

Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Postprocessing and sparse blind source separation of positive and partially overlapped data

Y. Sun^{a,*}, C. Ridge^b, F. del Rio^b, A.J. Shaka^b, J. Xin^a

^a Department of Mathematics, University of California at Irvine, Irvine, CA 92697, USA

^b Department of Chemistry, University of California at Irvine, Irvine, CA 92697, USA

ARTICLE INFO

Article history:

Received 5 February 2010

Received in revised form

10 November 2010

Accepted 11 February 2011

Available online 22 February 2011

Keywords:

Blind source separation

Postprocessing

Noise detection and reduction

Peak-based corrections

Nuclear magnetic resonance

ABSTRACT

We study sparse blind source separation (BSS) for a class of positive and partially overlapped signals. The signals are only allowed to have nonoverlapping at certain locations, while they could overlap with each other elsewhere. For nonnegative data, a novel approach has been proposed by Naanaa and Nuzillard (NN) assuming that nonoverlapping exists for each source signal at some location of acquisition variable. However, the NN method introduces errors (spurious peaks) in the output when their nonoverlapping condition is not satisfied. To resolve this problem and improve robustness of separation, postprocessing techniques are developed in two aspects. One is to detect coherent and uncertain components from NN outputs by using multiple mixture data, then removing the uncertain portion to enhance signals. The other is to find better estimation of mixing matrix by leveraging reliable source peak structures in NN output. Numerical results on examples including NMR spectra of a ¹³C-1-acetylated carbohydrate with overlapping proton spin multiplets show satisfactory performance of the postprocessed sparse BSS and offer promise to resolve complex spectra without using multidimensional NMR methods.

Published by Elsevier B.V.

1. Introduction

A blind source separation (BSS) problem arises when one attempts to recover source signals from their linear mixtures without detailed knowledge of the mixing process. BSS has received considerable attention in signal analysis and processing of speech, image, and biomedical signals [3,5,6,11,13,15,17,19,20]. The linear BSS model takes the following algebraic form:

$$X = AS + N, \quad (1.1)$$

where $X \in \mathbb{R}^{m \times p}$ is the known mixed signals, $S \in \mathbb{R}^{n \times p}$ represents the unknown source signals, and $A \in \mathbb{R}^{m \times n}$ is the unknown mixing matrix, and $N \in \mathbb{R}^{m \times p}$ is an unknown

matrix representing errors or sensor noise. The integer p is the number of available samples, m is the number of mixture signals, and n is the number of sources. In the paper we shall consider the (over)-determined case ($n \leq m$). The X , S and N are sampled functions of an acquisition variable which may be time, frequency, position, wavenumber, etc. depending on the underlying physical process. The objective of BSS is to estimate the original source signals (S) without knowing the parameters of mixing process (A), which is known as a nonconvex problem. In fact, a unique estimation is never possible without some *a priori* knowledge. However, one can usually estimate them up to certain indeterminacies which can be expressed as arbitrary scaling, permutation, etc. If P is a permutation matrix and Λ an invertible diagonal matrix then $AS = (AP\Lambda)(\Lambda^{-1}P^{-1}S)$. For BSS problems, S and $\Lambda^{-1}P^{-1}S$ are equivalent.

Various approaches have been proposed to solve BSS problems by exploiting *a priori* information about mixing

* Corresponding author.

E-mail address: sunyuanc@gmail.com (Y. Sun).

process or properties of source signals. For instance, independent component analysis (ICA) [7,8]) 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 nonnegative constraints are imposed for the mixing matrix A and/or estimated source signals S [2,10,21,22,24]. Ref. [2] develops a method for BSS when there are fewer mixtures than sources (underdetermined BSS), and the method is specially designed for two mixtures. The source matrix S is allowed to have only one nonzero entry in each column, which is a very restrictive condition. Under this condition the estimation of A can be achieved by clustering (e.g., K-means). Ref. [10] studies the solvability conditions and algorithms on separating sparse/nonnegative sources in an underdetermined case. Ref. [24] reports an application of ICA for signal artifacts removal from nuclear magnetic resonance (NMR) data, while [21,22] offer theoretical and algorithmic studies of nonnegative ICA.

