SIAM J. IMAGING SCIENCES
Vol. 5, No. 3, pp. 886–911

Nonnegative Sparse Blind Source Separation for NMR Spectroscopy by Data Clustering, Model Reduction, and ℓ_1 Minimization*

Yuanchang Sun† and Jack Xin†

Abstract. Motivated by applications in nuclear magnetic resonance (NMR) spectroscopy, we introduce a novel blind source separation (BSS) approach to treat nonnegative and correlated data. We consider the (over)-determined case where n sources are to be separated from m linear mixtures ($m \geq n$). Among the n source signals, there are $n-1$ partially overlapping (Po) sources and one positive everywhere (Pe) source. This condition is applicable for many real-world signals such as NMR spectra of urine and blood serum for metabolic fingerprinting and disease diagnosis. The geometric properties of the mixture matrix and the sparseness structure of the source signals (in a transformed domain) are crucial to the identification of the mixing matrix and the sources. The method first identifies the mixing coefficients of the Pe source by exploiting geometry in data clustering. Then subsequent elimination of variables leads to a sub-BSS problem of the Po sources solvable by the minimal cone method and related linear programming. The last step is based on solving a convex ℓ_1 minimization problem to extract the Pe source signals. Numerical results on NMR spectra show satisfactory performance of the method.

Key words. nonnegative sources, blind separation, sparseness, ℓ_1 minimization, clustering

AMS subject classifications. 94A12, 65H10, 65K10, 90C05

DOI. 10.1137/110827223

1. Introduction.

The goal of this paper is to study nonnegative blind source separation (BSS) problems arising from nuclear magnetic resonance (NMR) spectroscopy. The motivation is the separation of NMR spectra of biofluids such as urine and blood for metabolic fingerprinting and disease diagnosis. BSS methods aim to recover source signals from a set of mixture signals without knowing the mixing process. BSS has received considerable attention in processing of speech, image, and biomedical signals [2, 6, 8, 9, 10, 11, 16, 19, 20, 23, 24, 29, 31, 32, 40, 45]. Recently, there have been considerable activities [3, 5, 15, 21, 22, 27, 28, 32, 33, 34, 36, 37, 38, 42] involving nonnegative BSS in computer tomography, biomedical image processing, remote sensing, analytical chemistry, and other areas. The nonnegative BSS problem is defined by the matrix model

$$X = A\,S, \quad \text{with} \quad A_{ij} \geq 0, \quad S_{ij} \geq 0, \tag{1.1}$$

where A_{ij}, S_{ij} are matrix entries of A and S, respectively. We shall use similar notation for other matrices. $X \in \mathbb{R}^{m\times p}$ is the mixture matrix containing known mixture signals as

*Received by the editors March 11, 2011; accepted for publication (in revised form) April 2, 2012; published electronically July 26, 2012. This work was partially supported by NSF-ADT grant DMS-0911277 and NSF grant DMS-0712881.
http://www.siam.org/journals/siims/5-3/82722.html
†Department of Mathematics, University of California at Irvine, Irvine, CA 92697 (sunyuanc@gmail.com, jxin@math.uci.edu).

its rows, $S \in \mathbb{R}^{n\times p}$ is the unknown source matrix, and $A \in \mathbb{R}^{m\times n}$ is the unknown mixing matrix. All the matrices are nonnegative. The dimensions of the matrices are expressed in terms of three numbers: (1) p, the number of available samples; (2) m, the number of mixture signals; and (3) n, 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 nonnegative A and S from X. The problem is also known as nonnegative matrix factorization (NMF) [22]. Similar to factorizing a composite number ($24 = 3*8 = 8*3 = 4*6 = 6*4$), there are permutation and scaling ambiguities in solutions to BSS. For any permutation matrix P and invertible diagonal matrix Λ, $(AP\Lambda,\ \Lambda^{-1}P^{-1}S)$ is another pair equivalent to the solution (A, S), since

$$X = AS = (AP\Lambda)(\Lambda^{-1}P^{-1}S). \tag{1.2}$$

Various BSS methods have been proposed based on a priori knowledge of source signals such as spatiotemporal decorrelation, statistical independence, sparseness, and nonnegativity [4, 8, 9, 12, 18, 19, 22, 23, 24, 25, 28, 36, 37]. Sparseness BSS has been receiving considerable attention recently. Among many others, Hoyer [18] studied nonnegative matrix factorization with sparseness constraint; Takigawa, Kudo, and Toyama [39] used the ℓ_1 norm to study underdetermined BSS which extracts more sources than the number of mixtures; sparsity in a transformed domain has been widely explored by Bobin et al. [4] in the problem of BSS. Despite some success in signal/image processing, their methods solve nonconvex optimizations and are potentially unreliable in some real-world applications such as NMR spectroscopy. Independent component analysis (ICA) and its derivatives [10, 11] recover statistically independent source signals and mixing matrix A. However, the source independence fails to be a good separation criterion for NMR spectra, because the data are correlated when molecules responsible for each source share common structural features. Besides, the properly phased absorption-mode NMR spectral signals from a single-pulse experiment are positive. Separation of such positive and correlated data (NMR spectra) is a fundamental problem in analytical chemistry. Analysis of NMR signals is playing a significant role in the identification of complex chemical compounds and in studying structures of large proteins [35]. ICA-based methods would not work for this class of data. A better working assumption for the data is the partial sparseness condition proposed by Naanaa and Nuzillard [28]. The source signals are required to be nonoverlapping at some locations of acquisition variable (e.g., frequency). Such a local sparseness condition leads to a dramatic mathematical simplification of a general nonnegative matrix factorization problem (1.1) which is nonconvex. Geometrically speaking, the problem of finding the mixing matrix A reduces to the identification of a minimal cone containing the columns of mixture matrix X. The latter can be done by linear programming. In fact, Naanaa and Nuzillard's sparseness assumption and the geometric construction of columns of A were known in the 1990s [3, 42] in the study of blind hyperspectral unmixing of remote sensing, where the same mathematical model (1.1) is used. The analogue of Naanaa and Nuzillard's assumption is called the pixel purity assumption (PPA) [7]. The resulting geometric (cone) method is the so-called N-findr [42], and it is now a benchmark in hyperspectral unmixing.

The minimal cone methods (Naanaa and Nuzillard's method, N-findr method) are efficient in separating the sources for determined and overdetermined BSS problems where their conditions are met. They are also robust in that major peaks can still be recovered when their

condition is violated to a certain extent. A recent study by the authors of the current paper investigated how to postprocess with the abundance of mixture data, and how to improve mixing matrix estimation with major peak–based corrections [36]. Here we are interested in seeking a new method to deal with data which are intractable to the cone method and postprocessing.

The present work originates from analytical chemistry and, more precisely, from NMR spectroscopy of biofluids such as urine and blood. The change of metabolites in biofluids is directly related to the biochemical changes in a living system. Identifying and quantifying metabolites in these two biofluids has been an important research topic in biological sciences. The NMR-based metabonomics have been proved to be a powerful technology for measuring the metabolic responses of living systems to internal or external stimuli (see [1, 43, 44] and the references therein). The NMR spectra of urine samples are normally assigned using a catalogue of two-dimensional (2D) NMR methods. However, the complexity of urinary composition makes the complete assignments of the urinary spectra difficult with spiking alone or even with the combination of 2D NMR and spiking methods. For example, some 2D NMR methods require the so-called metabolite standards (reference spectra), which are not necessarily available all the time. In particular, the references for unknown metabolites, by default, are not available at all. Consequently, as of now, only about one-third of detectable urinary metabolites have been assigned unambiguously [44]. A similar situation exists in the NMR spectroscopy of blood serum. Our method can be useful for separating out unknown sources from the residuals after any known reference spectra have been first deployed to fit the data. Our hope is to offer a helpful tool for identifying and quantifying the remaining unknown sources.

The mathematical challenges of this problem lie in two aspects. First, the ideal stand-alone peak (SAP) [28] is no longer satisfied. The complicated NMR spectra contain both wide-peak source signals and narrow-peak source signals. In Figure 1, blood serum has constituent with very wide peaks extending over the acquisition regime; it is also observed that the narrow-peak sources nowhere dominate the wide-peak signal. As a result, the convex cone method and its postprocessing fail to be good separation approaches for this type of data. Second, a new and effective algorithm needs to be developed that is more than the minimal cone construction in [28]. The present work addresses these issues and makes an initial attempt for the blind separation of sources in serum and urine types of NMR data.

The paper is organized as follows. In section 2, we review the essentials of the convex cone method and its source sparseness assumption, and then propose a new source condition motivated by NMR spectra of urine and blood. In section 3, we introduce our BSS method consisting of a combination of data clustering, model reduction via elimination of variables, and solving a convex ℓ_1 minimization problem. The first two steps reduce a general nonconvex NMF problem to a convex problem. Numerical examples are shown in section 4, and concluding remarks are in section 5.

The notation A^j stands for the jth column of matrix A, S^j for the jth column of matrix S, and X^j for the jth column of matrix X, while S_j and X_j are the jth rows of matrix S and X, or the jth source and mixture, respectively.

Remark. For the purpose of illustration, we shall use the simulated data (Figures 2–16) for which we did not specify units. The unit of the x-axis can be understood as number of

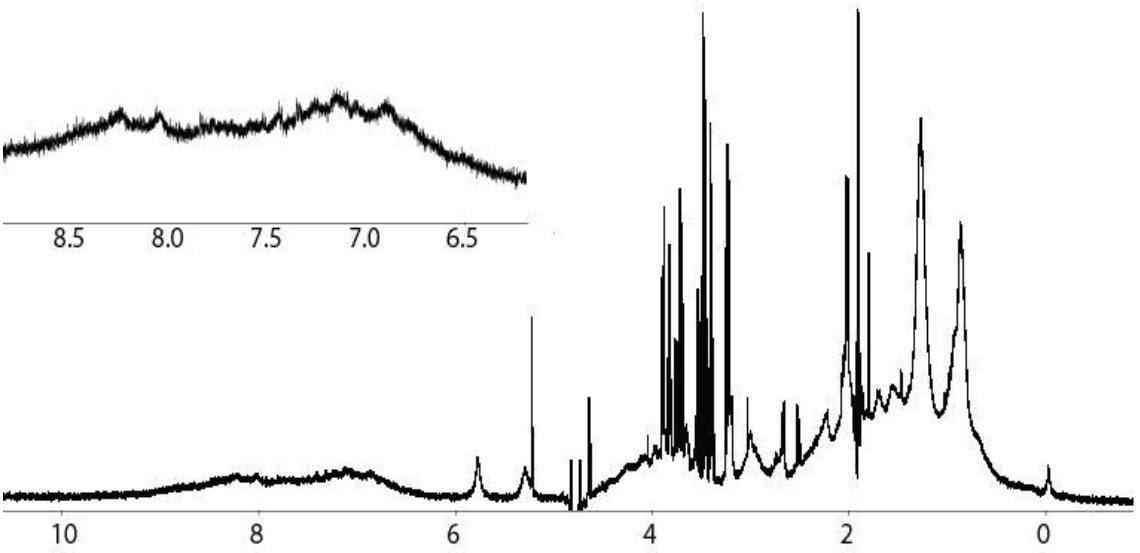

Figure 1. *NMR spectrum of a human serum sample from* [41]*, where the unit of the x-axis is ppm. This plot shows that the blood serum has a constituent with very wide spectral peaks which overlap others everywhere. As a result, the mixing matrix A cannot be recovered from data matrix X independently of S, and so A and S are much more coupled.*

pixels. A real-world NMR spectrum usually has ppm (parts per million) as its unit (see the data in Figure 1, Figures 17–21). The multidimensional figures are the scatter plots of the columns of the mixture matrix.

