A^3$ in Fig. 3.4). The construction of the hyperplane is discussed in Sec. 4.1.

Consider X's columns that are contained inside Γ_A, those interior points can be identified via linear programming. MOC-NNA implies that the interior points lie on different $(m-1)$-hyperplanes, and each hyperplane is spanned by vector A^{m+1}, and any $m-2$ vectors from $A^1, \ldots, A^m$. The intersection of these hyperplanes is the line connecting point A^{m+1} and the origin. Then the intersection point of the line with hyperplane determined by $A^1, \ldots, A^m$ is recognized as the estimation of A^{m+1}. For example, Fig. 3.5 shows that all the interior points are contained in three planes $OA^1A^4, OA^2A^4, OA^3A^4$, and the intersection of these planes is a line OA^4. Then the intersection of line OA^4 with plane $A^1A^2A^3$ is an estimate of A^4. In higher dimension, same idea applies for the determination of the degenerate column. In addition, the uniqueness of A^{m+1} can be expected. Thus the unique solution of A is achieved up to scaling.

The source recovery is equivalent to finding a sparse representation of X under A. Consider a column vector X^k of X, it is either inside the Γ_A or on its face. If X^k is located on the face, then it has the following form

$$X^k = s_1A^1 + s_2A^2 + \cdots + s_iA^i + \cdots + s_{m-1}A^{m-1} \ . \tag{3.3}$$

The question is whether this representation is unique. Apparently, it is unique in

terms of $A^1, \ldots, A^{m-1}$ which are linearly independent. However, we may have

$$X^k = s'_1 A^1 + s'_2 A^2 + \cdots + s'_{i-1} A^{i-1} + s_m A^m + s'_{i+1} A^{i+1} + s'_{m-1} A^{m-1} \, , m \neq i \, , \quad (3.4)$$

where $i \in \{1, \ldots, m-1\}$.

Subtracting (3.4) from (3.3) leads to

$$\begin{aligned} 0 \;\; = \;\; & (s_1 - s'_1) A^1 + \cdots + (s_{i-1} - s'_{i-1}) A^{i-1} + (s_{i+1} - s'_{i+1}) A^{i+1} \\ & + \cdots + (s_{m-1} - s'_{m-1}) A^{m-1} + s_i A^i - s_m A^m \, , \end{aligned}$$

which implies that $s_j = s'_j$ ($j = 1, \ldots, m-1$, and $j \neq i$), $s_i = s_m = 0$. Thus the representation (3.3) is unique.

Now suppose that X^k is a point inside the cone Γ_A. Without loss of generality, we assume

$$X^k = \sum_{i=1}^{m-2} c_i A^i + c A^{m+1} \, , \quad (3.5)$$

where $c_i \geq 0, c > 0$.

Then X^k has a unique representation under $A^1, \ldots, A^{m-2}, A^{m+1}$. In fact, suppose that

$$X^k = \sum_{i=1}^{m-2} c'_i A^i + c' A^{m+1} \, . \quad (3.6)$$

Then

$$\begin{aligned} 0 \;\; = \;\; & \sum_{i=1}^{m-2} (c_i - c'_i) A^i + (c - c') \sum_{i=1}^{m} \lambda_i A^i \\ = \;\; & \sum_{i=1}^{m-2} \big[(c_i - c'_i) + \lambda_i (c - c') \big] A^i + \sum_{i=m-1}^{m} \lambda_i (c - c') A^i \, . \end{aligned}$$

It follows from $\lambda_i > 0$ that $c = c', c_i = c'_i$.

Next assume that X^k can be represented by a different set of $m-2$ column vectors of $A^1, \ldots, A^m$ and A^{m+1}, or

$$X^k = c_{i_1} A^{i_1} + c_{i_2} A^{i_2} + \cdots + c_{i_{m-2}} A^{i_{m-2}} + c' A^{m+1} \, , \quad (3.7)$$

where $\{i_1, i_2, \ldots, i_{m-2}\} \subset \{1, 2, \ldots, m\}$. We subtract equation (3.7) from (3.5) to get

$$\begin{aligned} 0 \;\; = \;\; & (c - c') A^{m+1} + \sum_{i=1}^{m} b_i A^i \\ = \;\; & \sum_{i=1}^{m} \big[(c - c') \lambda_i + b_i \big] A^i \, , \end{aligned}$$

where b_i changes signs because $b_i > 0$ if i belongs to $\{1, \ldots, m\}$, but not in $\{i_1, \ldots, i_{m-2}\}$; $b_i < 0$ if the other way around.

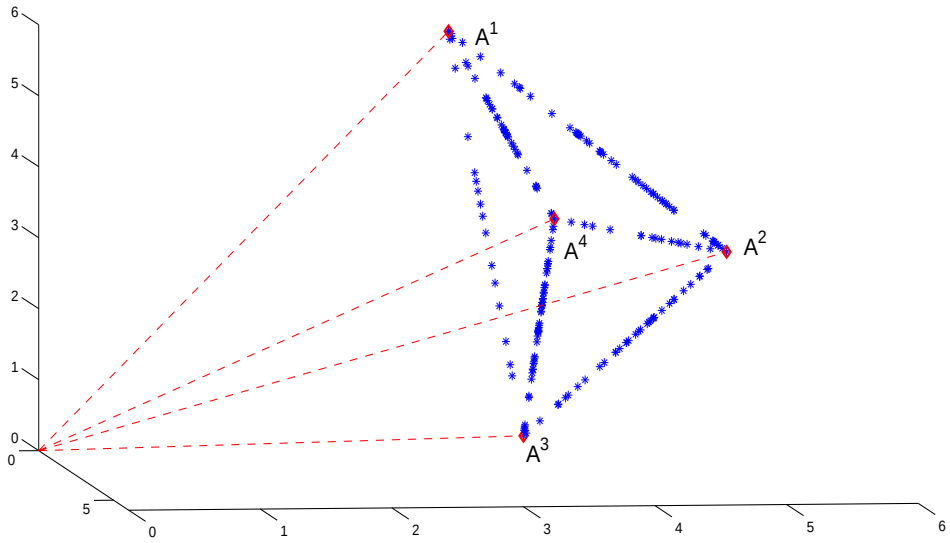

Figure 3.5: The interior red diamond is the degenerate column A^4 of A. The column vectors (blue stars) of X are located in several lines. Three inside lines intersect at A^4.

From the linearly independence of $A^1, \ldots, A^m$, it follows that

$$(c - c')\lambda_i + b_i = 0, \quad i = 1, 2, \ldots, m \; . \tag{3.8}$$

Suppose $c \neq c'$, say $c > c'$, then $b_i < 0$ since $\lambda_i > 0$. This is a contradiction because $b_1, \ldots, b_m$ change signs. As a consequence, c must be equal to c', and $b_i = 0$. This result implies that: except for $X^k = cA^{m+1}$, any other interior point of Γ_A lies on only one of the hyperplanes defined by A^{m+1}, and any $m-2$ vectors of $A^1, \ldots, A^m$. In other words, among the interior points only cA^{m+1} lies in the intersection of any two hyperplanes. In fact, this result provides a way to identify A^{m+1}. □

The MOC-NNA sources and OCDC mixing matrix guarantee the unique solution of uBSS. In next section, we propose an approach to retrieve A and S.

4 Geometric uBSS Method

Given $m \geq 3$ mixtures X, suppose that there are $m+1$ MOC-NNA sources S and a degenerate mixing system A. We propose a two-stage algorithm to determine A and S.

4.1 Degenerate Column

The m non-degenerate columns $A^1, \ldots, A^m$ of A can be identified from X by NN's method. MOC-NNA assumption implies that the degenerate column A^{m+1} is among the columns of X, and it is inside the cone Γ_A spanned by column vectors $A^1, \ldots, A^m$.

With $A^1, \ldots, A^m$ being recovered, we present the following algorithm to identify A^{m+1}:

1 Determine the $m-1$ dimensional hyperplane $\mathbf{H}$ defined by the points $A^1, \ldots, A^m$ in $\mathbb{R}^m$ (e.g. the plane $A^1A^2A^3$ in Fig. 3.5 for m = 3).

2 Scale the column vectors of mixture matrix X so that the scaled vectors lie on the hyperplane $\mathbf{H}$.

3 Construct two $m-1$ dimensional hyperplanes. Each hyperplane is spanned by a particular interior point, and a set of $m-2$ vectors from $A^1, \ldots, A^m$ such that this hyperplane contains A^{m+1}. The details are given in remark iii below.

4 Identify the points that are inside Γ_A.

5 Test all interior points, the one that lies on both hyperplanes in step 3 is taken as an estimate of A^{m+1} (see the second part of the proof in section 3).