The present work is motivated by the NMR spectroscopy data in which it should not be assumed that there is statistical independence, especially when the molecules responsible for each source may share common structural features. Besides, the properly phased absorption-mode NMR spectral signals from a single-pulse experiment are positive. Therefore, ICA-based methods mentioned above 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 [16]. The source signals are only 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 [14]. 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, NN's sparseness assumption and the geometric construction of columns of A were known in the 1990s [1,26] in the problem of blind hyperspectral unmixing, where the same mathematical model (1.1) is used. The analogue of NN's assumption is called pixel purity assumption [4]. The resulting geometric (cone) method is the so-called N-findr [26], and is now a benchmark in hyperspectral unmixing.

In this paper, we consider how to generalize NN method if their partial sparseness condition is not satisfied. We are concerned with the regime where source signals are nowhere nonoverlapping yet one source signal dominates others at certain locations of acquisition variable. The idea is to apply the NN method first, and then postprocess by using either extra mixture data or reliable peak structures of source signals in the NN output. In the first case, additional mixtures allow multiple NN outputs which can be decomposed into a sum of coherent and uncertain components with the help of a similarity measure. By filtering out the uncertain components, we enhance the signal contents of the NN output and

reduce the errors due to lack of nonoverlapping (partial sparseness) of the sources. In the second case, we observed that the main spectral peaks are often reliably recovered by NN method even though its partial sparseness assumption is not strictly valid, and that these robust peaks can be used as anchor points to re-estimate columns of mixing matrix A by simple algebraic manipulations.

The outline of the paper is as follows. In Section 2, we review the essentials of the NN method and its sparseness assumptions, then state our relaxed and more general source assumptions permitting source signal overlap. In Section 3, we introduce the decomposition of multiple NN outputs into coherent and random components for error reduction. The similarity measure is based on the cross-correlation function. In Section 4, we show that with partial information on locations of signal peaks, a method of peak-based corrections (PBC) improves the mixing matrix estimation. In Section 5, numerical experiments on NMR mixture data are performed to demonstrate the significant enhancements by the proposed postprocessing methods. Concluding remarks are in Section 6.

2. Sparse BSS and source assumptions

2.1. NN's source conditions and method

In [16], Naanaa and Nuzillard (NN) presented an efficient sparse BSS method and its mathematical analysis for nonnegative and partially orthogonal signals such as NMR spectra. Consider the over-determined or determined regime, where the number of mixtures is not less than the number of sources ($m \geq n$) and the mixing matrix A is full rank. In simple terms, NN's key sparseness assumption (referred to as NNA below) on source signals is that each source has a stand alone peak at some location of the 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 $i \in \{1, 2, \dots, n\}$ there exists a $j_i \in \{1, 2, \dots, p\}$ such that $s_{i,j_i} > 0$ and $s_{k,j_i} = 0$ ($k = 1, \dots, i-1, i+1, \dots, n$).

Consider the noiseless case ($N=0$), then Eq. (1.1) can be rewritten in terms of columns as

$$X^j = \sum_{k=1}^n s_{k,j} A^k, \quad j = 1, \dots, p, \quad (2.1)$$

where X^j denote the j th column of X , and A^k the k th column of A . Assumption NNA implies that $X^j = s_{i,j_i} A^i$, $i = 1, \dots, n$ or $A^i = (1/s_{i,j_i}) X^j$. Hence Eq. (2.1) is rewritten as

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

which says that every column of X is a nonnegative linear combination of the columns of \hat{A} . Here $\hat{A} = [X^{j_1}, \dots, X^{j_n}]$ 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 ($i=1, \dots, n$) 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 [9]. 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 the noiseless case, the following constrained equations are formulated for the identification of \hat{A} :

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

Then any column X^k will be a column of \hat{A} if and only if the constrained equation (2.3) is inconsistent. However, if noises are present, the following optimization problems are suggested to estimate the mixing matrix:

$$\text{minimize score} = \left\| \sum_{j=1, j \neq k}^p X^j \lambda_j - X^k \right\|_2, \quad k = 1, \dots, p$$

subject to $\lambda_j \geq 0$.