2. The source assumption and the method.

2.1. Convex cone method. Naanaa and Nuzillard [28] proposed an efficient sparse BSS method and its mathematical analysis for nonnegative and partially overlapped signals. Consider the (over)-determined cases of model (1.1), where $m \geq n$ and the mixing matrix A is full rank. Simply speaking, the key sparseness assumption on source signals is that each source has an SAP at some location of acquisition variable where the other sources are identically zero. We shall call this assumption the SAP condition. More precisely, the source matrix is assumed to satisfy the following condition.

Assumption [SAP]. For each $i \in \{1, 2, \ldots, n\}$ there exists a $j_i \in \{1, 2, \ldots, p\}$ such that $S_{ij_i} > 0$ and $S_{kj_i} = 0$ $(k = 1, \ldots, i-1, i+1, \ldots, n)$.

If (1.1) is written in terms of columns as

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

then the SAP implies that $X^{j_i} = S_{ij_i} A^i$, $i = 1, \ldots, n$, or $A^i = \frac{1}{S_{ij_i}} X^{j_i}$. Hence (2.1) can be rewritten as

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

which means that every column of X is a nonnegative 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 are found, an estimation of the mixing matrix is obtained. The identification of $\hat{A}$'s columns is equivalent to identifying a minimal cone of a finite collection of vectors. The cone encloses the data columns in matrix X and is the smallest of such cones. The edges of the minimal cone are columns of $\hat{A}$. To find the edges, the following optimization is suggested:

$$c = \min \sum_{j=1, j\neq k}^{p} \lambda_j$$
$$\text{subject to (s.t.)} \quad \sum_{j=1, j\neq k}^{p} X^j \lambda_j = X^k, \lambda_j \geq 0.$$

Reference [13] shows that X^k *is an edge of the convex cone if and only if the optimal objective function value of* c^* *is greater than* 1.

If the data are contaminated by noise, the following optimization problems is suggested to estimate the mixing matrix:

$$\min \text{ score} = \left\| \sum_{j=1, j\neq k}^{p} X^j \lambda_j - X^k \right\|_2, \quad \text{s.t. } \lambda_j \geq 0. \tag{2.3}$$

A score is associated with each column of X. A column with a low score is unlikely to be a column of $\hat{A}$ because this column is roughly a nonnegative 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 nonnegative linear combination of other columns. The n rows from X with highest scores are selected to form $\hat{A}$, the mixing matrix.

Once the mixing matrix is extracted from the data, the Moore–Penrose inverse $\hat{A}^+$ of $\hat{A}$ is then calculated, and an estimate of S is achieved: $\hat{S} = \hat{A}^+ X$. However, the recovered S contains negative values due to the error in the estimation of A. Realizing that the columns of S possess sparsity due to the SAP condition, we solve a nonnegative ℓ_1 optimization problem for each column S^i of S:

$$\min_{S^i \geq 0} \mu \|S^i\|_1 + \frac{1}{2} \|X^i - \hat{A} S^i\|_2^2 . \tag{2.4}$$

One can also solve the constrained problem; however, the unconstrained problem is more reasonable because the data in general contain noise. When there is minimal measurement error, one assigns a tiny value to μ to heavily weigh the fidelity term $\|X^i - \hat{A} S^i\|_2^2$ in order for $\hat{A} S^i = X^i$ to be nearly satisfied. The solution can be obtained by the Bregman iterative method [17, 46] with a proper projection onto a nonnegative convex set.

2.2. PePoDi source condition. The convex cone method is very efficient in separating the SAP sources for determined and overdetermined BSS problems. However, new methods have

to be found for separating the source signals violating SAP. Among the non-SAP data, we are particularly concerned with the urine (serum)-type NMR data. An examination of NMR structures of the constituents suggests a better working condition: among the n sources, $n-1$ sources are assumed to satisfy SAP, or they are allowed to have partial overlapping with each other. These $n-1$ sources are said to satisfy the Po condition, which essentially is SAP. The last source has positive response everywhere over the whole acquisition region. This source is called a Pe source. In addition, the Pe source is required to have dominant interval(s) (Di) over other Po sources. We shall call this source condition the PePoDi assumption. To simplify the notation, we assume that $S_1, \ldots, S_{n-1}$ are the Po sources, while S_n is the Pe source. Note that these assignments are not known and have to be computed. Mathematically speaking, the source matrix S is assumed to satisfy the following condition.

Assumption [PePoDi]. For each $i \in \{1, 2, \ldots, n-1\}$ there exists a $j_i \in \{1, 2, \ldots, p\}$ such that $S_{ij_i} \gg S_{kj_i}$ $(k = 1, \ldots, i-1, i+1, \ldots, n)$. Moreover, $S_{nj} > 0$ for $j \in \{1, 2, \ldots, p\}$, and there is a set $\mathcal{I} \subset \{1, 2, \ldots, p\}$ such that $S_{nk} \gg S_{ik}$ for $k \in \mathcal{I}$ and $i \in \{1, 2, \ldots, n-1\}$.

Simply stated, sources $S_1, \ldots, S_{n-1}$ satisfy a relaxed SAP condition which is more appropriate due to the noise in the real-world data. The postprocessing techniques developed in [36] in fact deal with this type of data. Source S_n is the Pe source which has dominant intervals over other sources. We further assume that there are no locations where $S_1, \ldots, S_{n-1}$ dominate S_n motivated by the NMR spectra of blood serum. Figure 2 is an example of such a source matrix S which contains three source signals. In the next section, we develop a BSS method to separate Po and Pe sources.

3. Proposed method.

In this section, we present a novel three-step BSS method. In the following, we shall assume that $m = n$ (determined case); however, the results can be easily extended to the case of $m > n$ (overdetermined case). Let us rewrite (1.1) as follows:

$$X^j = \sum_{k=1}^{n} S_{kj} A^k, \qquad j = 1, \ldots, p\,. \tag{3.1}$$

For the purpose of illustration, we shall continue to assume that the first $n-1$ rows of S are the Po sources, and the last row S_n corresponds to the Pe source. However, these assignments are not known and have to be identified in the computation. If SAP holds, the column vectors of the mixing matrix A span a minimal cone which contains all the column vectors of X. SAP implies that the spanning edges of the cone lie in the column vectors of X, and their identification can be achieved by linear programming. From a geometric perspective, if SAP is satisfied and A is full rank, the column vectors of A span the minimal cone containing all the columns of X. However, if the mixture matrix X comes from mixing sources that violate SAP, none of the columns of X are parallel to those of A. The estimated mixing matrix whose columns are extracted from columns of X according to the convex cone method will then deviate from the true A, causing errors in the recovered source signals. The second plot in Figure 3 shows two mixtures of two sources satisfying the PePoDi condition. The true cone, spanned by the two red diamond dots, is wider than the set of column vectors of X (the collection of stars). The convex cone method, however, identifies the smallest cone containing the mixtures (spanned by the two blue circles). As illustrated in Figure 3, the convex cone method fails to retrieve the mixing matrix A for the PePoDi data. However, the

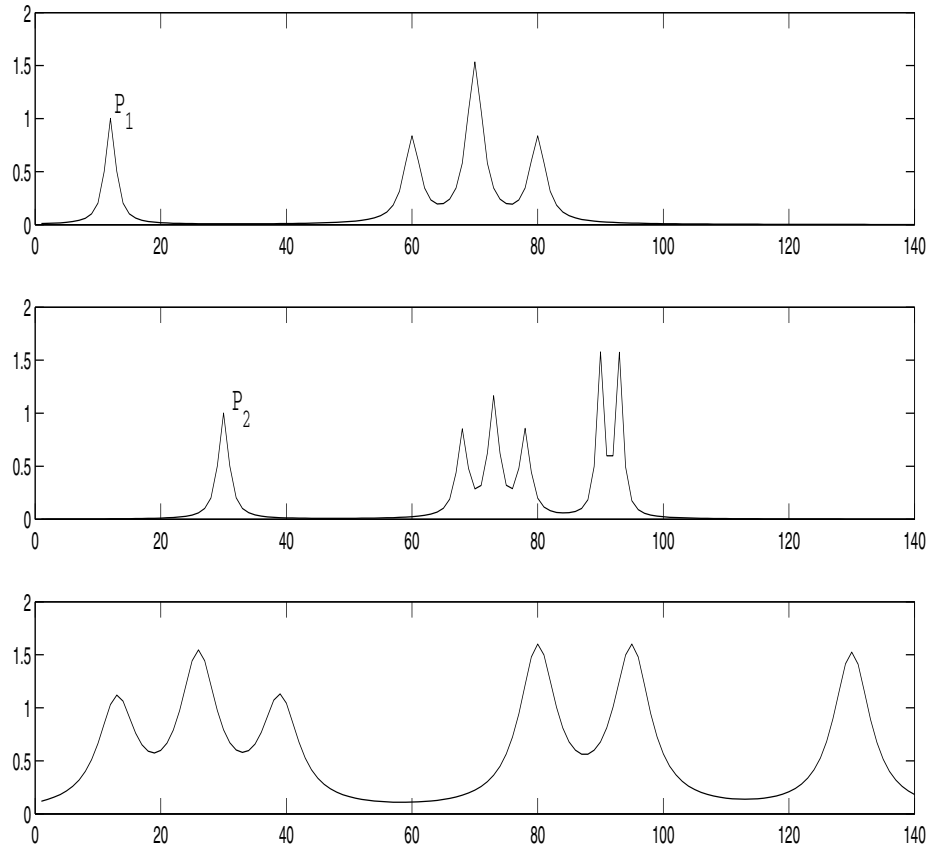

Figure 2. *Sources* 1 *and* 2 *are two Po sources which satisfy SAP, and* P_1, P_2 *are two SAPs. Source* 3 *is a Pe source which is Pe. Note also that source* 3 *has a dominant interval* $[120, 140]$.

dominant interval(s) (Di) in the last (nth) source implies that A^n could be estimated by a clustering technique combined with linear programming. We shall explain the idea in detail next.

3.1. Identification of A^n. In fact, the dominant interval(s) in source S_n imply that there are columns of X such that

$$X^k = S_{nk}A^n + \sum_{i=1}^{n-1} O_{ik}A^i \,, \tag{3.2}$$

where S_{nk} dominate O_{ik} $(i = 1, \ldots, n-1)$, i.e., $S_{nk} \gg O_{ik}$. The identification of A^n is equivalent to finding these columns X^k that form a cluster due to the Di condition. As seen in Figure 2, mixtures in the sampling region [120, 140] satisfy (3.2); hence a cluster is formed in Figure 4. Next we shall discuss how to locate the cluster (thus A^n). All of X's columns form a set of points $\mathcal{S} = \{X^1, X^2, \ldots, X^p\}$ in n-dimensional space. The convex hull of $\mathcal{S}$ is a polytope, $\mathcal{A} \in \mathbb{R}^n$. The frame $\mathcal{F}$ of these points is the set of extreme points of the