Remarks:

i: Step 2 reduces the problem to a lower dimensional manifold. Fig. 3.4 provides a geometric illustration of this step. It should be noted that the data points can be scaled to lie on a different $(m-1)$-hyperplane, for example we can scale them to lie on $x_1 + \cdots + x_m = 1$.

iii. In step 3, the following constrained equations are solved for $\lambda = (\lambda_1, \ldots, \lambda_m)$,

$$\sum_{j=1}^{m} A^j \lambda_j = X^k, \quad \lambda_j > 0, \quad k = 1, \ldots, p \; . \tag{4.1}$$

Any column X^k will be an interior point if and only if the constrained equation (4.1) is consistent. If there is only one interior point, then it is A^{m+1}. When noises are present, we propose an alternative way to locate the interior points. The idea is as follows: we shall first identify the points lying on the faces of the cone. To achieve this, the following optimization problems are suggested to solve

$$\text{minimize} \;\; \text{score} = \| \sum_{j \in \{i_1, \ldots, i_{m-1}\}} A^j \lambda_j - X^k \|_2 \; , k = 1, \ldots, p \tag{4.2}$$

$$\text{subject to} \;\; \lambda_j > 0 \; , \tag{4.3}$$

where $\{i_1, \ldots, i_{m-1}\} \subset \{1, \ldots, m\}$. We set a tolerance ϵ for the score, and a column with a score lower than ϵ is recognized as a face point. Thus we can locate all the points that lie on the face spanned by $A^{i_1}, \ldots, A^{i_{m-1}}$. The value of ϵ depends on the noise level and needs to be tuned manually. We then repeat this process for all other different sets of $m-1$ vectors from $A^1, \ldots, A^m$. Finally, all the face points are obtained, and the retrieval of interior points is followed.

iii. Here we shall describe how to construct the hyperplanes and select the particular interior point in step 4. Any $m-2$ vectors from $A^1, \ldots, A^m$ plus an interior column vector of X span an $(m-1)$-hyperplane. For example, We consider to find the normal equation $\mathbf{n} \cdot (\mathbf{Y} - \mathbf{Y_0}) = 0$ of the hyperplane

spanned by $A^1, \ldots, A^{m-2}$ and an interior vector X^I. $\mathbf{n}$ is the normal vector which can be obtained by the singular value decomposition of the matrix $B = [A^1, A^2, \ldots, A^{m-2}, X^I]$. Note that B is an $m \times (m-1)$ matrix of rank $m-1$, and it has the factorization $B = VDU^{\mathrm{T}}$, where $U = [u_1, \ldots, u_{m-1}]$ and $V = [v_1, \ldots, v_m]$ are orthogonal $(m-1) \times (m-1)$ and $m \times m$ matrices, respectively, and where D is an $m \times (m-1)$ diagonal matrix with entries $d_{jj} > 0$ for $j = 1, \ldots, m-1$ and $d_{jk} = 0$ otherwise. Hence $B^{\mathrm{T}} v_m = 0$, which implies that v_m is the normal vector of the hyperplane spanned by B's column vectors. Furthermore, the hyperplane is through point Y_0, and a choice is to set $Y_0 = \frac{1}{m-1}\left(\sum_{j=1}^{m-2} A^j + X^I\right)$ to reduce the influence of noises in data. We continue to use $A^1, \ldots, A^{m-2}$ to illustrate how to select the particular interior point. The criterion is that the point needs to lie on the hyperplane spanned by $A^1, \ldots, A^{m-2}$ and A^{m+1}. Suppose the set of all interior points is denoted as $\mathcal{I}$, the set $\{A^1, \ldots, A^{m-2}\}$ as $\mathcal{S}$. The selecting process is described as follows:

(1) Set $\mathcal{J} = \mathcal{I}$. Pick $P \in \mathcal{I}$, define a hyperplane $\mathbf{H}$ by P and $\mathcal{S}$. Set $\mathcal{I} = \mathcal{I} - P$.

(2) Take $Q \in \mathcal{I}$ and set $\mathcal{I} = \mathcal{I} - Q$,

  (a) if Q lies on $\mathbf{H}$, then P is the desired interior point. Go to (3)

  (b) Otherwise,

    · if $\mathcal{I} \neq \emptyset$, go to (2),

    · if $\mathcal{I}$ is empty, set $\mathcal{I} = \mathcal{J} - P$ then go to (1).

(3) Output P and $\mathbf{H}$, stop the process.

Generally there is at least one interior point (excluding A^{m+1}) on each of the hyperplanes spanned by A^{m+1} and any $m-2$ vectors from $A^1, \ldots, A^m$. Therefore, the selection of the particular interior point will be successful and the output hyperplane $\mathbf{H}$ by this process contains A^{m+1}. Repeat the same process for a different set of $m-2$ vectors from $A^1, \ldots, A^m$, we obtain another hyperplane $\mathbf{G}$. With these two hyperplanes, we can identify A^{m+1} in step 5.

4.2 Recovery of Sparse Sources

For the recovery of S, we solve $X = AS$ for S, given A and X. Suppose the source signals satisfy the MOC-NNA sparseness condition, the theoretical result in section 3 guarantees a unique solution of S. We seek the sparsest solution for each column S^i of S as:

$$\min \|S^i\|_0 \quad \text{subject to } AS^i = X^i \ . \tag{4.4}$$

Here $\|\cdot\|_0$ (0-norm) represents the number of nonzeros. Because of the non-convexity of the 0-norm, we minimize the l_1-norm:

$$\min \|S^i\|_1 \quad \text{subject to } AS^i = X^i \ , \tag{4.5}$$

which is a linear program [13] because S^i is non-negative. Under certain conditions of matrix A, it is known [8, 36] that solution of l_1-minimization (4.5) gives the exact recovery of sufficiently sparse signal, or solution to (4.4), [8, 36]. Though our numerical

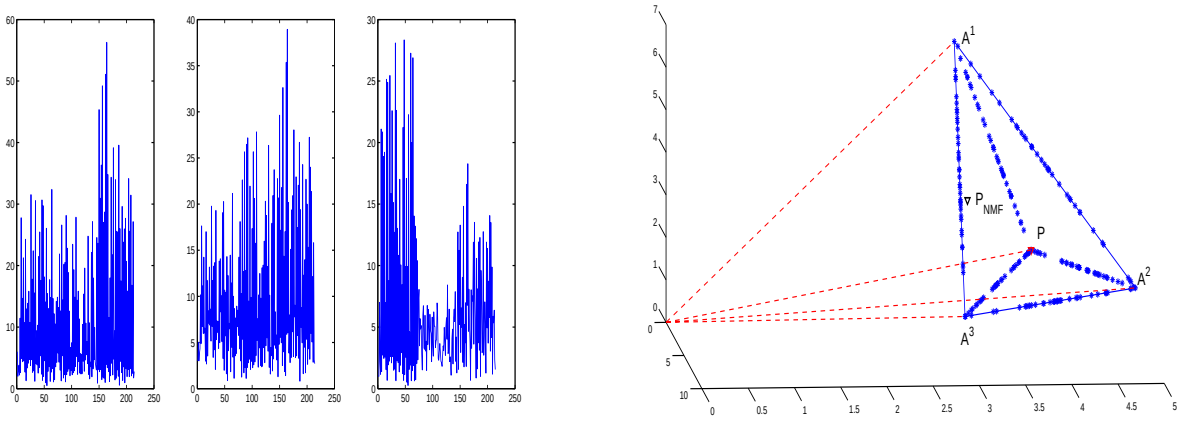

Figure 5.1: When plotted in $\mathbb{R}^3$, the three mixtures (left) has the geometric structure shown in the right plot. The read diamond is the degenerate column detected by the geometric approach, and it is the intersection of the two lines. The black diamond represents the result using NMF algorithm, which apparently deviates from the correct solution.

results support the equivalence of l_1 and l_0 minimizations, the mixing matrix A does not satisfy the existing sufficient conditions [8, 36]. Linear programming is sufficient for the examples we studied here, however larger size problems (larger values of m and n) may require an efficient iterative method such as Bregman iteration [35] for solving (4.5).

5 Numerical Experiments

We present numerical results to test our method, also to validate the unique solvability condition proposed in the paper. Two examples are presented. Example one is to retrieve four sources from three mixtures, while example two recovers five sources from four mixtures. The non-negative sources in both examples satisfy MOC-NNA assumption, and the positive mixing matrices satisfy OCDC.