A score is associated with each column. 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 of X . Practically, the n columns from X with the 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 to S is obtained: $\hat{S} = \hat{A}^+ X$.

Implementation of NN method showed that it is both accurate and efficient if NNA condition holds [16]. However, if the condition is not satisfied, errors and artifacts may occur because the true mixing matrix is no longer the smallest enclosing convex cone of columns of data matrix. This is the main issue we shall address in this paper.

2.2. A more general source condition

A more general assumption applicable to NMR source signals S is the following relaxed NNA condition (referred to as rNNA):

Assumption (rNNA). For each $i \in \{1, 2, \dots, n\}$ there exists a $j_i \in \{1, 2, \dots, p\}$ such that $s_{i, j_i} > 0$ and $s_{k, j_i} = \varepsilon_k$ ($k = 1, \dots, i-1, i+1, \dots, n$), where $\varepsilon_k \ll s_{i, j_i}$.

Simply said, each source signal has a dominant peak at acquisition position where the other sources are allowed to be nonzero. For example, rNNA condition holds for NMR data consisting of positive-valued absorption-mode Lorentzian-shaped peaks with tails extending over the whole range of acquisition variable, see the first plot in Fig. 2.1 for example. From a geometric perspective, if NNA is satisfied and A is full rank, the column vectors of A expand the minimal cone containing all the columns of X . When the mixture matrix X comes from mixing sources satisfying rNNA, the columns of X 's are not parallel to those of A 's. The estimated mixing matrix extracted from columns of X 's according to NN method will thus deviate from the true A , causing errors in the recovered source signals. The second plot in Fig. 2.1 shows two mixtures of two sources satisfying rNNA. The true cone, spanned by the two black diamond dots, is wider than the set of column vectors of X (the collection of circles). The NN method identifies the smallest cone containing the mixtures (spanned by the two black circular dots).

We realize that without NNA, the problem of blind signal recovery or nonnegative matrix factorization is not unique, as there are infinitely many solutions for columns of mixing matrix A even up to scaling factors. In the second plot of Fig. 2.1, any cone containing the columns of X may be a choice of columns of A . In the next section, two postprocessing techniques are proposed to correct NN output based on either the random nature of NN errors when NNA is violated or additional prior information about the possible sources.

3. Random error detection and removal

Let us consider how to correct the output of NN method to reduce its errors.

3.1. Random error detection and reduction

Consider the case of more mixtures than the number of sources ($m \geq n+1, n \geq 2$), and assume that multi-peak source signals satisfy condition rNNA. NN method may produce spurious peaks on rNNA data, although it is still able to recover main features such as large peak locations. The idea is to detect coherent and incoherent components of the NN output. The coherent component

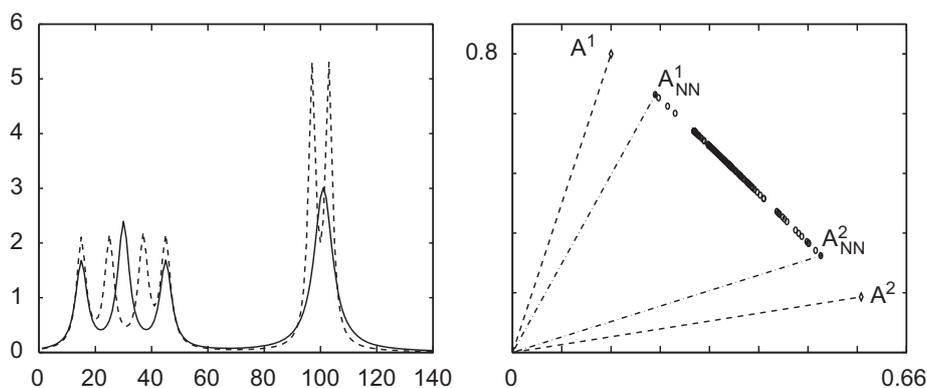


Fig. 2.1. Left: Two positive and overlapped Lorentzian source signals. At each acquisition point, either source has nonzero response. Right: Comparison of the two columns of mixing matrix recovered from NN with those of the true mixing matrix A . NN method identifies the columns of mixing matrix as the edges of a minimal cone enclosing the mixtures. The deviation of NN's results is due to the violation of the condition NNA.

is kept while the incoherent component is discarded as errors. Details of implementation are described below.