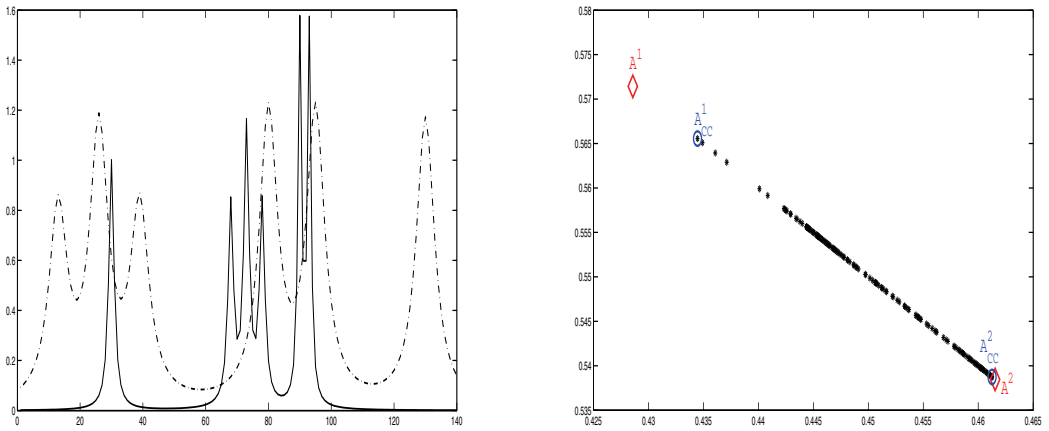

Figure 3. *A Pe source (dash-dot line) and a Po source (solid line) are shown in the left panel. Notice that the Pe source has a dominant interval* $[100, 140]$. *Right: Comparison of the two columns of the mixing matrix (blue circles) recovered from the convex cone method with those of the true mixing matrix A (red diamonds). The cone method identifies the columns of the mixing matrix as the edges of a minimal cone enclosing the mixtures. The deviation of the convex cone's results is due to the violation of the SAP condition.*

convex hull. To determine whether the element X^k of the set $\mathcal{S}$ constitutes an element of $\mathcal{F}$, the constrained equations (2.3) can be used. X^k belongs to $\mathcal{F}$ if it cannot be written as a linear combination of other points of $\mathcal{S}$; i.e., (2.3) is inconsistent. Finding the convex hull and its frame for a finite collection of points is a basic problem in computational geometry; the standard approach consists of solving linear programs [13]. Apparently, $\mathcal{F}$ contains a vector which is approximately parallel to A^n; this can be seen from (3.2). Let us denote it as $\hat{A}^n$. Among the elements of $\mathcal{F}$, $\hat{A}^n$ is the one attracting a cluster. To identify $\hat{A}^n$, an ϵ-ball is created with the center at each element of $\mathcal{F}$. The center of the ball that contains most of the data points is recognized as $\hat{A}^n$. To summarize, the following algorithm is suggested for identifying $\hat{A}^n$:

1. Select a submatrix $\hat{X}$ of X consisting of all the mutually noncollinear columns of X.
2. Find the frame $\mathcal{F}$ of the set of points $\hat{X}$.
3. Identify A^n from $\mathcal{F}$ by the geometry of the mixture X (clustering).

For example, the plot in Figure 4 is an illustration of the algorithm. In this example, we used the spectra given in Figure 2 as source signals. The mixtures are constructed by simulating the mixing process described by model (1.1). The mixing matrix

$$A = \begin{pmatrix} 6 & 2 & 3 \\ 3 & 5 & 4 \\ 2 & 3 & 3.5 \end{pmatrix},$$

and the recovered $\hat{A}^3$ is $[0.2866, 0.3809, 0.3326]^{\mathrm{t}}$, which is approximately parallel to $[3, 4, 3.5]^{\mathrm{t}}$ (it can be normalized to $[0.2858, 0.3810, 0.3333]^{\mathrm{t}}$); hence it is taken as the estimate of A^3. Once $\hat{A}^n$ is obtained, we eliminate the last source S_n (Pe source) from X. The resulting mixtures contain only Po sources, and the reduced BSS problem could be solved by the convex cone method.

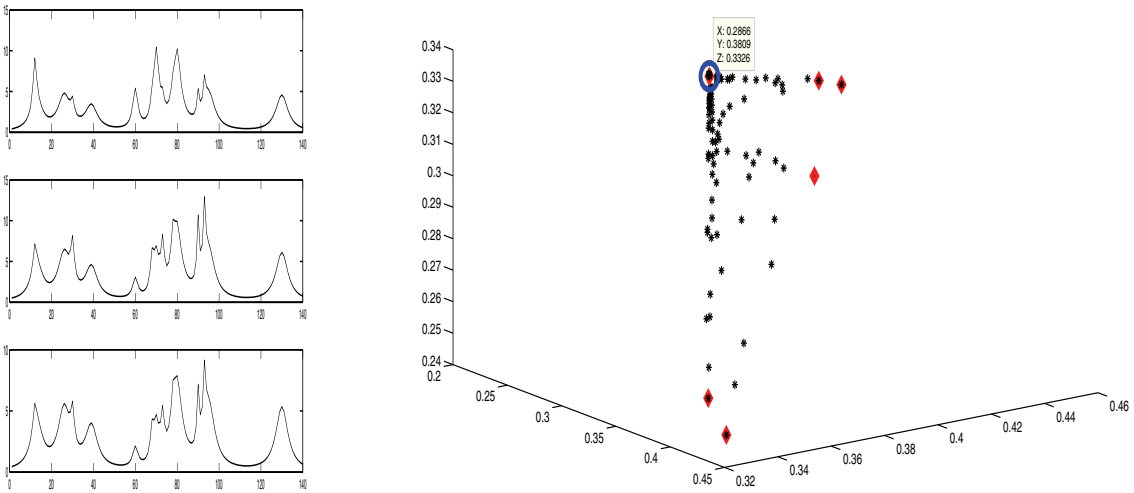

Figure 4. *Left: The three mixtures from the sources in Figure 2. Right: The scatter plot of columns of the mixtures X; a cluster can be seen (indicated by blue circle). The red diamonds are the convex hull frame of X's columns.*

3.2. Recovery of Po sources. Denoting the n mixture data as row vectors $X_1, \ldots, X_n$, we have

(3.3)

$$X_1 = \left[\sum_{j=1}^{n-1} A_{1j}S_{j1} + A_{1n}S_{n1},\ \sum_{j=1}^{n-1} A_{1j}S_{j2} + A_{1n}S_{n2}, \ldots,\ \sum_{j=1}^{n-1} A_{1j}S_{jp} + A_{1n}S_{np}\right],$$

$$X_2 = \left[\sum_{j=1}^{n-1} A_{2j}S_{j1} + A_{2n}S_{n1},\ \sum_{j=1}^{n-1} A_{2j}S_{j2} + A_{2n}S_{n2}, \ldots,\ \sum_{j=1}^{n-1} A_{2j}S_{jp} + A_{2n}S_{np}\right]$$

$$\vdots$$

$$X_n = \left[\sum_{j=1}^{n-1} A_{nj}S_{j1} + A_{nn}S_{n1},\ \sum_{j=1}^{n-1} A_{nj}S_{j2} + A_{nn}S_{n2}, \ldots,\ \sum_{j=1}^{n-1} A_{nj}S_{jp} + A_{nn}S_{np}\right],$$

where $A^n = [A_{1n}, A_{2n}, \ldots, A_{nn}]^{\mathrm{T}}$ are obtained from the last step (actually a parallel vector $\hat{A}^n$ is extracted, and it is considered equivalent to A^n). Using the last equation of (3.3) we transform $X_i \to X_i - \frac{A_{in}}{A_{nn}}X_n$, $i = 1, 2, \ldots, n-1$. Then a new mixture matrix is formed as

$$\hat{X} = \begin{pmatrix} X_1 - \frac{A_{1n}}{A_{nn}}X_n \\ X_2 - \frac{A_{2n}}{A_{nn}}X_n \\ \vdots \\ X_{n-1} - \frac{A_{n-1\,n}}{A_{nn}}X_n \end{pmatrix}, \tag{3.4}$$

which depends only on the first $(n-1)$ Po sources satisfying the SAP condition. Note that the matrix $\hat{X}$ may contain negative entries, but the convex cone method is still able to identify the minimal cone which contains the columns of $\hat{X}$; the cone, however, may not lie in the sector consisting of nonnegative vectors. In this step, we separate Po sources from the reduced mixtures.

At this point, let us summarize the known and unknown variables by writing the model (1.1) in terms of components:

$$\text{(3.5)}\quad \begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1p} \\ X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{np} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} \cdot \begin{pmatrix} S_{11} & S_{12} & \cdots & S_{1p} \\ S_{21} & S_{22} & \cdots & S_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ S_{n1} & S_{n2} & \cdots & S_{np} \end{pmatrix},$$

where the known components are denoted in black, while the unknown variables are in red. To solve the unknown A_{ij} (there are $n(n-1)$ in total), we select the first n columns of X to set up the following n^2 equations:

$$\text{(3.6)}\quad \begin{cases} X_{11} = A_{11}S_{11} + A_{12}S_{21} + \cdots + A_{1n}S_{n1} \\ \cdots \\ X_{1n} = A_{11}S_{1n} + A_{12}S_{2n} + \cdots + A_{1n}S_{nn}, \\ X_{21} = A_{21}S_{11} + A_{22}S_{21} + \cdots + A_{2n}S_{n1} \\ \cdots \\ X_{2n} = A_{21}S_{1n} + A_{22}S_{2n} + \cdots + A_{2n}S_{nn} \\ \vdots \\ X_{n1} = A_{n1}S_{11} + A_{n2}S_{21} + \cdots + A_{nn}S_{nn} \\ \cdots \\ X_{nn} = A_{n1}S_{1n} + A_{n2}S_{2n} + \cdots + A_{nn}S_{nn}, \end{cases}$$

which can be written in matrix form $M\ y = b$ with

$$M = \begin{pmatrix} V & O & O & \cdots & A_{1n}I_n \\ O & V & O & \cdots & A_{2n}I_n \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ O & O & \cdots & V & A_{nn}I_n \end{pmatrix} \in \mathbb{R}^{n^2 \times n^2},$$

$$V = \begin{pmatrix} S_{11} & S_{21} & \cdots & S_{n-11} \\ S_{12} & S_{22} & \cdots & S_{n-12} \\ \vdots & \vdots & \vdots & \vdots \\ S_{1n} & S_{2n} & \cdots & S_{n-1n} \end{pmatrix} \in \mathbb{R}^{n \times (n-1)},$$

where O is a zero matrix in $\mathbb{R}^{n\times(n-1)}$,

$$y = (A_{11}, \ldots, A_{1n-1}, A_{21}, \ldots, A_{2n-1}, \ldots, A_{n1}, \ldots, A_{nn-1}, S_{n1}, \ldots, S_{nn})^{\mathrm{T}},$$

$$b = (X_{11}, \ldots, X_{1n}, X_{21}, \ldots, X_{2n}, \ldots, X_{n1}, \ldots, X_{nn})^{\mathrm{T}}.$$

In general, the matrix M is noninvertible. In fact, the following result holds for M.

Theorem. *For $n \geq 2$, M is singular, and its kernel has dimension $n-1$.*

Proof. It can be shown that column reductions transform M into the following form:

$$\begin{pmatrix} O & O & O & \cdots & A_{1n}I_n \\ O & V & O & \cdots & A_{2n}I_n \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ O & O & \cdots & V & A_{nn}I_n \end{pmatrix}. \tag{3.7}$$

In fact, performing column addition operations by $A_{1n}I_n$ to eliminate V in the top left corner, we will obtain

$$\begin{pmatrix} O & O & O & \cdots & A_{1n}I_n \\ \alpha_2 V & V & O & \cdots & A_{2n}I_n \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \alpha_n V & O & \cdots & V & A_{nn}I_n \end{pmatrix},$$

where $\alpha_j = -\frac{A_{jn}}{A_{1n}} \neq 0$, $j = 2, \ldots, n$. Next, the submatrix $\alpha_2 V$ can be eliminated by V, and similar operations can be performed until $\alpha_n V$ is canceled by V. Therefore, W is transformed into the form (3.7), which has rank $n^2 - (n-1)$.