The left plot in Fig. 5.1 shows the three mixtures of the first example, the plot on the right is the geometric structure of the mixtures in $\mathbb{R}^3$. The geometric method provides an exact recovery A_{GM} of A up to a permutation. As a comparison, the NMF's result is also presented here:

$$A = \begin{pmatrix} 6 & 2 & 3 & 3.8 \\ 3 & 5 & 3 & 2.8 \\ 1 & 1 & 7 & 2.2 \end{pmatrix}, A_{\text{GM}} = \begin{pmatrix} 2 & 6 & 3 & 3.8 \\ 5 & 3 & 3 & 2.8 \\ 1 & 1 & 7 & 2.2 \end{pmatrix}, A_{\text{NMF}} = \begin{pmatrix} 2 & 6 & 3 & 3.8 \\ 5 & 3 & 3 & 2.790 \\ 1 & 1 & 7 & 3.117 \end{pmatrix}.$$

Once the mixing matrix is obtained, the sources are recovered by l_1 minimization. Fig. 5.2 shows that the recovered sources agree very well with the ground truth. For example 2, the geometry of the four mixtures in Fig. 5.3 are difficult to visualize. Yet

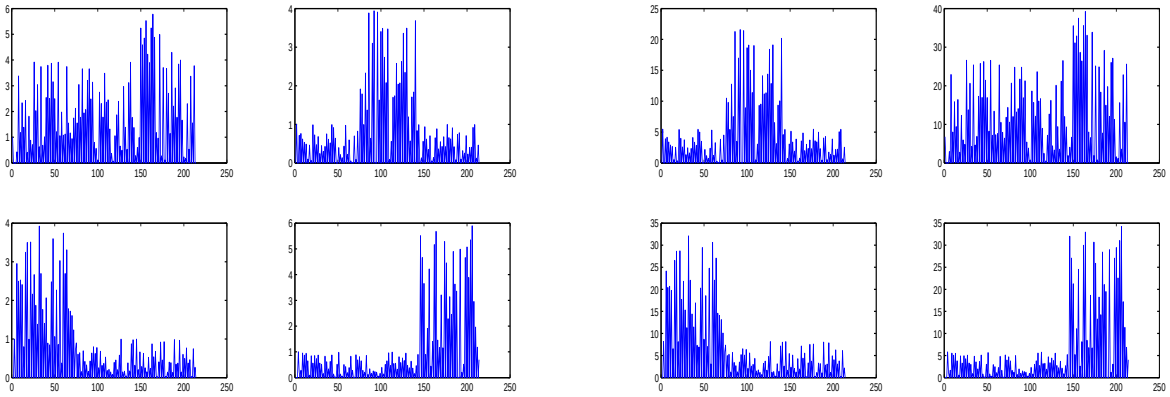

Figure 5.2: Left: the four true sources. Right: recovery by l_1 minimization.

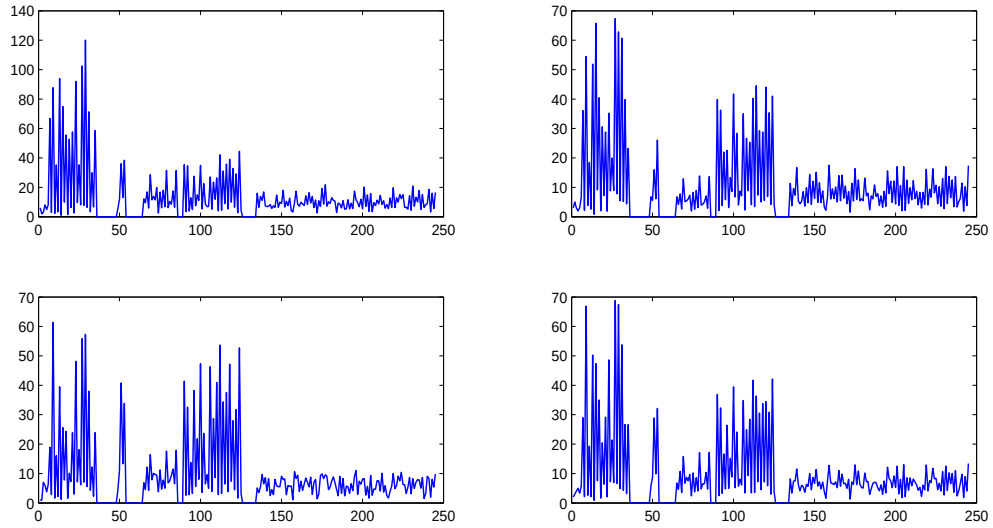

Figure 5.3: The four mixtures of example 2.

our algorithm still produced an exact recovery of the mixing matrix:

$$A = \begin{pmatrix} 6 & 2 & 3 & 8 & 4.9 \\ 3 & 5 & 3 & 2 & 3.1 \\ 1 & 1 & 7 & 6 & 3.8 \\ 2 & 3 & 4 & 5 & 3.35 \end{pmatrix}, \; A_{\text{GM}} = \begin{pmatrix} 2 & 6 & 3 & 8 & 4.9 \\ 5 & 3 & 3 & 2 & 3.1 \\ 1 & 1 & 7 & 6 & 3.8 \\ 3 & 2 & 4 & 5 & 3.35 \end{pmatrix}, \; A_{\text{NMF}} = \begin{pmatrix} 2 & 6 & 3 & 8 & 4.9 \\ 5 & 3 & 3 & 2 & 3.696 \\ 1 & 1 & 7 & 6 & 2.391 \\ 3 & 2 & 4 & 5 & 3.025 \end{pmatrix}.$$

The recovered sources in Fig. 5.4 shows again the reliable performance of the l_1 minimization.

5.1 Robustness

In this subsection, various examples are carried out to support the reliability of our approach. To test its robustness in the presence of noises, we varied the signal-to-noise

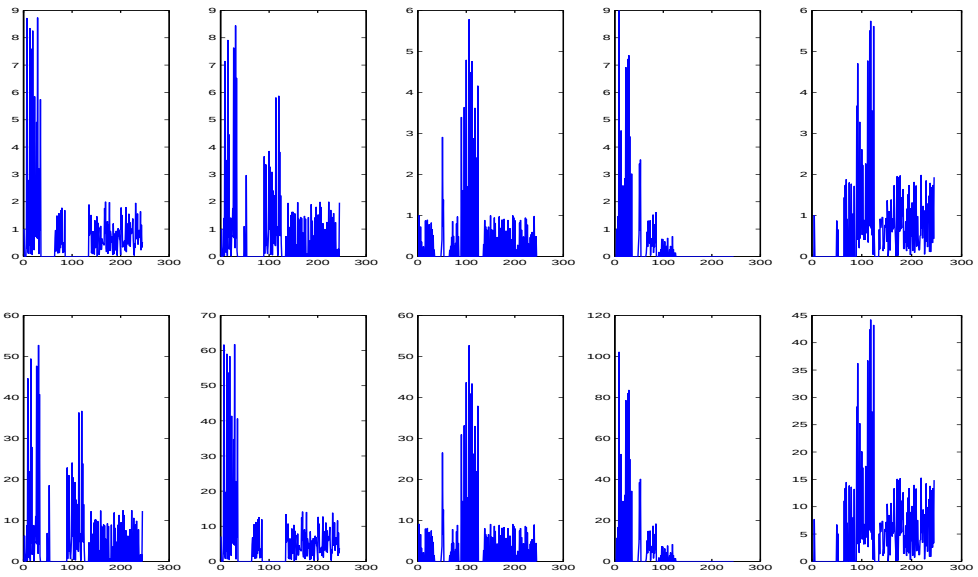

Figure 5.4: Top: The five true sources. Bottom: Recovered five sources.

ratio (SNR) when adding white Gaussian noise to the data. Fig. 5.5 is an example of three mixtures from four sources, which are obtained by adding Gaussian noise with SNR = 30 dB. The recovery of sources are presented in Fig. 5.6. In a second example, we used five sources to construct four noisy mixture signals (Fig. 5.7) by adding Gaussian noise with SNR = 40 dB. The result of the separation is presented in Fig. 5.8.

In a third example we apply the method to real world data. We used true Nuclear Magnetic Resonance (NMR) spectra of four compounds (mannitol, β-cyclodextrine,β-sitosterol) and menthol as source signals (data are from [24]). The NMR spectrum of a chemical compound is produced by the Fourier transformation of a time-domain signal which is a sum of sine functions with exponentially decaying envelopes. The real part of the spectrum can be presented as the sum of symmetrical, positive valued, Lorentzian-shaped peaks (see Fig. 5.10). Hence an NMR spectrum has nonzero responses everywhere. Therefore, the source signals in this case only satisfy a relaxed MOC-NNA condition:

Assumption. *For each column of the source matrix S, there are at most $m-1$ dominant entries. Furthermore, for each $i \in \{1, 2, \ldots, n\}$ there exists an $j_i \in \{1, 2, \ldots, p\}$ such that $s_{i,j_i} > 0$ dominates that column.*

The mixed signals were generated by (1.1) which is simulated by a 3×4 OCDC mixing matrix. The second plot in Fig. 5.9 is the geometric structure of the mixtures, where the degenerate column A^4 is identified as the intersection of the two lines. The separation result is presented in Fig. 5.10. The performance of the method can be seen clearly by comparison of the spectra in the two plots. The above examples demonstrate that our method is reliable in the regime where either the source conditions are violated to certain extent or the mixtures are noisy. However, every method has its limitation, our approach may fail to identify the degenerate column if higher level noise is present (SNR $\leq$ 25 dB). In this situation, some statistical techniques should be

used. Since the data points lie on different hyperplanes, the clustering analysis shall be applied to assign the points into clusters so that points in a cluster are regarded as in a hyperplane. The hyperplanes can then be constructed by least-square data fitting. In the following, we shall apply two clustering methods, Hough transform and spectral clustering, to recognize the hyperplanes in the noisy data. For the details of Hough transform and spectral clustering, the readers are refereed to [2, 3, 14, 26, 29].