To detect incoherence, we shall make use of the abundance of mixtures available. Theoretically we have C_m^n (number of combinations of choosing n from m) groups with n mixtures in each group. The problem is of determined type (n mixture and n sources) in each group. It is unnecessary to use all C_m^n groups, although the results are better when more groups are employed. Let us select k groups of mixtures ($k \geq n$) and label them G_1, \dots, G_k . Next NN's method is applied to the k groups and suppose $S(1), \dots, S(k)$ are the output source matrices. Each row of $S(i)$ ($i=1, \dots, k$) represents a source signal, there are k rows which correspond to one source signal. But these k rows have to be automatically selected from the rows of $S(1), \dots, S(k)$ via a selective criterion. If two rows correspond to the same source, they should be correlated. The cross-correlation function can be used as a measure of similarity. The process of selecting k similar rows is as follows:

1. Normalize the NN output signals and subtract their means.
2. Take a row from $S(1)$ and calculate its inner product with all rows of $S(2)$.
3. Select the row of $S(2)$ with the largest product and put them in one group.
4. Repeat steps 2 and 3 for $S(3), \dots, S(k)$.
5. Repeat steps 2–4 for other rows of $S(1)$.

After the selection process is finished, there should be n groups with each of them representing one of the n source signals.

Next, we take one group to illustrate how to filter out the incoherence and recover a better source signal. The intermediate signals in the group typically share common and robust features (large peak structures) though they are recovered from different mixtures. On the other hand, they have differences which are random in nature. To remove the uncertain part, the following decomposition is performed:

1. Divide every intermediate signal into segments with length L . Normalize the segments and subtract their means.
2. Take the first segments from all intermediate signals and compute the inner product between any two of them. If the arithmetic average of pairwise inner products exceeds a preset threshold value, redefine the segment by the arithmetic average of all the first segments in the group; otherwise the NN output in this segment is regarded as containing sufficient random error and is redefined as zero.
3. Repeat step 2 for other segments.
4. Piece together processed segments in steps 2 and 3 to form the postprocessed signal.

In step 1, length L should be chosen to be at least the width of the smallest peak in the intermediate signal. The threshold value is set to a number near 1 (above 0.95) with the help of synthetic mixtures of known chemical compounds.

The above postprocessing is designed for better source signals S . It may not improve the estimation of mixing coefficients. The next postprocessing method aims to extract a better estimation of A from NN output based on additional peak information.

4. Peak-based correction

In order to establish the theoretical basis of the method of peak-based correction, we propose to strengthen condition rNNA by the pairwise overlap condition (POC) on the source signals ($n \geq 2$).

Assumption (POC). Each source signal has a dominant peak at some acquisition location, where other source signals are allowed to be nonzero. Furthermore, there exist different acquisition regions where the source signals overlap each other pairwise.

Clearly POC is a relaxation to condition NNA. Simply said, this condition requires not only that each source has a dominant peak at some location of acquisition variable, but also that there are only two active sources at certain acquisition region. Below we shall study two determined cases and introduce the method via examples.

4.1. $(m,n)=(2,2)$

In the case of two mixtures and two sources, the condition POC is equivalent to rNNA. The linear model can be written as

$$[X^1, \dots, X^p] = [A^1, A^2] \times \begin{pmatrix} S_{11} & \dots & S_{1p} \\ S_{21} & \dots & S_{2p} \end{pmatrix}, \quad (4.1)$$

where A^1, A^2 are the two columns of A . Based on recovered source signals using NN method, we intend to improve the estimate of the mixing matrix, which in turn improves the source signals. It has been noted that NN still recovers the major peak locations of source signals for rNNA data. As shown in Fig. 4.1, the first peaks from S_1 and S_2 are used to identify A^1 and A^2 . The points S_{1i} and S_{1j} of equal heights are located on the shoulder of the peak from S_1 , likewise points S_{2e} and S_{2f} on the peak of S_2 . Consider the