Consider an example of $n = 3$. In this case, M reads as

$$\begin{pmatrix} S_{11} & S_{21} & 0 & 0 & 0 & 0 & A_{13} & 0 & 0 \\ S_{12} & S_{22} & 0 & 0 & 0 & 0 & 0 & A_{13} & 0 \\ S_{13} & S_{23} & 0 & 0 & 0 & 0 & 0 & 0 & A_{13} \\ 0 & 0 & S_{11} & S_{21} & 0 & 0 & A_{23} & 0 & 0 \\ 0 & 0 & S_{12} & S_{22} & 0 & 0 & 0 & A_{23} & 0 \\ 0 & 0 & S_{13} & S_{23} & 0 & 0 & 0 & 0 & A_{23} \\ 0 & 0 & 0 & 0 & S_{11} & S_{21} & A_{33} & 0 & 0 \\ 0 & 0 & 0 & 0 & S_{12} & S_{22} & 0 & A_{33} & 0 \\ 0 & 0 & 0 & 0 & S_{13} & S_{23} & 0 & 0 & A_{33} \end{pmatrix}.$$

After applying the column addition operations by the last three columns, we first eliminate the submatrix in the top left corner (blue):

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & A_{13} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & A_{13} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & A_{13} \\ \alpha_2 S_{11} & \alpha_2 S_{21} & S_{11} & S_{21} & 0 & 0 & A_{23} & 0 & 0 \\ \alpha_2 S_{12} & \alpha_2 S_{22} & S_{12} & S_{22} & 0 & 0 & 0 & A_{23} & 0 \\ \alpha_2 S_{13} & \alpha_2 S_{23} & S_{13} & S_{23} & 0 & 0 & 0 & 0 & A_{23} \\ \alpha_3 S_{11} & \alpha_3 S_{21} & 0 & 0 & S_{11} & S_{21} & A_{33} & 0 & 0 \\ \alpha_3 S_{12} & \alpha_3 S_{22} & 0 & 0 & S_{12} & S_{22} & 0 & A_{33} & 0 \\ \alpha_3 S_{13} & \alpha_3 S_{23} & 0 & 0 & S_{13} & S_{23} & 0 & 0 & A_{33} \end{pmatrix},$$

where $\alpha_2 = -\frac{A_{23}}{A_{13}}, \alpha_3 = -\frac{A_{33}}{A_{13}}$. Next, the third and fourth columns can be used to eliminate the column entries marked in blue. Likewise, column entries in red can be canceled by columns five and six. Hence M becomes

$$\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & A_{13} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & A_{13} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & A_{13} \\
 & & & & & & & & \\
0 & 0 & S_{11} & S_{21} & 0 & 0 & A_{23} & 0 & 0 \\
0 & 0 & S_{12} & S_{22} & 0 & 0 & 0 & A_{23} & 0 \\
0 & 0 & S_{13} & S_{23} & 0 & 0 & 0 & 0 & A_{23} \\
 & & & & & & & & \\
0 & 0 & 0 & 0 & S_{11} & S_{21} & A_{33} & 0 & 0 \\
0 & 0 & 0 & 0 & S_{12} & S_{22} & 0 & A_{33} & 0 \\
0 & 0 & 0 & 0 & S_{13} & S_{23} & 0 & 0 & A_{33}
\end{pmatrix},$$

which has rank $7 = (3^2 - (3-1))$. ■

It appears that solving the equations exactly for S^n is hopeless due to M being singular. However, a meaningful solution is possible if the actual source signals are structurally compressible, meaning that they essentially depend on a low number of degrees of freedom. For instance, if our source signal is sparse in some transformed domain, then the problem is readily simplified, and the search for solutions becomes feasible. According to analytical chemistry [14], an NMR spectrum can be presented as a sum of symmetrical, positive valued, Lorentzian-shaped peaks. Therefore, the NMR spectrum can be thought as a convolution of the Lorentzian kernel with some sparse function consisting of a few sharp peaks or, more precisely,

$$S = \hat{S} * \mathcal{L}(x, w),$$

where $\mathcal{L}(x, w) = \frac{1}{\pi} \frac{\frac{1}{2}w}{x^2 + (\frac{1}{2}w)^2}$, w specifies its width, and $\hat{S}$ is a sparse function. In Figure 5, the NMR signal (dash-dot line) is a convolution of a sparse (solid line) with Lorentzian kernel of width $= 4$. The sparsity under the Lorentzian kernel suggests that an ℓ_1 minimization problem can be formulated to recover the Pe source.

3.3. Recovery of the Pe source. Let matrix $S_{\mathrm{Po}} \in \mathbb{R}^{(n-1)\times p}$ denote the $(n-1)$ Po sources (each of dimension p) which are already recovered by the minimal cone method. The facts that the Pe source is sparse under the Lorentzian kernel and that mixture signals X may in general contain noise suggest solving the following nonnegatively constrained optimization problem:

$$\min_{\substack{A_{\mathrm{Po}} \in \mathbb{R}^{n\times(n-1)},\ A_{\mathrm{Po}} \geq 0, \\ \hat{S} \in \mathbb{R}^{n\times p},\ \hat{S} \geq 0}} \mu \|\hat{S}\|_1 + \frac{1}{2} \|X - A_{\mathrm{Po}} S_{\mathrm{Po}} - \hat{S} * \mathcal{L}\|_2^2 , \tag{3.8}$$

where n is the number of source signals and p is the number of available samples. Here A_{Po} is an $n \times (n-1)$ matrix whose columns are the mixing coefficients of the Po sources in X. Due to the nonnegativity constraint, the l_1 norm of $\hat{S}$ is a linear function. Each row of $\hat{S}$

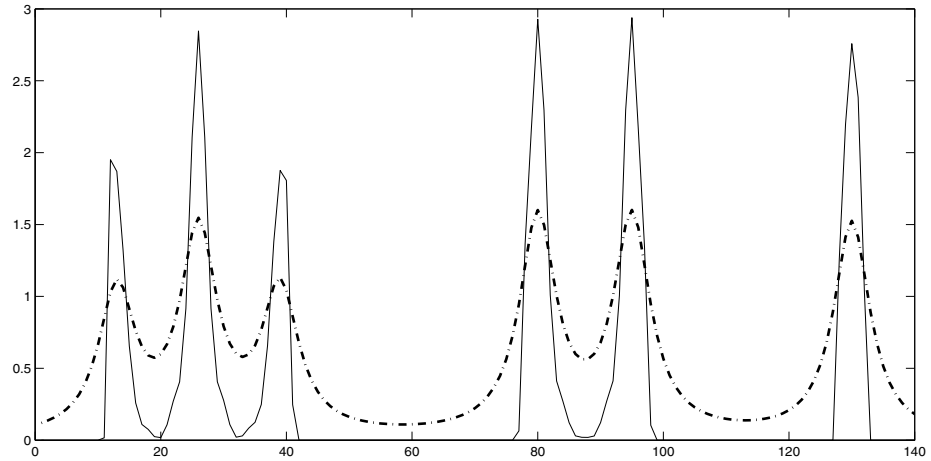

Figure 5. *NMR signal (dash-dot line) is a convolution of a sparse signal (solid line) with a Lorentzian kernel of width* $= 4$.

stands for a sparse function with a few sharp peaks, and the rows of $\hat{S} * \mathcal{L}$ are the multiples of the Pe source in the mixtures. The linear convolution $\hat{S} * \mathcal{L}$ is approximated by a matrix multiplication $\hat{S}\ L$ in the computation; here $L \in \mathbb{R}^{p\times p}$ is the discretized Lorentzian kernel. Because (3.8) allows the constraint $A_{\rm Po}\, S_{\rm Po} + \hat{S} * \mathcal{L} = X$ to be relaxed, it is applicable when the mixtures X are contaminated by measurement errors such as instrument noise. When there is minimal measurement error, one assigns a tiny value to μ to heavily weigh the fidelity term $\|X - A_{\rm Po}\, S_{\rm Po} - \hat{S} * \mathcal{L}\|_2^2$ in order for $A_{\rm Po}\, S_{\rm Po} + \hat{S} * \mathcal{L} = X$ to be nearly satisfied. In this case, one could also formulate an optimization problem as

$$\min \|\hat{S}\|_1 \ \text{ s.t. } \ A_{\rm Po}\, S_{\rm Po} + \hat{S} * \mathcal{L} = X, \ \ A_{\rm Po} \geq 0, \ \ \hat{S} \geq 0, \tag{3.9}$$

for which the Bregman iterative method [17, 46] with a proper projection onto a nonnegative convex subset can be used to obtain a solution. The equality constraint of (3.9) contains $n \times p$ equations and $n \times (n - 1 + p)$ variables, which implies that it is underdetermined. When measurement noise is minimal, the region defined by the equality and inequality constraints is nonempty.

For simplicity, we solve (3.8) by the projected gradient descent approach. The following multiplicative update rules preserve the nonnegativity constraints for nonnegative initial data:

$$a_{ij} \leftarrow a_{ij}\, \frac{\Big[\big[X\, S_{\rm Po}^{\rm T}\big]_{ij} - \big[\hat{S}\, L\, S_{\rm Po}^{\rm T}\big]_{ij}\Big]_+}{\big[A_{\rm Po}\, S_{\rm Po}\, S_{\rm Po}^{\rm T}\big]_{ij}}, \qquad \hat{s}_{jk} \leftarrow \hat{s}_{jk}\, \frac{\Big[\big[X\, L - A_{\rm Po}\, S\, L\big]_{jk} - \mu\Big]_+}{\big[\hat{S}\, L\, L\big]_{jk}}\,, \tag{3.10}$$

where nonlinear operator $[x]_+ = \max\{0, x\}$, $a_{ij} = (A_{\rm Po})_{ij}$, $\hat{s}_{jk} = (\hat{S})_{jk}$. For the convex objective (3.8), the iterations in (3.10) converge to a global minimum.

For the peak width parameter w in the computation, an estimate of an upper bound suffices. A straightforward way of estimating an upper bound is to read off an approximate value from the mixture signals if the dominant interval(s) happen to contain a peak. For more

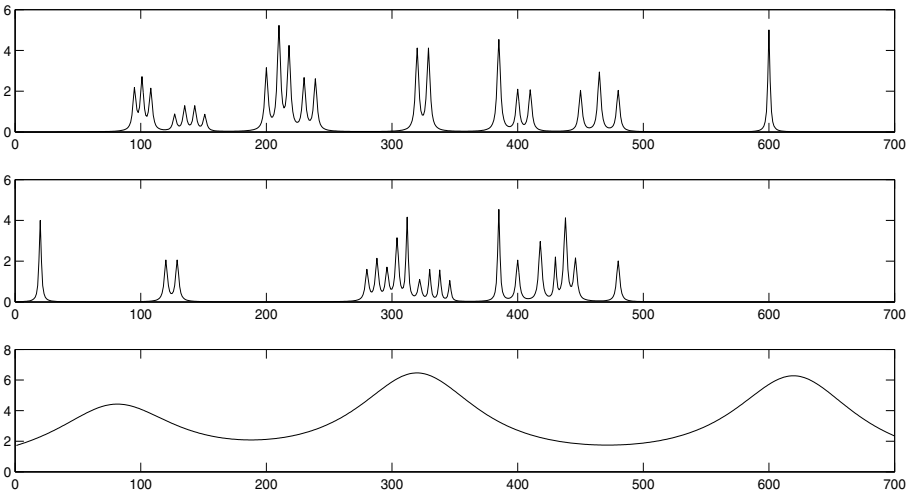

Figure 6. *The three source signals which are the idealized NMR spectra of some urine components. The first two have very narrow peaks, while the last one has wide peaks and dominant intervals in the region* $[500, 700]$ *over the other sources.*

complicated mixture signals, a low pass filter could be designed to extract a more accurate estimation. In the practice of NMR, the expertise of an analytical chemist may also be helpful in estimating this parameter.

4. Numerical results.

4.1. Examples. We present here numerical examples computed by our method. The data used in the first example are synthetic; we use the source signals in Figure 6 and generate mixtures by simulating the mixing process (1.1). The results are shown in a series of plots in Figures 7–10. The plots in Figure 7 are the three mixtures and the scatter plot of X's columns. The convex hull frame contains six points, and $\hat{A}^3$ is identified as the one attracting a cluster. Figure 8 shows the second step of the method. The first plot is the two reduced mixtures after eliminating the Pe source; note that the new mixtures contain negative values. However, the minimal cone method still recovers the two Po sources shown in the right column. In the last step of recovering the Pe source via ℓ_1 minimization, a Butterworth low pass filter is designed to estimate the upper bound of the peak width w. From Figure 9, a value $w = 130$ is read off and used in the ℓ_1 minimization. The recovered Pe source is plotted in Figure 10.