The Hough transform is a feature extraction technique used in image analysis, computer vision, and digital image processing [29]. The simplest case of Hough transform is a linear transform for detecting straight lines, which can be used in the case of three mixtures four sources (three dimension). In the Hough transform, a main idea is to consider the characteristics of the straight line not as points x or y, but in terms of its parameters, here the slope parameter m and the intercept parameter b. Based on that fact, the straight line $y = mx + b$ can be represented as a point (b, m) in the parameter space. Computationally, it is common to use a different pair of parameters, denoted by r and θ, for the lines in the Hough transform. The parameter r represents the distance between the line and the origin, while θ is the angle of the vector from the origin to the closest point (see Fig. 5.11). Using this parametrization, the equation of the line can be written as

$$y = -\frac{\cos\theta}{\sin\theta}x + \frac{r}{\sin\theta} \tag{5.1}$$

which can be rearranged to $r = x\cos\theta + y\sin\theta$. It is possible to associate to each line a pair (r, θ) which is unique if $\theta \in [0, \pi)$ and $r \in \mathbb{R}$, An infinite number of lines can pass through a single point of the plane. If that point has coordinates (x_0, y_0) in the original plane, all the lines passing through it obey

$$r(\theta) = x_0\cos\theta + y_0\sin\theta,$$

where r is the same as the one in the equation (5.1). This corresponds to a sinusoidal curve in the (r, θ) plane, which is unique to that point. If the curves corresponding to two points are superimposed, the location (in the (r, θ) space) where they cross correspond to lines (in the original plane) that pass through both points. More generally, a set of points that form a straight line will produce sinusoids which cross at the parameters for that line. Thus, the problem of detecting collinear points can be converted to the problem of finding concurrent curves. This idea is demonstrated by the second plot in Fig. 5.11. In an example of three noisy mixtures from four sources, Hough transform is used to detect lines from the mixture geometry. The results are presented in Fig. 5.12. The results of this transform are stored in a matrix. Cell value represents the number of sinusoidal curves through any point. Higher cell values are rendered brighter. The six bright spots are the Hough parameters (r, θ) of the six lines. From the positions of these spots, (r, θ) values can be read off. Equations of the lines are given by (5.1), and columns of A are approximated as the intersections of the lines (see Fig. 5.13). The estimation of A by Hough transform (A_{HT}) is as follows (first row of A_{HT} is scaled to be same as that of A)

$$A = \begin{pmatrix} 0.8847 & 0.3651 & 0.3665 & 0.6544 \\ 0.4423 & 0.9129 & 0.3665 & 0.6544 \\ 0.1474 & 0.1826 & 0.8552 & 0.3789 \end{pmatrix},\ A_{\mathrm{HT}} = \begin{pmatrix} 0.8847 & 0.3651 & 0.3665 & 0.6544 \\ 0.4380 & 0.9357 & 0.3668 & 0.6459 \\ 0.1519 & 0.1892 & 0.8783 & 0.4039 \end{pmatrix},$$

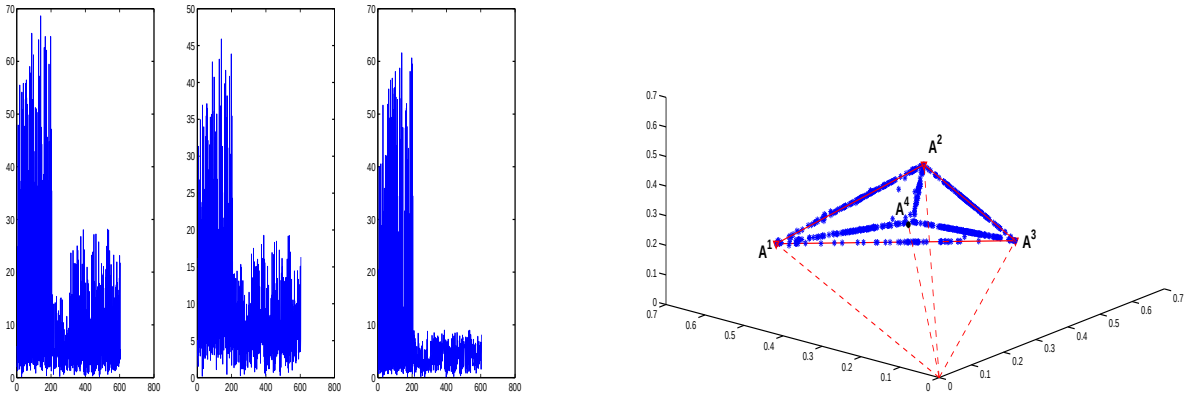

Figure 5.5: Signals of four source mixtures with noises. Three mixing spectra in the left plot and their geometric structure in the right plot where the black dot represents the approximation of the degenerate column.

and the source recovery is shown in Fig. 5.13. Note that the Hough transform detects lines in two dimensions. For higher dimensional data, we shall use spectral clustering technique to identify data points.

Spectral clustering has become one of the most popular clustering algorithms recently [3, 26, 29]. It is simple to implement, can be solved efficiently, and very often outperforms traditional clustering algorithms such as the k-means algorithm. Next we shall combine NN method and spectral clustering to retrieve columns of mixing matrix from the rather noisy data. We first run NN method to identify A's non-degenerate columns, then identify the interior data points by solving (4.2). Secondly, spectral clustering is applied to assign the interior points to different groups so that the points in the same group lie on the same hyperplane. Moreover, equations of the hyperplanes can be obtained by least-square data fitting. Finally, the degenerate column is identified as the intersection of these hyperplanes. We present here an example of four mixtures from five sources. The true mixing matrix and its estimation via NN and spectral clustering are (for ease of comparison, the first row of $(A_{\rm NS})$ is scaled to be same as that of A)

$$A = \begin{pmatrix} 0.8485 & 0.3203 & 0.3293 & 0.7044 & 0.6154 \\ 0.4243 & 0.8006 & 0.3293 & 0.1761 & 0.4190 \\ 0.1414 & 0.1601 & 0.7683 & 0.5283 & 0.4976 \\ 0.2828 & 0.4804 & 0.4391 & 0.4402 & 0.4452 \end{pmatrix}$$

$$A_{\rm NS} = \begin{pmatrix} 0.8485 & 0.3203 & 0.3293 & 0.7044 & 0.6154 \\ 0.4249 & 0.8138 & 0.3383 & 0.1756 & 0.4520 \\ 0.1408 & 0.1564 & 0.7854 & 0.5326 & 0.6116 \\ 0.2754 & 0.4826 & 0.4462 & 0.4394 & 0.5023 \end{pmatrix},$$

where the first four columns are non-degenerate and the last column is degenerate. The source separation results are shown in Fig. 5.14, the quality of the separation can be seen from the comparison between the real sources and their recovery.

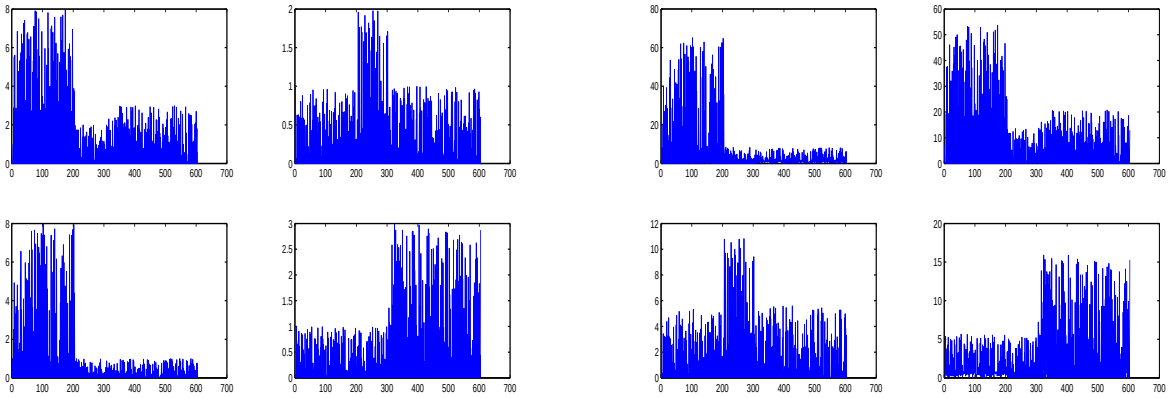

Figure 5.6: Left: Four true sources. Right: Recovered sources.

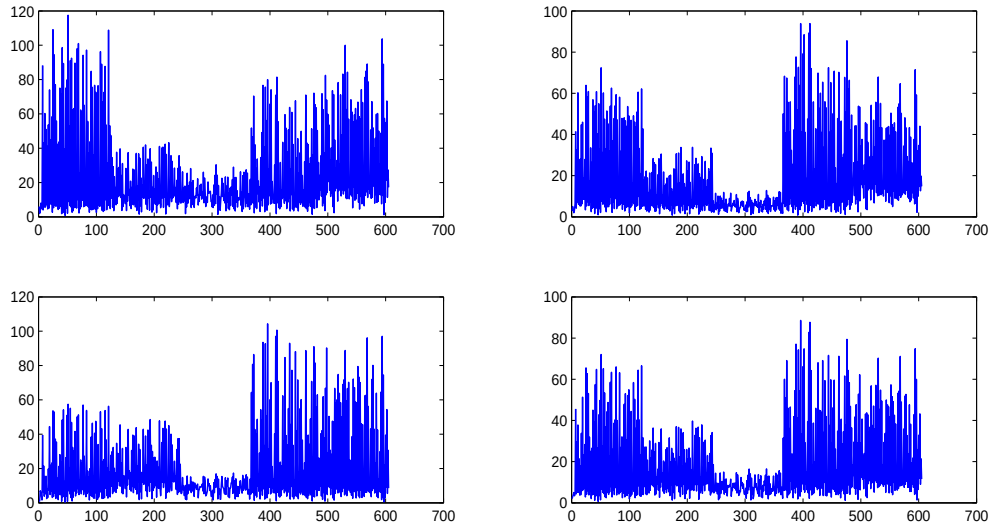

Figure 5.7: Four noisy mixtures.

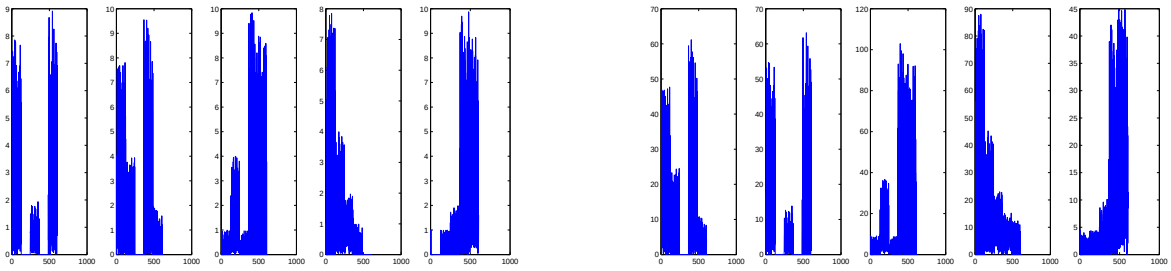

Figure 5.8: Left: Four true sources. Right: Recovered sources via our method.

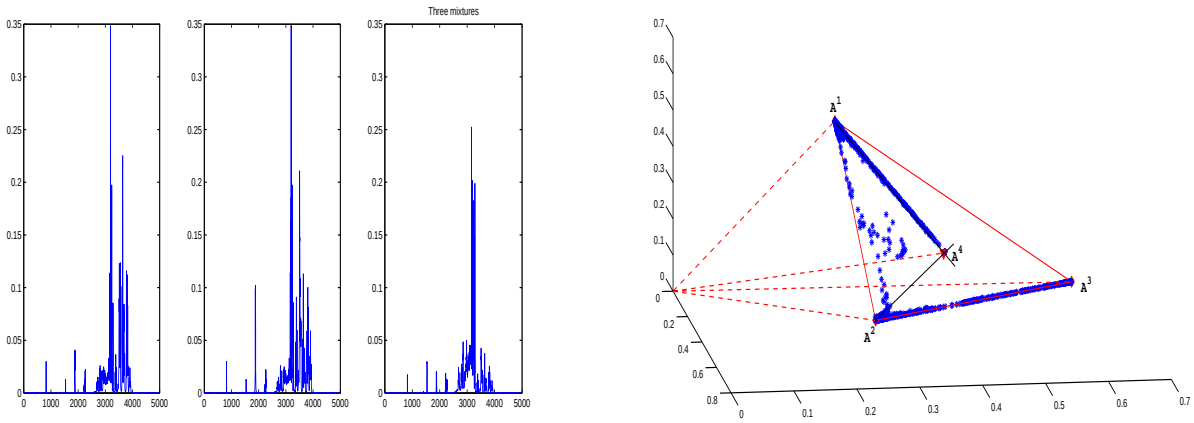

Figure 5.9: Three mixtures obtained by combining the spectra of menthol, β-sitosterol,mannitol,and β-cyclodextrine. Left: Three mixture spectra. Right: Their geometric structure.

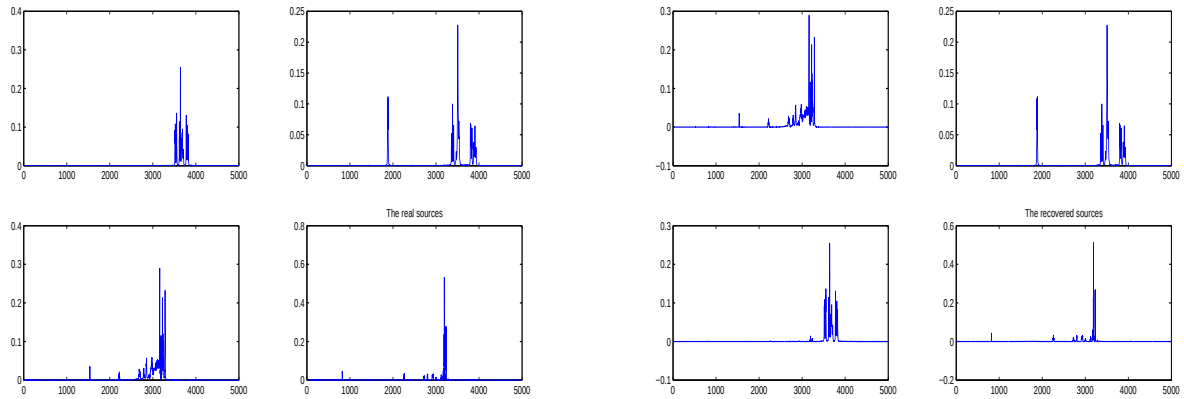

Figure 5.10: Left: The true source signals. Right: Sources signals recovered by our method.

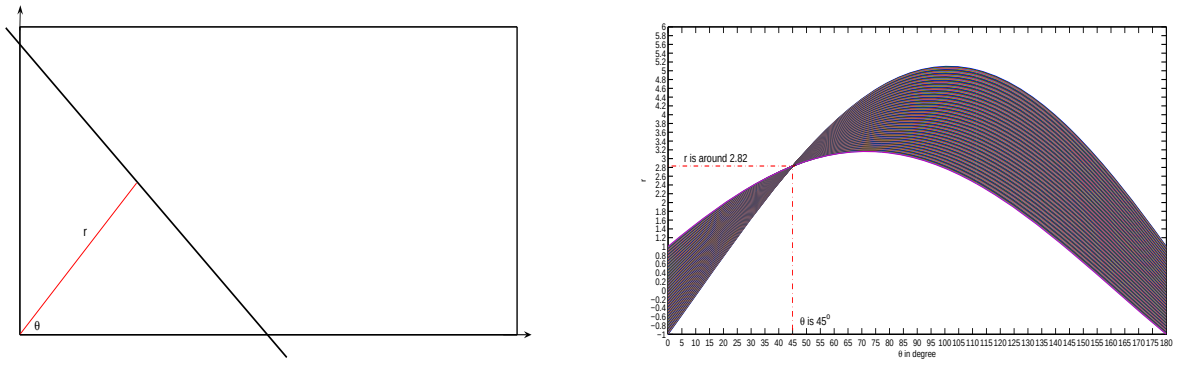

Figure 5.11: A straight line in (x, y) plane and its sinusoidal curves in the (r, θ) plane. The line's (approximate) geometric parameters (r, θ) are read off first, then its equation is obtained by formula (5.1).

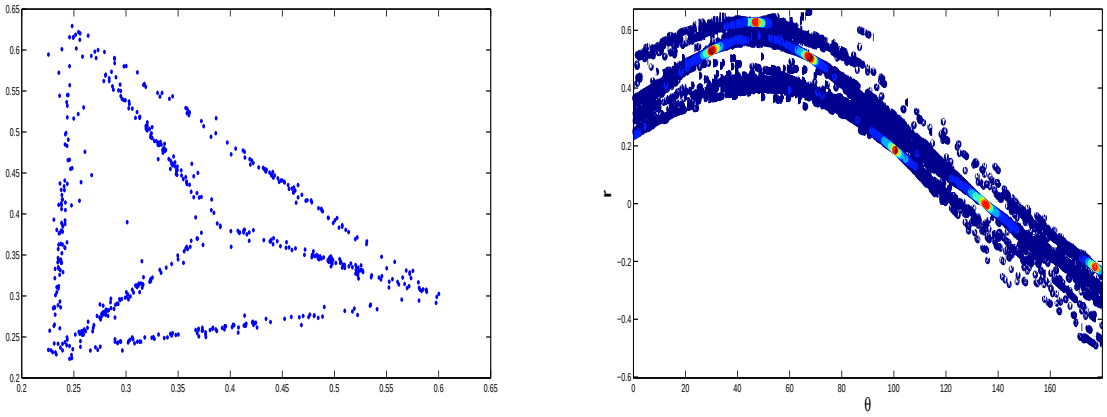

Figure 5.12: The geometry of the noisy mixture data is projected in the (x, y) space and its Hough transform in the (r, θ) space. The six bright color spots (red) imply that there are six lines, and their Hough parameters can be easily read off.