In the second and third examples the sources are real-world data, and the mixtures are constructed by mixing the sources with random matrices. The sources are the NMR spectra of quinine and urea; their noisy mixtures are in Figure 11. The result is shown in Figure 12. The quality of the separation can be seen from the comparison between the real sources and their recovery. The third example is to separate NMR spectra of menthol (Po source) and β-sitosterol (Pe source) (their spectral references are plotted in Figure 13) from the mixture signals shown in Figure 14. The calculated spectra are in Figures 15–16. Comparing the two calculated spectra of β-sitosterol in Figure 16, ℓ_1 minimization generates a sparser solution and proves to be robust to noise.

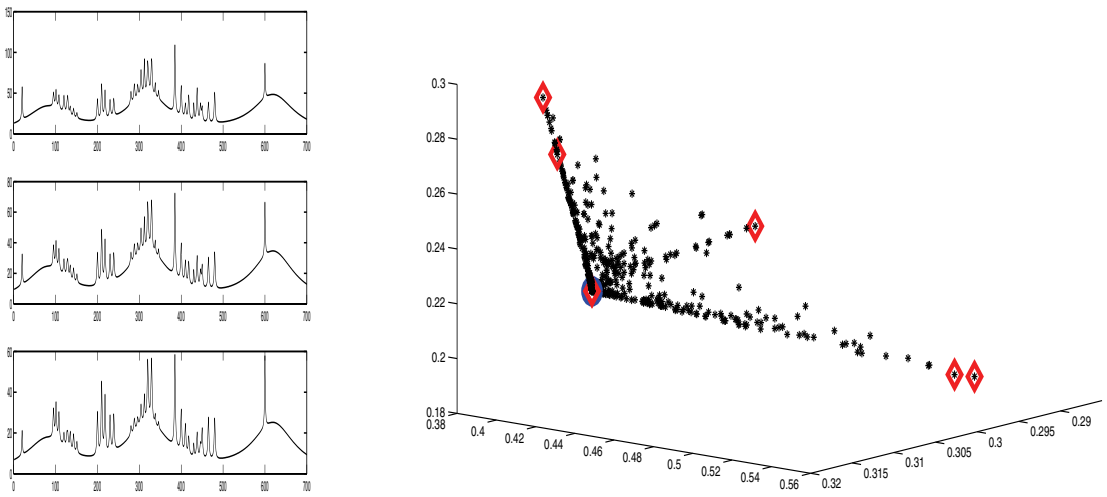

Figure 7. *First step of our method: A^3 is identified from the column vectors of X. The left plot shows the three mixtures, and the right plot is their geometry. The red diamonds are the convex hull frame of X, and A^3 is the one attracting a cluster (in the blue circle).*

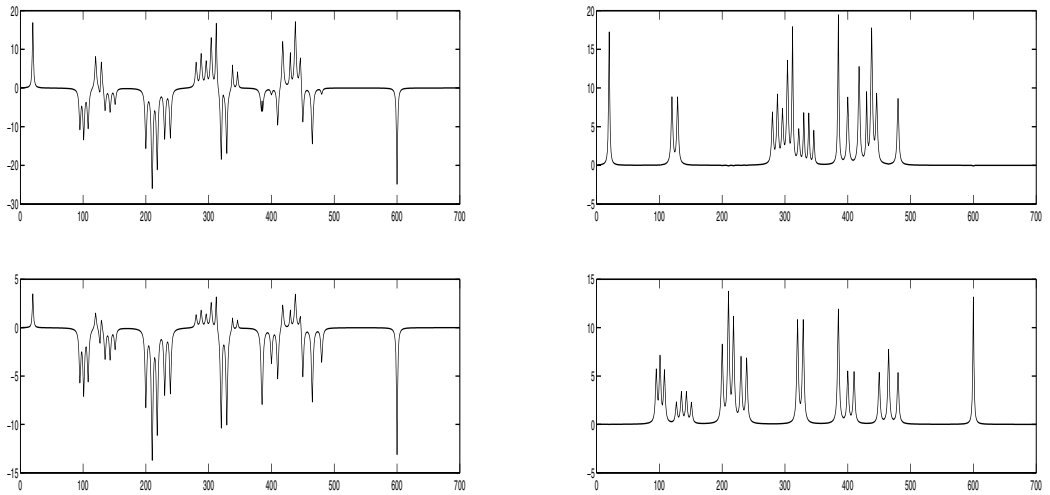

Figure 8. *Second step of our method: Two new mixtures are obtained by eliminating S_3 (left); the recovered Po sources are in the right column.*

We remark that although the proposed method is motivated by the NMR spectroscopy of biofluids, the idea can be applied to separating other NMR data. In practice, we might need to perform only data clustering and ℓ_1 minimization. Next we shall process an experimental data set. We shall use clustering to retrieve all the columns of the mixing matrix, and then retrieve the sources by solving a nonnegative ℓ_1 optimization problem. The elimination of variables (model reduction) is unnecessary for this example.

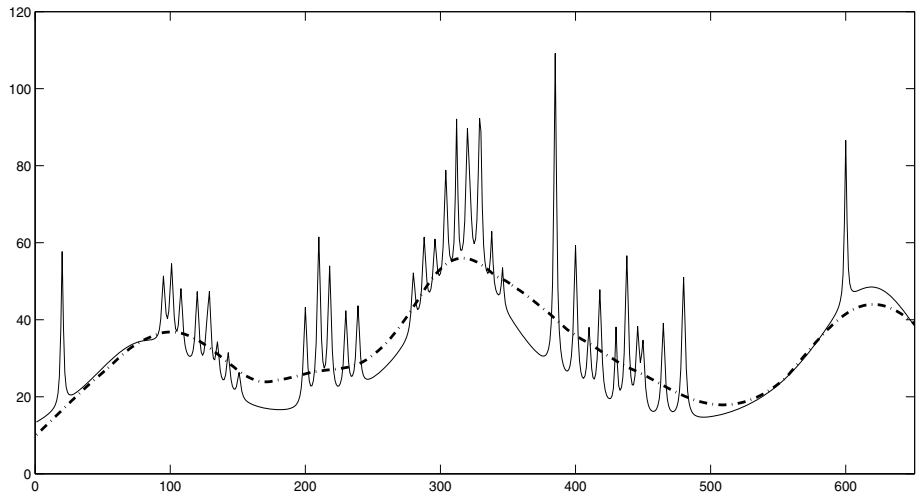

Figure 9. *A second order Butterworth low pass filter is designed to estimate the peak width of the Pe source signal. The solid curve is the mixture signal, and the dashed curve is the filtered signal.*

In this example, we provide a set of real-world data to test our method. The data are produced by diffusion ordered spectroscopy (DOSY), which is an NMR spectroscopy technique used by chemists for mixture separation [26]. However, the three compounds used in the experiment (quinine, geraniol, and camphor) have similar chemical functional groups (i.e., there is overlap in their NMR spectra) [30], for which DOSY fails to separate them. Though our working hypothesis is not satisfied completely, it is known that each of the three sources has dominant interval(s) over others in its NMR spectrum. This can also be verified from the three isolated clusters formed in their mixed NMR spectra (see the geometry of their mixtures in Figure 17). Here we separate three sources from three mixtures. Figure 17 plots the mixtures (rows of X) and their geometry (columns of X), where three clusters of points can be spotted. Then the columns of A are identified as the center points of three clusters. Note that elimination of variables is not necessary since all of A's columns have been obtained by one time clustering (K-means). The three isolated clusters imply that the source matrix S possesses column sparsity; therefore we consider solving the nonnegative ℓ_1 optimization (2.4) for each column S^i of S. The solutions are presented in Figure 18; the results are satisfactory compared to those of the ground truth. As a comparison, the source signals recovered by the convex cone method are shown in Figure 19, where $S = \text{inverse}(A)\, X$, which produces some negative (erroneous) peaks in S.

4.2. 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 ratio (SNR) when adding white Gaussian noise to the data. We also present the comparison with the convex cone method. To test the performance of the methods, we compute Comon's index [10]. The index is defined as follows: Let A and $\hat{A}$ be two nonsingular matrices with normalized columns. Then the distance between A and $\hat{A}$ is denoted by $\epsilon(A, \hat{A})$

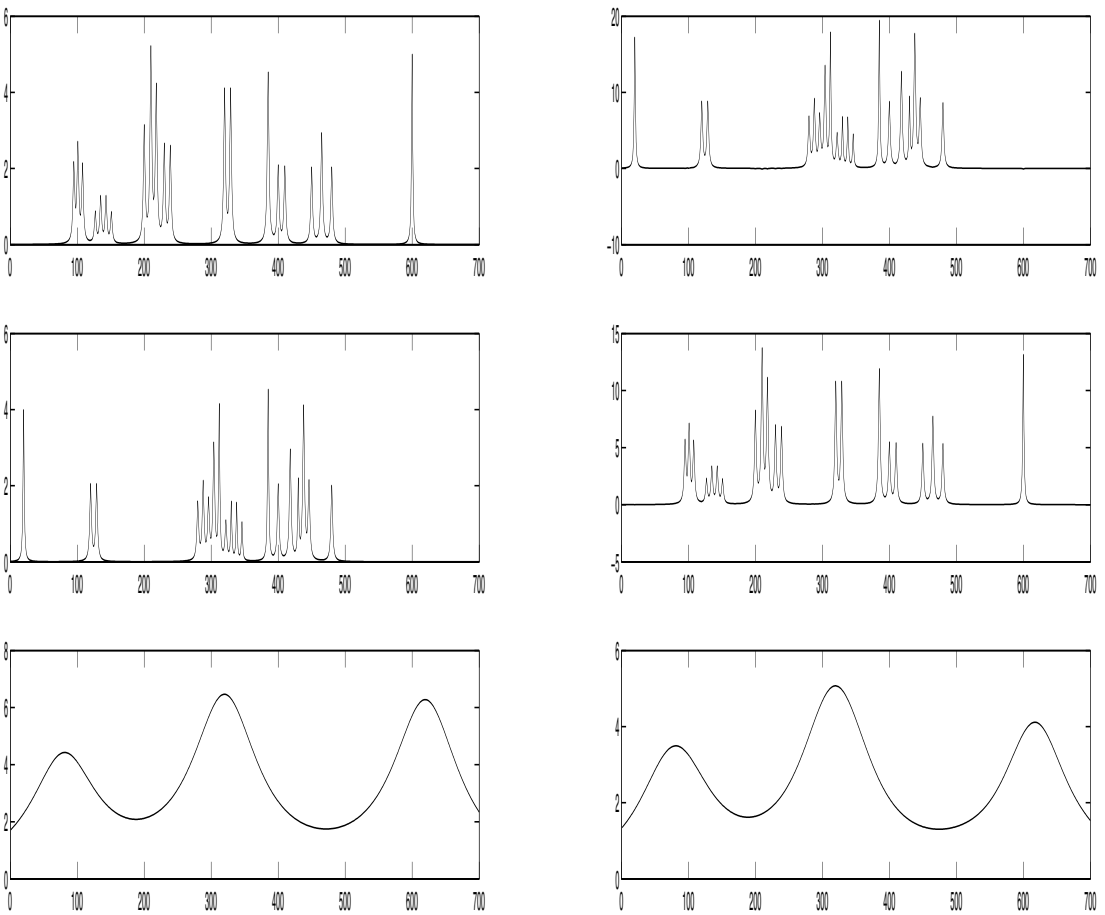

Figure 10. *Comparison: The left column shows the real sources, and the right column is the solution by our method.*

which reads as

$$\epsilon(A,\hat{A}) = \sum_i \left| \sum_j |d_{ij}| - 1 \right|^2 + \sum_j \left| \sum_i |d_{ij}| - 1 \right|^2 + \sum_i \left| \sum_j |d_{ij}|^2 - 1 \right| + \sum_j \left| \sum_i |d_{ij}|^2 - 1 \right|,$$

where $D = A^{-1}\hat{A}$ and d_{ij} is the entry of D. In [10] Comon proved that A and $\hat{A}$ are treated as nearly equivalent in the sense of BSS (i.e., $\hat{A} = A\,P\,\Lambda$) if $\epsilon(A,\hat{A}) \approx 0$. The results are shown in Figure 20; the proposed method outperforms the convex cone method, and it is robust to noise. For example, at SNR = 30 dB, for the true mixing matrix A_{True}, its estimations via

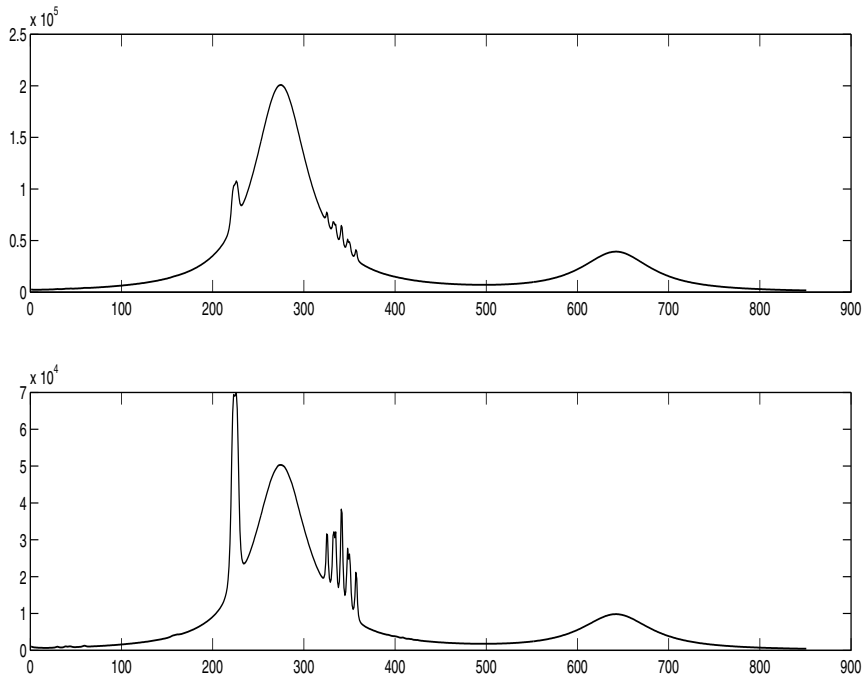

Figure 11. *Two mixtures of quinine and urea.*

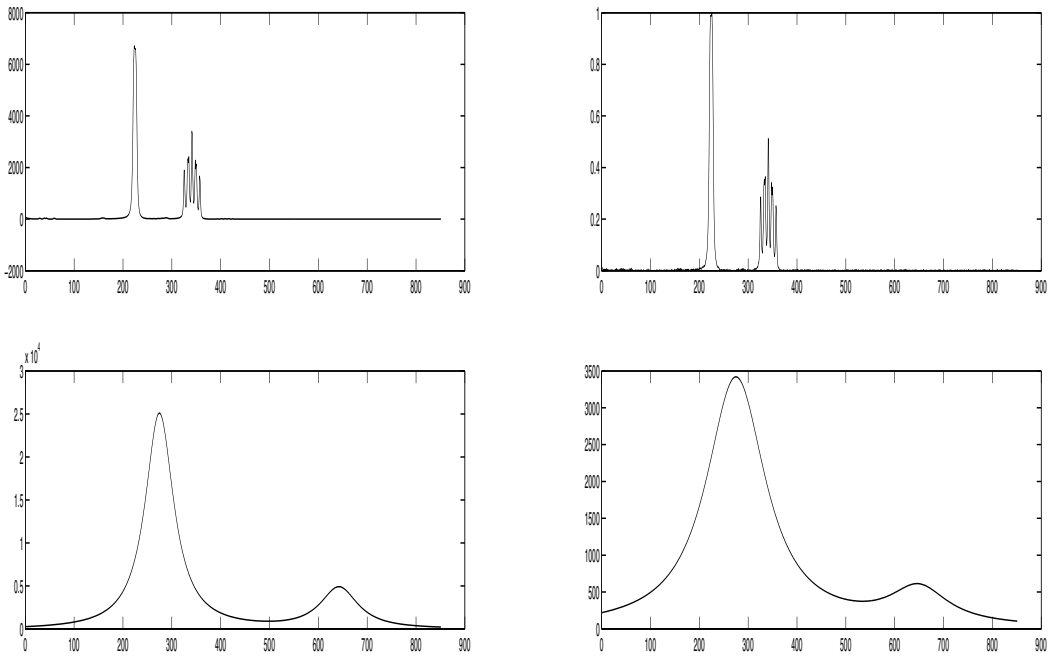

Figure 12. *Left: Reference spectra of quinine and urea. Right: The recovery by our method.*

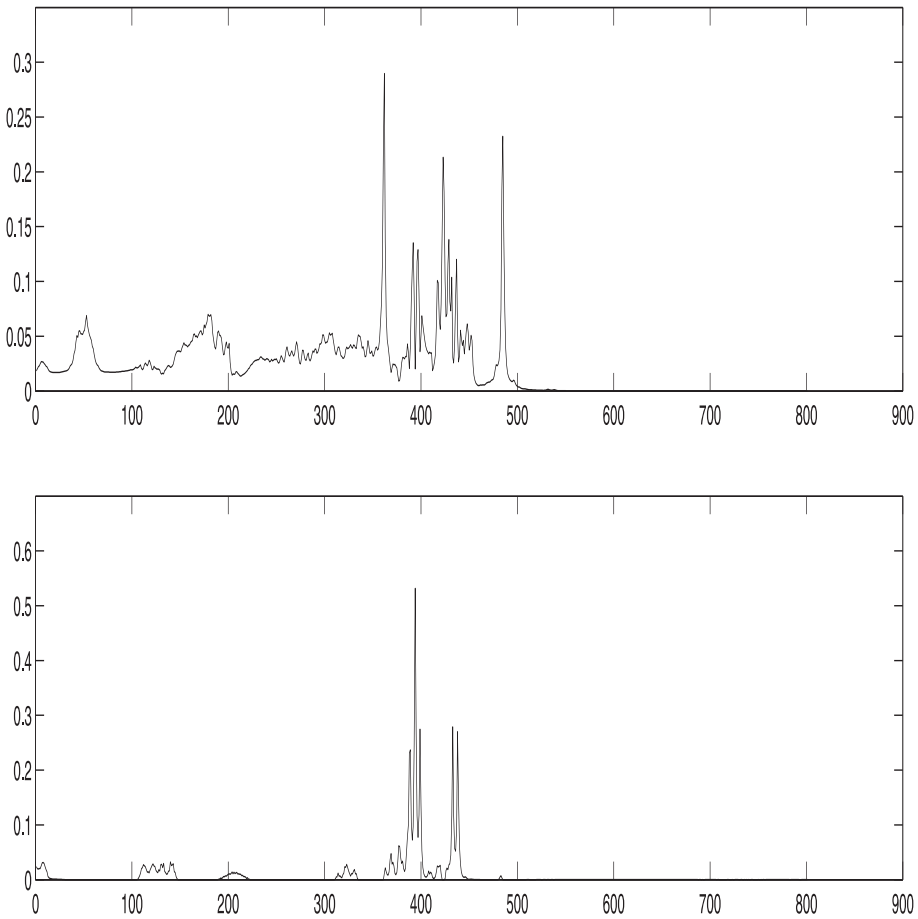

Figure 13. *Reference spectra of β-sitosterol and menthol.*

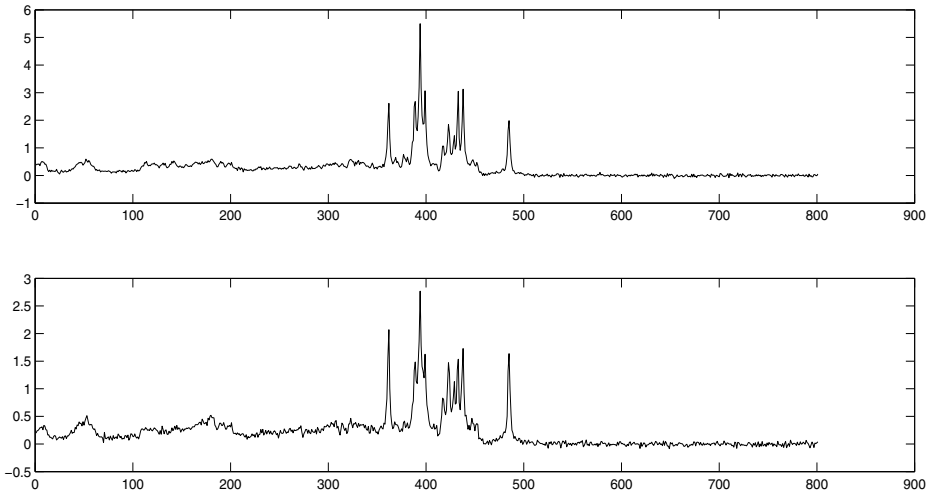

Figure 14. *Spectra of two mixture samples of menthol and β-sitosterol.*

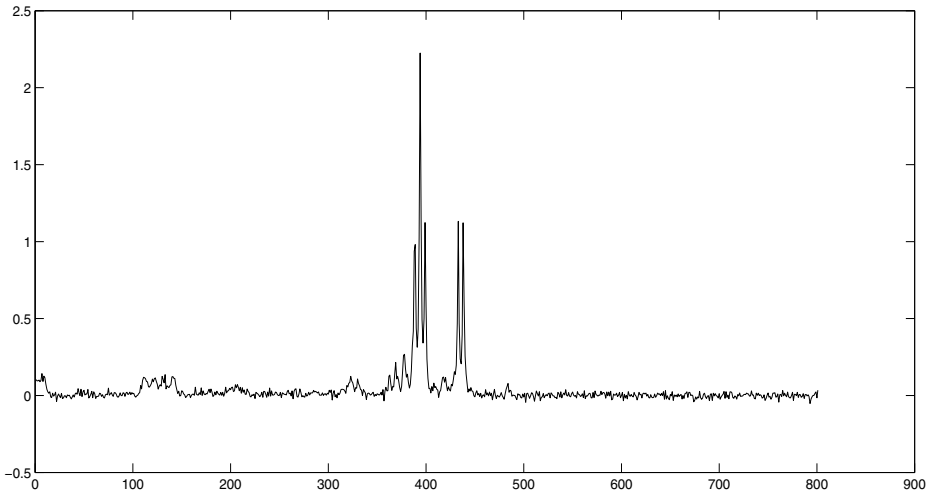

Figure 15. *Computed source spectrum of menthol (Po source).*

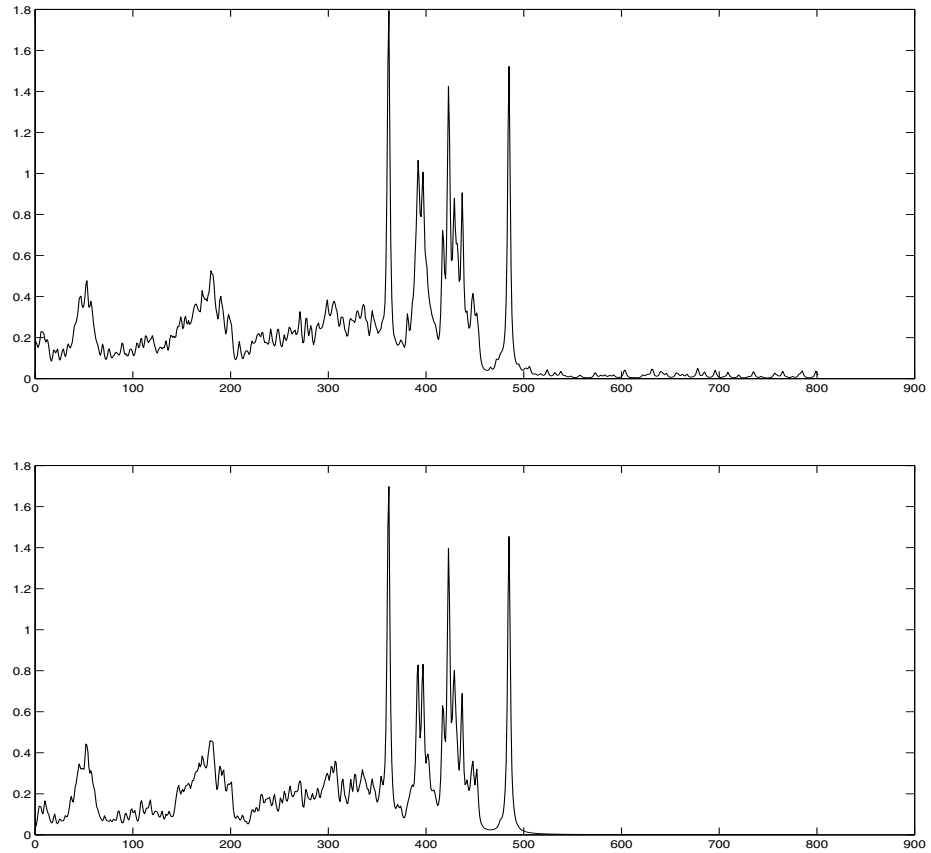

Figure 16. *Computed source spectrum of β-sitosterol (Pe source) by ℓ_1 minimization (3.8) using different μ values. Top: $\mu = 0$. Bottom: $\mu = 0.04$. The ℓ_1 penalty provides a sparser solution in the bottom panel.*