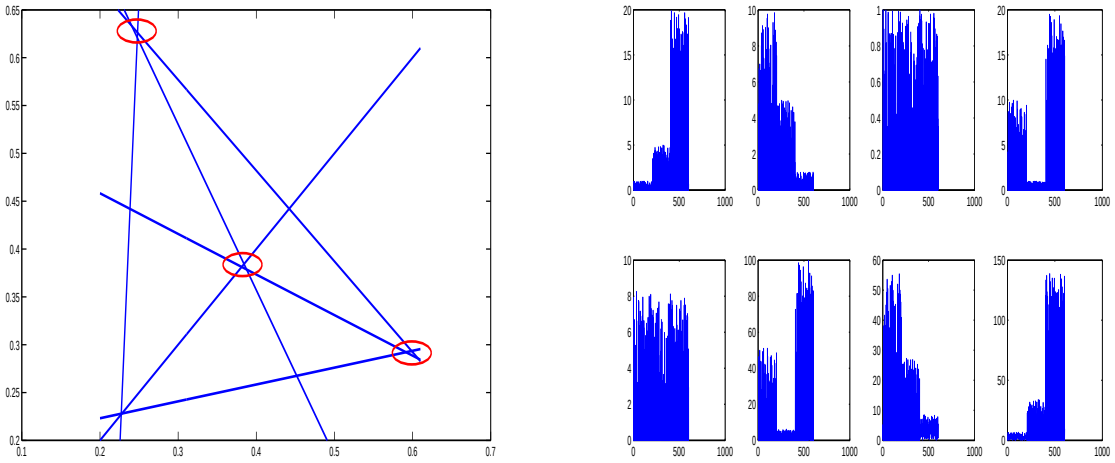

Figure 5.13: Left: Six lines are detected in Fig. 5.12, and notice that they are not concurrent in the circled region which is normal due to the presence of noise. All the neighboring intersecting points are computed and their average is taken to be the intersection point. Right: Original sources (up) and the recovery by method in section 4.2.

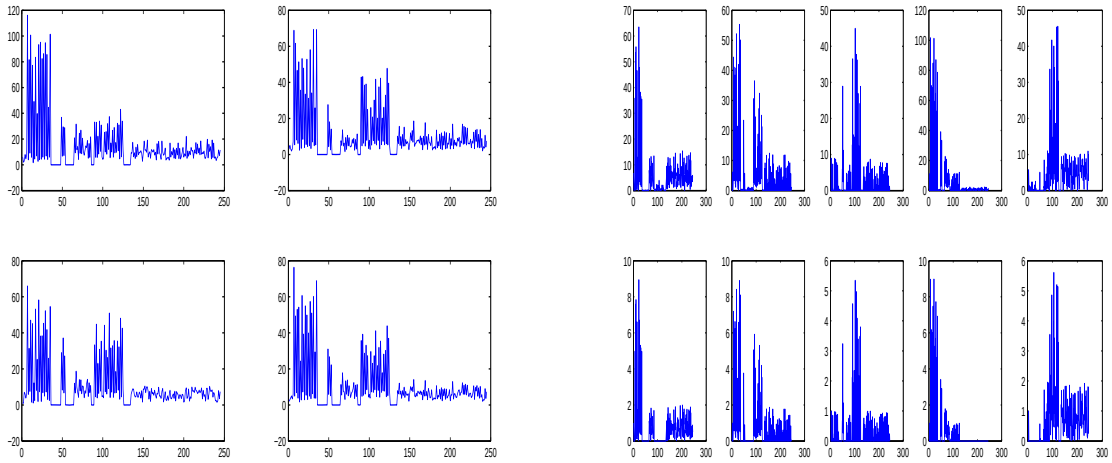

Figure 5.14: Left: The four mixtures. Right: the five true sources (up) and their recovery (down).

6 Extension to General Cases

In this section, we extend our uBSS geometric method from treating mixing matrix of order $m \times (m+1)$ to any order $m \times n$, $3 \leq m < n$. The extension is based on the degree of degeneracy of the columns of the mixing matrix, and allows multiple solutions. Note that the MOC condition is not needed for constructing columns of mixing matrix in the absence of degeneracy. In the degenerate regime, it is needed to search for interior intersection points by subspaces and their translations.

6.1 Degeneracy of Degree Zero

If there is no column of A that is a non-negative linear combination of other columns (zero degeneracy), then the columns form the edges of a convex cone in $\mathbb{R}^m$ under NN sparseness condition. The computation reduces to the identification of spanning vectors of the minimal cone containing the data set, which can be achieved by linear programming. Note that there may be infinitely many solutions for the sources because the mixing matrix is non-invertible. The l_1 norm minimization shall be used to ensure a sparse source solution. The numerical results are presented as follows. The first example is to separate out 4 sources from 3 mixtures, where the sources satisfy NNA condition and the mixing matrix has zero degeneracy. The mixtures and their geometry are shown in Fig. 6.1. It can be seen that the four columns are identified as the spanning edges of the convex cone containing the data set. The l_1 solution of the sources is in Fig. 6.2. Compared to the ground truth, the recovery via l_1 optimization is very satisfactory: source signals 1 and 3 are almost exactly recovered. Although there are some peaks missing, almost all the major peaks are captured in recovered source signals 2 and 4. The second example aims to extract 5 sources from 3 mixtures, which is a more under-determined than example one. The results are presented in Fig. 6.3 and Fig. 6.4, where the spanning edges of the cone are identified and l_1 norm minimization delivers a partially correct source separation. Although several spurious

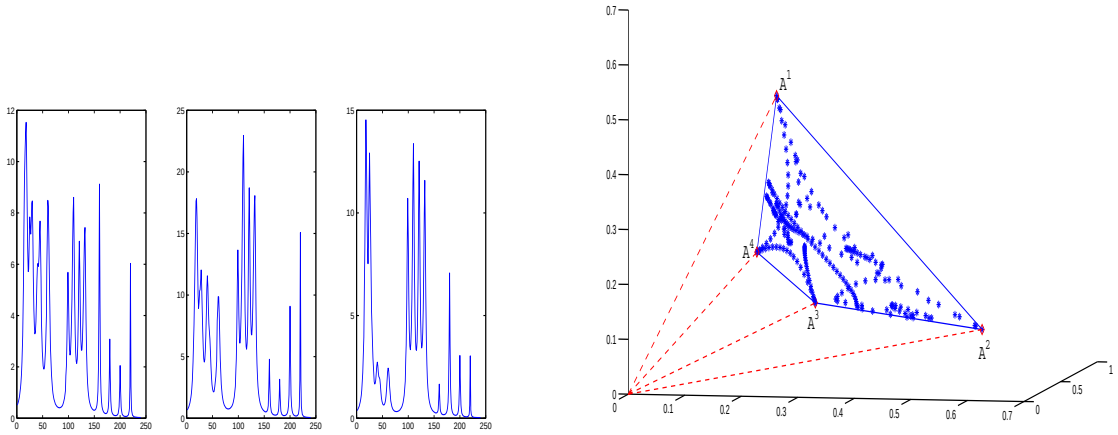

Figure 6.1: The three mixtures (left), and their geometric structure (right).

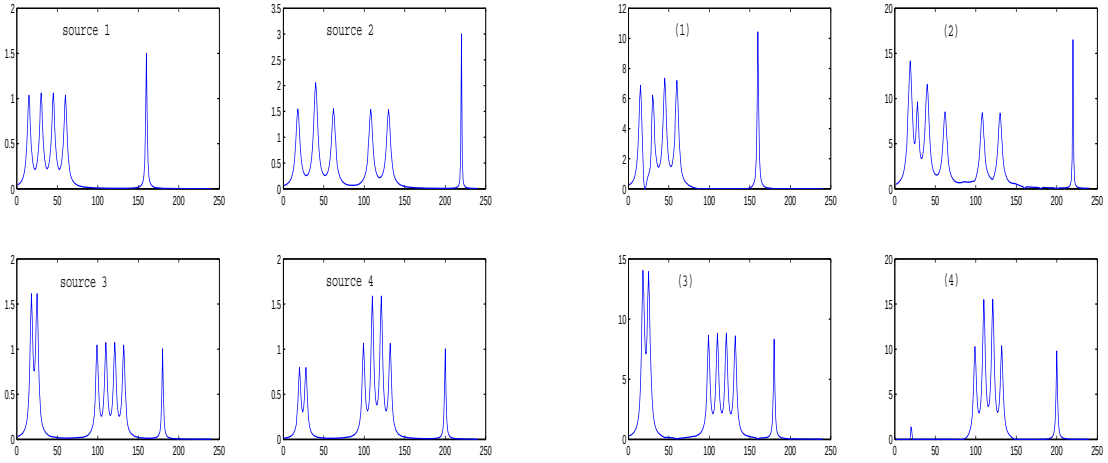

Figure 6.2: Left panel is the four real sources, and the right is their l_1 solutions.

peaks are introduced (for example, in the recovered source 1), the major characteristic peaks are captured. In the practice of NMR, such results, though imperfect, still provide valuable clues and assistance for NMR chemists to recognize chemicals from a template.

6.2 Degeneracy of Degree $r \geq 1$

If there are r degenerate columns ($r \geq 2$, the case of $r = 1$ is discussed already in detail in previous sections), then under MOC condition, one must also search for intersections of translated subspaces of dimensions $m - 2$ (lines when $m = 3$) in the interior of the cone. Consider $m = 3$ for simplicity and ease of visualization (Fig. 6.5). There are at least r intersections, each of which is associated with a positive integer (degree) equal to the number of concurrent lines passing through it. The

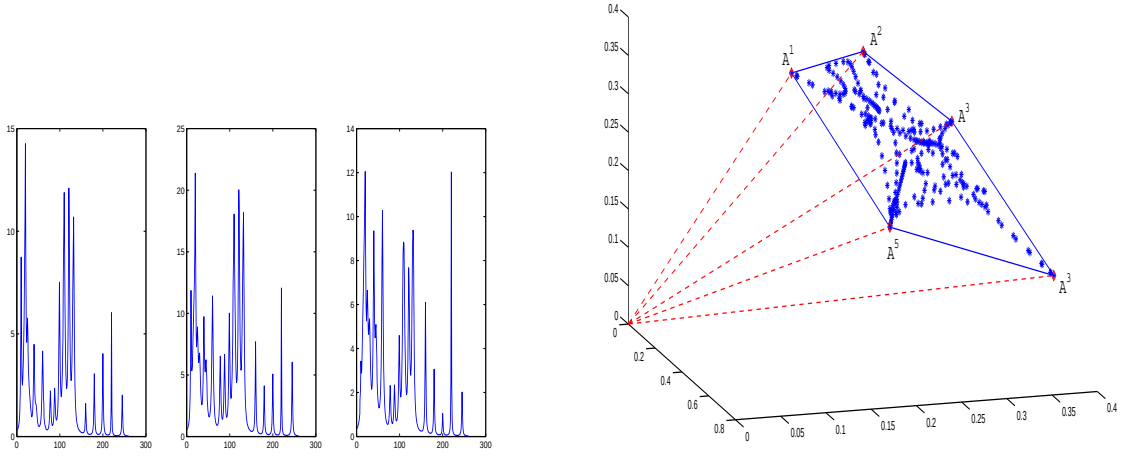

Figure 6.3: The three mixtures (left), and their geometric structure (right).

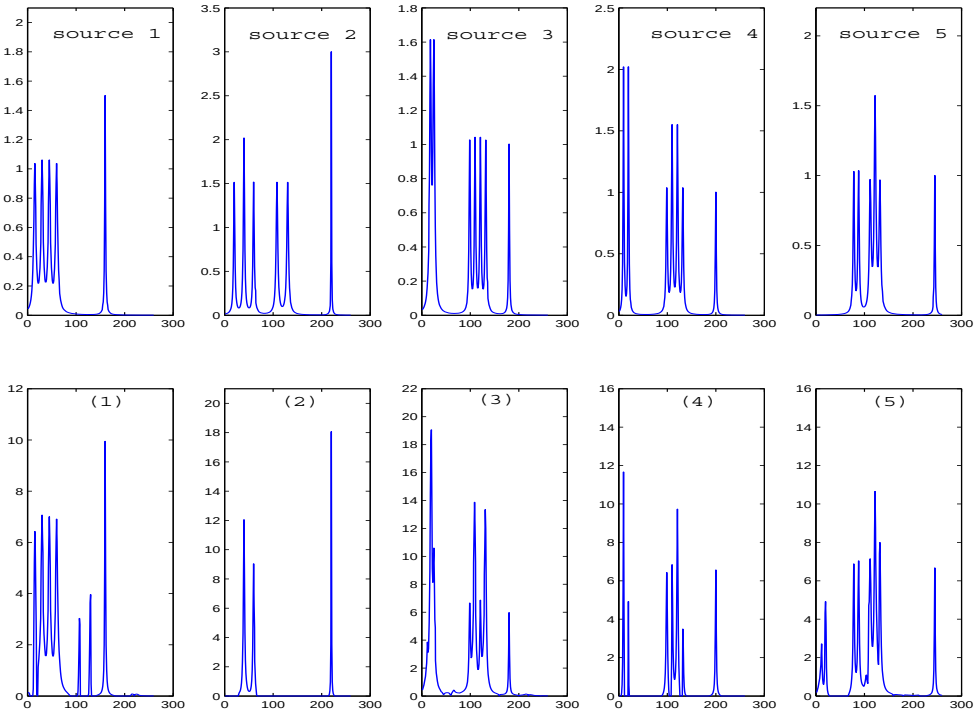

Figure 6.4: The four real sources (top), and their l_1 solutions (bottom).

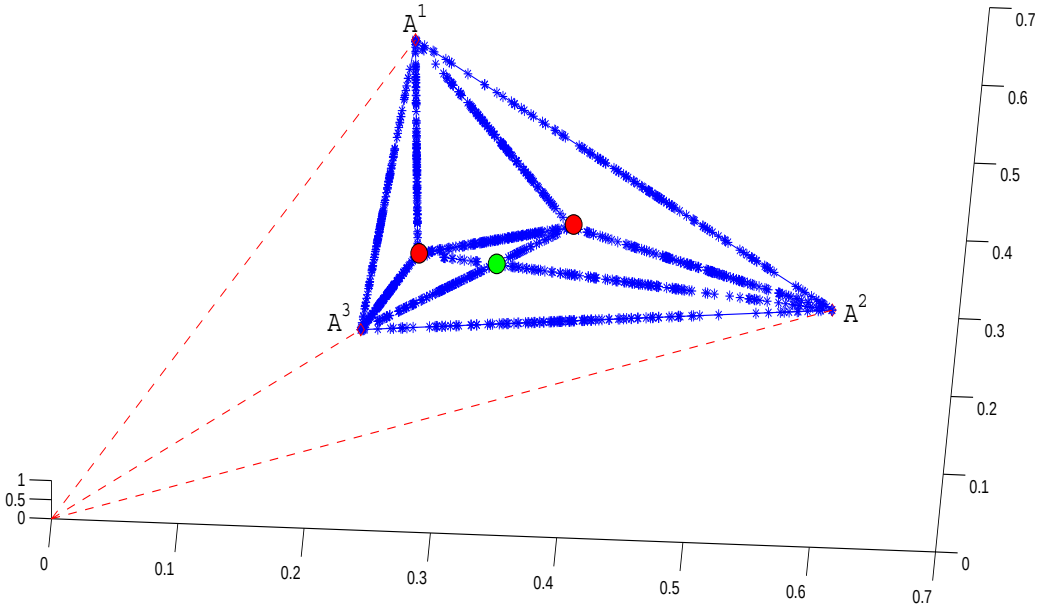

Figure 6.5: The geometry of the mixtures: among the three intersections (red and green dots), the two red ones have degree 3 and 4 respectively, while the green one has degree 2.

intersections are ordered from high to low in terms of the degree of intersection. The higher degree ones will be chosen first to fill in degenerate columns of A. If identical degree appears at different intersections, one may encounter multiple solutions. In practice, if the number of sources is unknown and is above the number of edges of the cone, we choose additional columns of the mixing matrix from the ordered list of interior intersections, and provide possibly multiple solutions for practitioners to analyze with their knowledge and experience.

Fig. 6.5 shows the geometry of the mixtures in case of $m = 3$. The spanning edges of the convex cone are identified using NN's method. Inside the cone, there are three intersections found by either Hough transform or spectral clustering. Suppose that there are two degenerate columns in the mixing matrix. The separation results are shown in Fig. 6.6, where a reasonably good recovery can be seen by comparison with the ground truth.

7 Concluding Remarks

We studied sparse blind source separation of non-negative sources when there are fewer number of mixtures than sources. Considering a geometric interpretation of the data reveals a great deal of information about unique solvability. We found necessary and sufficient conditions on the uniqueness of the uBSS problem up to scaling and permutation in the case of recovering $m + 1$ source signals from m mixture data.

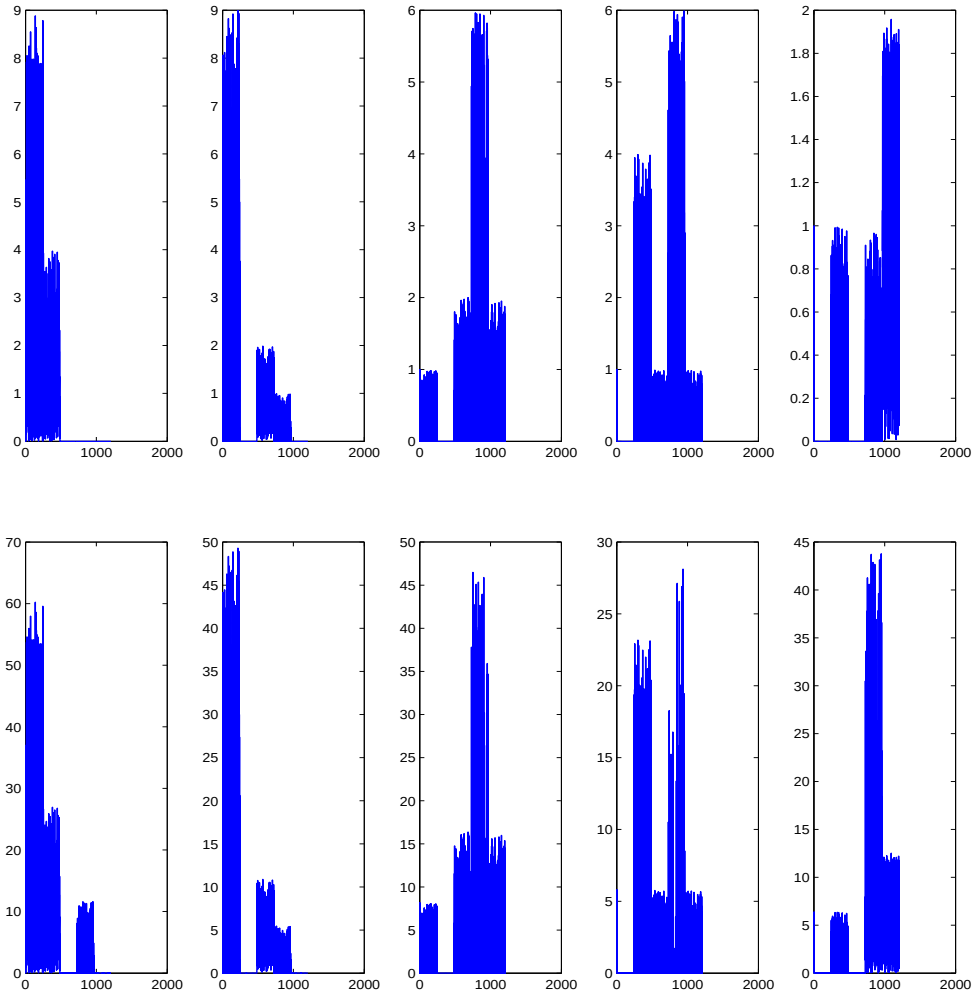

Figure 6.6: The ground truth of the five sources (top). The l_1 recovery (bottom).