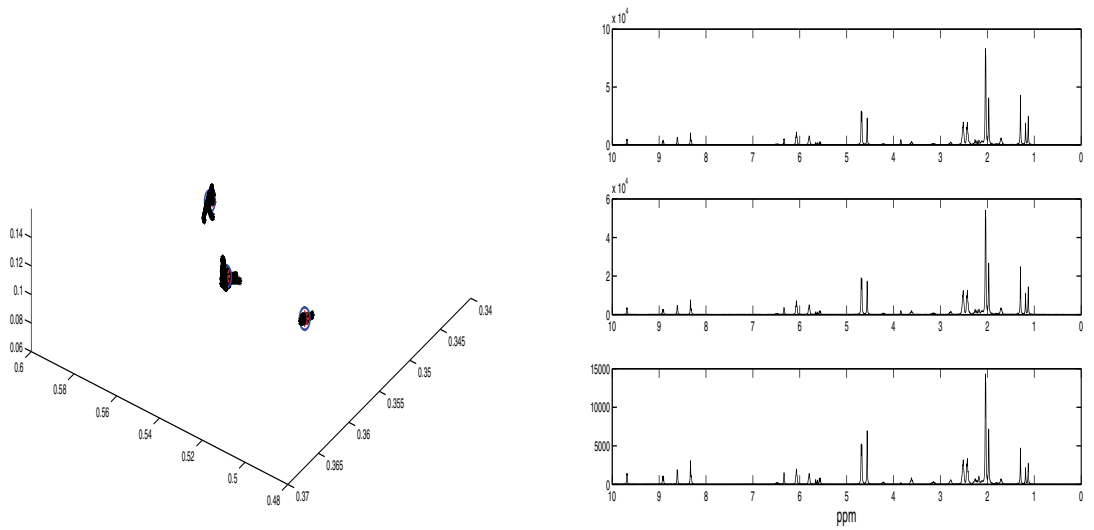

Figure 17. *Three columns of A are identified as the three center points in blue circles attracting the most points in scatter plots of the columns of X (left), and the three rows of X (right).*

the cone method (A_{cone}) and the proposed approach A_{pr} are

$$A_{True} = \begin{pmatrix} 0.8571 & 0.3244 & 0.4915 \\ 0.4286 & 0.8111 & 0.6554 \\ 0.2857 & 0.4867 & 0.5735 \end{pmatrix},$$

$$A_{cone} = \begin{pmatrix} 0.6654 & 0.4255 & 0.4917 \\ 0.5893 & 0.7279 & 0.6552 \\ 0.4582 & 0.5377 & 0.5733 \end{pmatrix},$$

$$A_{Pr} = \begin{pmatrix} 0.8433 & 0.3291 & 0.4923 \\ 0.4515 & 0.8050 & 0.6553 \\ 0.2914 & 0.4936 & 0.5728 \end{pmatrix}.$$

Figure 21 shows the comparison results of the recovered sources by the two methods. Even though two sources recovered by the cone method are recognizable compared to the ground truth, the third has spurious peaks which make it hard to identify the reliable peaks needed to perform the postprocessing. As a matter of fact, postprocessing developed in [36] deals with the regime when the SAP condition is violated to some extent; clearly the PePoDi data are intractable to the postprocessing. The proposed method in this paper becomes useful.

We also test the dependence of our method on the width parameter w. The result is presented in Figure 22. The true value is $w = 7$. It can be seen that the method achieves fairly good results if the estimation of w falls in $(4.5, 8.5)$.

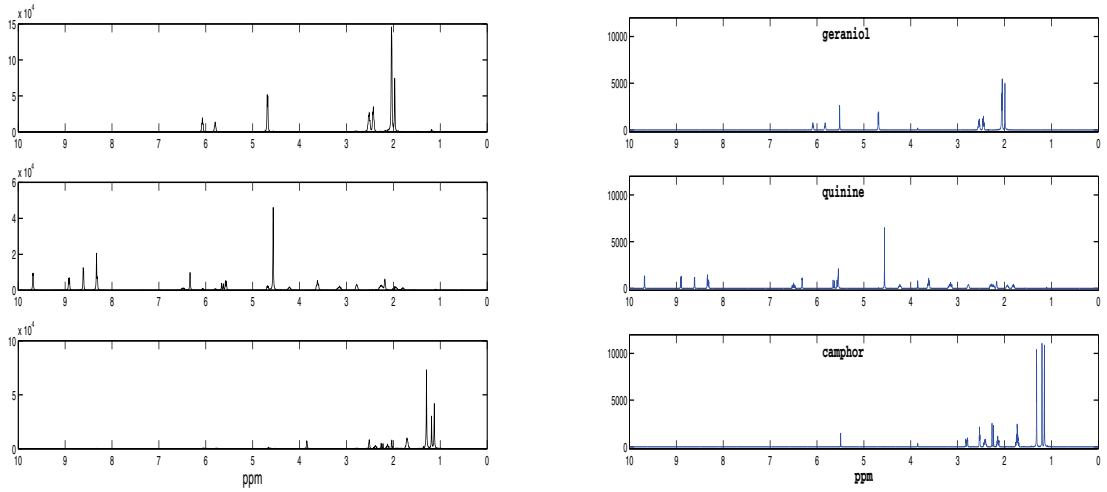

Figure 18. *The recovered source signals by nonnegative ℓ_1 minimization (left) and the ground truth (right).*

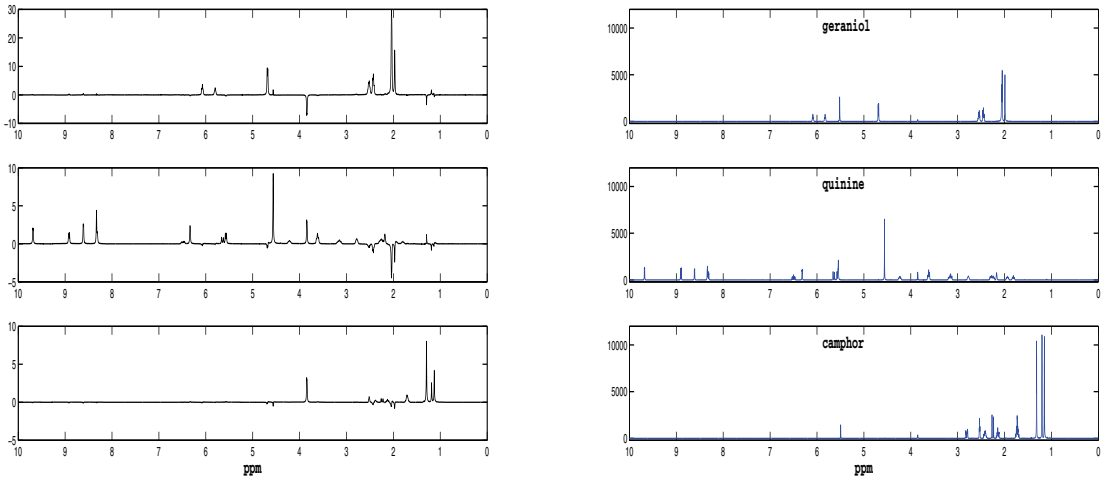

Figure 19. *The recovered source signals using the convex cone method (left) and the ground truth (right).*

5. Concluding remarks. This paper presented a novel BSS method for nonnegative and correlated data. This work is motivated by the study of the urine (serum) metabolites using NMR spectroscopy. The separation of urine-type NMR data is challenging mainly due to the heavy overlapping of some constituents (Pe sources). To deal with this class of data, this paper proposed a viable working hypothesis which enables the development of a new BSS method. The method includes three main steps. Exploiting the dominant regions in the Pe sources, the first step identifies the mixing columns of these sources. Eliminating the Pe sources is then possible, based on which the second step constructs new mixtures containing only the Po sources, for which the convex cone method is applicable. The sparsity of Pe sources under

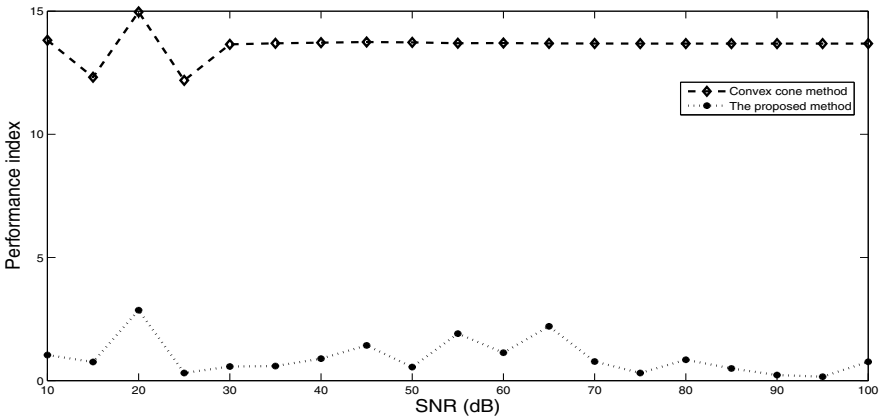

Figure 20. *Comparison of the convex cone method and the proposed approach in the presence of noises.*

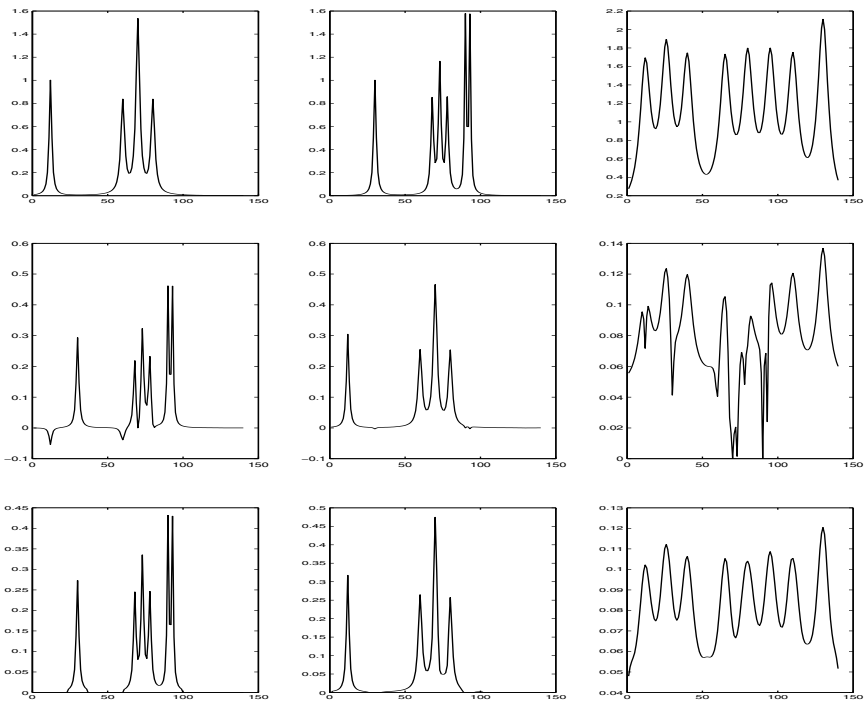

Figure 21. *The top row is the ground truth; the middle row is the result of the convex cone method; and the bottom row is the result of the proposed method.*

the Lorentzian kernel allows an efficient recovery by ℓ_1 minimization. The effectiveness of the method is validated by NMR spectra data.