Our approach exploited the geometry of data matrix and the sparsity of the source signals. Numerical results validate the solvability condition, and show satisfactory performance of the resulting uBSS. In order to deal with noisy data, an initial attempt has been made by combining clustering analysis and the geometric approach, and the idea proved to be successful.

Based on the degree of the degeneracy of the mixing matrix, we extend our method to the general case of extracting n sources from m mixtures with $m < n$, $m \geq 3$. The degenerate columns of the mixing matrix may be recovered from intersections of data hyperplanes (or translated subspaces) inside the minimal cone containing the mixture data set. The intersections may be ordered by degrees. It often requires additional knowledge to determine the actual number of degenerate columns of the mixing matrix from the mixture data. One way to go is to examine whether the recovered source signals are chemically meaningful. The geometric method developed here only provides a short list of possible sparse solutions satisfying the mixing model. In the practice of NMR, the computed short list may reveal valuable clues for a knowledged chemist to pursue further analysis. In this sense, the uBSS method is a valuable assistive computational tool.

References

[1] J. Boardman, *Automated spectral unmixing of AVRIS data using convex geometry concepts*, in Summaries of the IV Annual JPL Airborne Geoscience Workshop, JPL Pub. 93-26, Vol. 1, 1993, pp 11-14.

[2] D.H. Ballard, *Generalizing the Hough Transform to Detect Arbitrary Shapes*, Pattern Recognition 13 (1981) 111–122.

[3] M. Belkin, P. Niyogi, *Laplacian Eigenmaps for Dimensionality Reduction and Data Representation*, Neural Computation 15 (2003) pp. 1373–1396.

[4] M.W. Berry, M. Brownea, A. N. Langvilleb, V.P. Paucac, and R.J. Plemmons, *Algorithms and applications for approximate nonnegative matrix factorization*, Computational Statistics & Data Analysis 52 (2007) pp.155–173.

[5] A. Bijaoui and D. Nuzillard, *Blind source separation of multispectral astronomical images*, in Mining the Sky: Proceedings of the MPA/ESO/MPE Workshop Held at Garching, Germany, July 31–August 4, 2000, A. J. Banday, S. Zaroubi, and M. Bartelmann, eds., Springer-Verlag, Berlin, 2001, p. 571.

[6] C. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006

[7] P. Boflla and M. Zibulevsky, *Underdetermined blind source separation using sparse representations*, Signal Processing, 81 (2001) pp. 2353–2362.

[8] E. Candés, J. Romberg, and T. Tao, *Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information*. IEEE Trans. Inform. Theory, 52 (2006) pp. 489–509.

[9] S. Choi, A. Cichocki, H. Park, and S. Lee, *Blind source separation and independent component analysis: A review*, Neural Inform. Process. Lett. Rev., 6 (2005), pp. 1–57.

[10] A. Cichocki and S. Amari, *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*, John Wiley and Sons, New York, 2005.

[11] P. Comon, *Independent component analysis–a new concept?*, Signal Processing, 36 (1994) pp. 287–314.

[12] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applications*, Academic Press, 2010.

[13] D. Donoho and J. Tanner, *Sparse nonnegative solutions of underdetermined linear equations by linear programming*, Proc Natl Acad Sci USA, 102 (2005) pp. 9446-9451.

[14] R.O. Duda, P.E. Hart, *Use of the Hough Transformation to Detect Lines and Curves in Pictures*, Comm. ACM 15 (1972) pp. 11-15.

[15] J.H. Dulà and R.V. Helgason, *A new procedure for identifying the frame of the convex hull of a finite collection of points in multidimensional space*, European J. Oper. Res. 92 (1996), pp. 352–367.

[16] P. Georgiev, F. Theis, and A. Cichocki, *Sparse component analysis and blind source separation of underdetermined mixtures*, IEEE Transactions on Neural Networks, 16(4) (2005) pp. 992–996.

[17] B. Klingenberg., J. Curry and A. Dougherty, *Non-negative matrix factorization:Ill-posedness anda geometric algorithm*, Pattern Recognition 42 (2009) pp. 918–928.

[18] J. Kolba and I. Jouny, *Blind source separation in tumor detection in mammograms*, in Proceedings of the IEEE 32nd Annual Northeast Bioengineering Conference, Easton, PA, 2006, pp. 65–66.

[19] I. Koprivaa, I. Jerić, and V. Smrečki, *Extraction of multiple pure component 1H and 13C NMR spectra from two mixtures: Novel solution obtained by sparse component analysis-based blind decomposition*, Analytica Chimica Acta, 653 (2009) pp. 143-153.

[20] D. D. Lee and H. S. Seung, *Learning of the parts of objects by non-negative matrix factorization*, Nature, 401 (1999) pp. 788791.

[21] J. Liu, J. Xin, Y-Y Qi, *A Dynamic Algorithm for Blind Separation of Convolutive Sound Mixtures*, Neurocomputing, 72(2008), pp 521-532.

[22] J. Liu, J. Xin, Y-Y Qi, *A Soft-Constrained Dynamic Iterative Method of Blind Source Separation*, SIAM J. Multiscale Modeling Simulations, Vol. 7, No. 4, pp 1795-1810, 2009.

[23] J. Liu, J. Xin, Y-Y Qi, F-G Zeng, *A Time Domain Algorithm for Blind Separation of Convolutive Sound Mixtures and L-1 Constrained Minimization of Cross Correlations*, Comm. Math Sci, Vol. 7, No. 1, 2009, pp 109-128.

[24] W. Naanaa and J.–M. Nuzillard, *Blind source separation of positive and partially correlated data*, Signal Processing 85 (9) (2005), pp. 1711–1722.

[25] M. Naceur, M. Loghmari, and M. Boussema, *The contribution of the sources separation method in the decomposition of mixed pixels*, IEEE Trans. Geosci. Remote Sensing, 42 (2004), pp. 2642–2653.

[26] A.Y. Ng, M.I. Jordan, and Y. Weiss,*On spectral clustering: Analysis and an algorithm*, Advances in Neural Information Processing Systems 14 MIT Press (2001), pp 849–856.

[27] M. Plumbley, *Conditions for non-negative independent component analysis*, IEEE Signal Processing Letters, 9 (2002) pp. 177–180.

[28] M. Plumbley, *Algorithms for nonnegative independent component analysis*, IEEE Transactions on Neural Networks, 4(3) (2003) pp. 534–543.

[29] L.G. Shapiro, G.C. Stockman,*Computer Vision*, Prentice-Hall Inc., 2001.

[30] R.M. Silverstein, F.X. Webster, and D.J. Kiemle, *Spectrometric Identification of Organic Compounds*, John Wiley and Sons, 2005.

[31] K. Stadlthanner, A. Tom, F. Theis, W. Gronwald ,H.-R. Kalbitzer, and E. Lang, *On the use of independent analysis to remove water artifacts of 2D NMR Protein Spectra*, In Proc. Bioeng'2003 (2003).

[32] Y. Sun, C. Ridge, F. del Rio, A.J. Shaka and J. Xin, *Postprocessing and Sparse Blind Source Separation of Positive and Partially Overlapped Data*, submitted.

[33] F.J. Theis, C.G. Puntonet and E. W. Lang, *A histogram-based overcomplete ICA algorithm*, in 4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA 2003), April 2003, Nara, Japan.

[34] Ö. Yilmaz and S. Rickard, *Blind separation of speech mixtures via time-frequency masking*, IEEE Trans. Signal Processing, 52 (2004) pp. 1830–1847.

[35] W. Yin, S. Osher, D. Goldfarb, J. Darbon, *Bregman iterative algorithm for l_1-minimization with applications to compressive sensing*, SIAM J. Image Sci, 1(143), pp 143-168, 2008.

[36] Y. Zhang, *Theory of compressive sensing via L_1-Minimization: A Non-RIP analysis and extensions*, Technical report, 2009, Rice University.

[37] M.E. Winter, *N-findr: an algorithm for fast autonomous spectral endmember determination in hyperspectral data*, in Proc. of the SPIE, vol. 3753, 1999, pp 266-275.