The method developed in this paper deals with one Pe source which is already applicable to many real-world data. The first novelty of this paper is the formulation of a clustering problem exploiting the geometry of the mixtures, which makes possible an accurate estimation

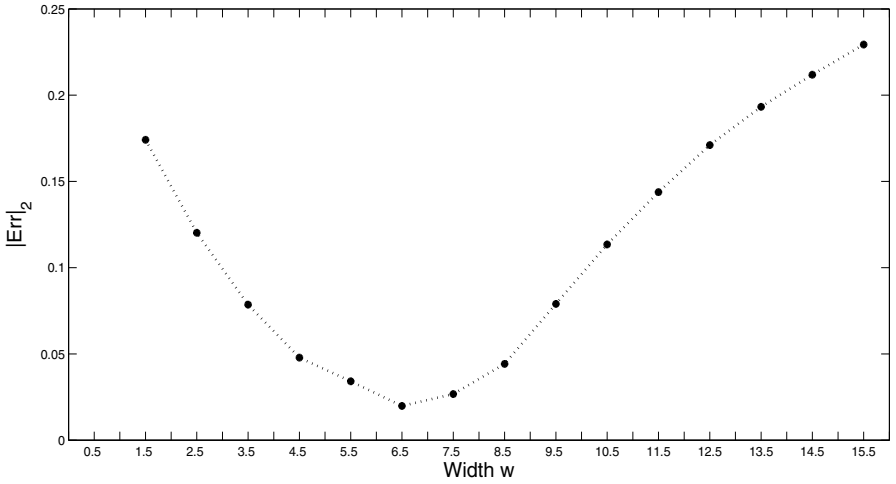

Figure 22. *The dependence of the error on the width parameter. Gaussian noise of SNR = 20 dB is added to the data.*

of a partial mixing matrix by means of linear programming and a clustering technique. The second novelty is the reduction of the model to a subproblem for which effective methods are available. The major strength of this work is in converting a nonconvex problem to a convex one by exploiting the data structures, which is nontrivial and requires a nonstandard approach. It is worthwhile to pursue this line of inquiry for more NMR data in the future by a combination of clustering, model reduction, and ℓ_1 minimization. In practice, more than one Pe NMR source could exist in the mixture signals. An ongoing project is to extend our method from treating one Pe source to multiple Pe sources. In this case, the current source condition PePoDi should be replaced by a constraint which requires dominant interval(s) from each source signal over some of the other source signals in a hierarchical manner. To this end, a more complicated method needs to be developed.

Acknowledgments. The authors thank Professor A. J. Shaka and Mr. Hasan Celik of the Chemistry Department of UC Irvine for helpful discussions.

REFERENCES

[1] R. Barton, J. Nicholson, P. Elliot, and E. Holmes, *High-throughput 1H NMR-based metabolic analysis of human serum and urine for large-scale epidemiological studies: Validation study*, Int. J. Epidemiol., 37 (2008), pp. i31–i40.

[2] A. Bijaoui and D. Nuzillard, *Blind source separation of multispectral astronomical images*, in Mining the Sky: Proceedings of the MPA/ESO/MPE Workshop, Garching, Germany, 2000, A. J. Banday, S. Zaroubi, and M. Bartelmann, eds., Springer-Verlag, Berlin, 2001, pp. 571–581.

[3] J. Boardman, *Automated spectral unmixing of AVIRIS data using convex geometry concepts*, in Summaries of the IV Annual JPL Airborne Geoscience Workshop, JPL Pub. 93-26, Vol. 1, Jet Propulsion Laboratory, Pasadena, CA, 1993, pp. 11–14.

[4] J. Bobin, J.-L. Starck, J. Fadili, and Y. Moudden, *Sparsity and morphological diversity in blind source separation*, IEEE Trans. Image Process., 16 (2007), pp. 2662–2674.

[5] P. Bofill and M. Zibulevsky, *Underdetermined blind source separation using sparse representations*, Signal Process., 81 (2001), pp. 2353–2362.
[6] G. Brown and D.-L. Wang, *Separation of speech by computational auditory scene analysis*, in Speech Enhancement, J. Benesty, S. Makino, and J. Chen, eds., Springer, New York, 2005, pp. 371–402.
[7] C.-I. Chang, ed., *Hyperspectral Data Exploitation: Theory and Applications*, Wiley-Interscience, Hoboken, NJ, 2007.
[8] 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.
[9] A. Cichocki and S. Amari, *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*, John Wiley and Sons, New York, 2005.
[10] P. Comon, *Independent component analysis, A new concept?,* Signal Process., 36 (1994), pp. 287–314.
[11] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applications*, Academic Press, Oxford, UK, 2010.
[12] I. Drori, *Fast ℓ_1 minimization by iterative thresholding for multidimensional NMR spectroscopy*, EURASIP J. Adv. Signal Process., No. 23 (2007), 20248.
[13] 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.
[14] R. Ernst, G. Bodenhausen, and A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Oxford University Press, Oxford, UK, 1987.
[15] P. Georgiev, F. Theis, and A. Cichocki, *Sparse component analysis and blind source separation of underdetermined mixtures*, IEEE Trans. Neural Netw., 16 (2005), pp. 992–996.
[16] X. Guo, C. Chang, and Y. Lam, *Blind separation of electron paramagnetic resonance signals using diversity minimization*, J. Magn. Reson., 204 (2010), pp. 26–36.
[17] Z. Guo and S. Osher, *Template Matching via l_1 Minimization and Its Application to Hyperspectral Target Detection*, Technical report 09-103, UCLA, Los Angeles, CA, 2009; available online from http://www.math.ucla.edu/applied/cam/.
[18] P. Hoyer, *Non-negative matrix factorization with sparseness constraints*, J. Mach. Learn. Res., 5 (2004), pp. 1457–1469.
[19] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley and Sons, New York, 2001.
[20] 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.
[21] 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*, Anal. Chim. Acta, 653 (2009), pp. 143–153.
[22] D. D. Lee and H. S. Seung, *Learning the parts of objects by non-negative matrix factorization*, Nature, 401 (1999), pp. 788–791.
[23] J. Liu, J. Xin, and Y.-Y. Qi, *A dynamic algorithm for blind separation of convolutive sound mixtures*, Neurocomput., 72 (2008), pp. 521–532.
[24] J. Liu, J. Xin, and Y. Qi, *A soft-constrained dynamic iterative method of blind source separation*, Multiscale Model. Simul., 7 (2009), pp. 1795–1810.
[25] J. Liu, J. Xin, Y.-Y. Qi, and F.-G. Zeng, *A time domain algorithm for blind separation of convolutive sound mixtures and L_1 constrained minimization of cross correlations*, Commun. Math. Sci., 7 (2009), pp. 109–128.
[26] G. Morris, *Diffusion-ordered spectroscopy (DOSY)*, in Encyclopedia of Nuclear Magnetic Resonance, D. M. Grant and R. K. Harris, eds., John Wiley and Sons, New York, 2002, pp. 35–44.
[27] S. Moussaoui, H. Hauksdóttir, F. Schmidt, C. Jutten, J. Chanussot, D. Briee, S. Douté, and J. A. Benediktsson, *On the decomposition of Mars hyperspectral data by ICA and Bayesian positive source separation*, Neurocomput., 71 (2008), pp. 2194–2208.
[28] W. Naanaa and J.-M. Nuzillard, *Blind source separation of positive and partially correlated data*, Signal Process., 85 (2005), pp. 1711–1722.
[29] M. Naceur, M. Loghmari, and M. Boussema, *The contribution of the sources separation method in the decomposition of mixed pixels*, IEEE Trans. Geosci. Remote Sens., 42 (2004), pp. 2642–2653.
[30] M. Nilsson, M. Connel, A. Davies, and G. Morris, *Biexponential fitting of diffusion-ordered NMR data: Practicalities and limitations*, Anal. Chem., 78 (2006), pp. 3040–3045.

[31] D. Nuzillard and A. Bijaoui, *Blind source separation and analysis of multispectral astronomical images*, Astron. Astrophys. Suppl. Ser., 147 (2000), pp. 129–138.
[32] D. Nuzillard, S. Bourg, and J.-M. Nuzillard, *Model-free analysis of mixtures by NMR using blind source separation*, J. Magn. Reson., 133 (1998), pp. 358–363.
[33] M. Plumbley, *Conditions for non-negative independent component analysis*, IEEE Signal Process. Lett., 9 (2002), pp. 177–180.
[34] M. Plumbley, *Algorithms for nonnegative independent component analysis*, IEEE Trans. Neural Netw., 4 (2003), pp. 534–543.
[35] R. M. Silverstein, F. X. Webster, and D. J. Kiemle, *Spectrometric Identification of Organic Compounds*, John Wiley and Sons, Hoboken, NJ, 2005.
[36] 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*, Signal Process., 91 (2011), pp. 1838–1851.
[37] Y. Sun and J. Xin, *Underdetermined sparse blind source separation of nonnegative and partially overlapped data*, SIAM J. Sci. Comput., 33 (2011), pp. 2063–2094.
[38] K. Stadlthanner, A. Tomé, F. Theis, W. Gronwald, H.-R. Kalbitzer, and E. Lang, *On the use of independent component analysis to remove water artifacts of 2D NMR protein spectra*, in Proceedings of the 7th Portugese Conference on Biomedical Engineering (BIOENG'2003), 2003.
[39] I. Takigawa, M. Kudo, and J. Toyama, *Performance analysis of minimum l_1-norm solutions for underdetermined source separation*, IEEE Trans. Signal Process., 52 (2004), pp. 582–591.
[40] F. J. Theis, C. G. Puntonet, and E. W. Lang, *A histogram-based overcomplete ICA algorithm*, in Proceedings of the 4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA 2003), Nara, Japan, 2003.
[41] C. Vitols and A. Weljie, *Identifying and Quantifying Metabolites in Blood Serum and Plasma*, Chenomx Application Note, Chenomx Inc., Edmonton, AB, Canada, 2006.
[42] M. E. Winter, *N-findr: An algorithm for fast autonomous spectral end-member determination in hyperspectral data*, in Proceedings of the SPIE, Vol. 3753, SPIE, Bellingham, WA, 1999, pp. 266–275.
[43] W. Wu, M. Daszykowski, B. Walczak, B. C. Sweatman, S. Connor, J. Haselden, D. Crowther, R. Gill, and M. Lutz, *Peak alignment of urine NMR spectra using fuzzy warping*, J. Chem. Inf. Model., 46 (2006), pp. 863–875.
[44] W. Yang, Y. Wang, Q. Zhou, and H. Tang, *Analysis of human urine metabolites using SPE and NMR spectroscopy*, Sci. China Ser. B-Chem., 51 (2008), pp. 218–225.
[45] Ö. Yilmaz and S. Rickard, *Blind separation of speech mixtures via time-frequency masking*, IEEE Trans. Signal Process., 52 (2004), pp. 1830–1847.
[46] W. Yin, S. Osher, D. Goldfarb, and J. Darbon, *Bregman iterative algorithm for ℓ_1-minimization with applications to compressive sensing*, SIAM J. Imaging Sci., 1 (2008), pp. 143–168.