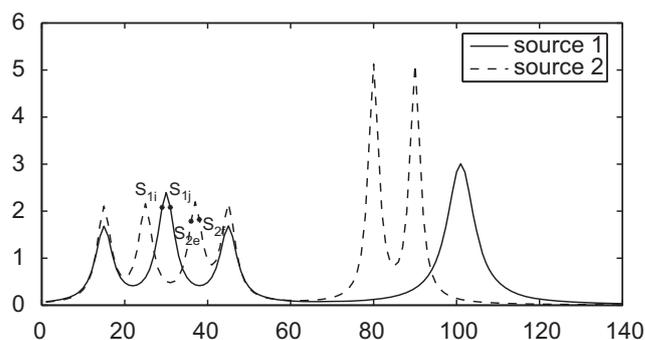


Fig. 4.1. The indices labelled in the NN recovered two sources are used for improving the 2×2 mixing matrix.

submatrix $[X^i, X^j, X^e, X^f]$ of X which is equal to

$$[A^1, A^2] \times \begin{pmatrix} S_{1i} & S_{1j} & S_{1e} & S_{1f} \\ S_{2i} & S_{2j} & S_{2e} & S_{2f} \end{pmatrix}. \quad (4.2)$$

The following equations are obtained in component form:

$$\begin{cases} X^i = S_{1i}A^1 + S_{2i}A^2 \\ X^j = S_{1j}A^1 + S_{2j}A^2 \\ X^e = S_{1e}A^1 + S_{2e}A^2 \\ X^f = S_{1f}A^1 + S_{2f}A^2 \end{cases} \quad \text{with } S_{1i} = S_{1j}, S_{2e} = S_{2f}. \quad (4.3)$$

It follows from Eq. (4.3) that A^1 is parallel to $X^m - X^n$ and A^2 parallel to $X^j - X^i$ since

$$X^e - X^f = (S_{1e} - S_{1f})A^1,$$

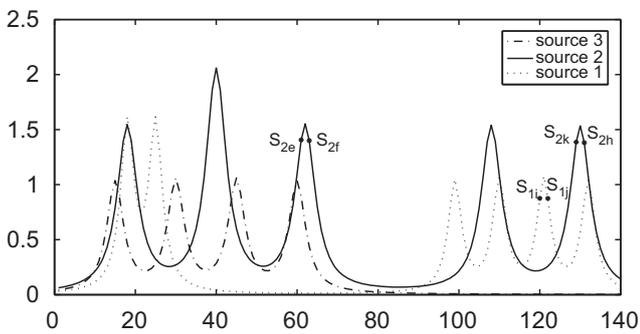


Fig. 4.2. The indices labelled in the NN recovered three sources are used for improving the 3×3 mixing matrix.

$$X^i - X^j = (S_{2i} - S_{2j})A^2,$$

and both $S_{1e} - S_{1f}$ and $S_{2i} - S_{2j}$ are positive, which can be seen from the plot. Therefore, the matrix $[X^m - X^n, X^j - X^i]$ should improve the estimation of A as numerical results will show in the next section.

Remark 1. The selection of PBC peaks depends on the NN outputs, and there may be more than one group of peaks for PBC. In this case, each estimated mixing matrix is normalized along column and the final estimation will be replaced by their average, which is a possible way to make the method more robust.

4.2. $(m, n) \geq (3, 3)$

First we shall demonstrate the method using an example of three mixtures and three sources. The sources are shown in Fig. 4.2. The mixing equation reads

$$[X^1, \dots, X^p] = [A^1, A^2, A^3] \times \begin{pmatrix} S_{11} & \dots & S_{1p} \\ S_{21} & \dots & S_{2p} \\ S_{31} & \dots & S_{3p} \end{pmatrix}.$$

Given the NN output where the peak locations are correctly recovered, first select the indices i and j which are from the level line of the third peak of S_1 , so $S_{1i} = S_{1j}$. Then the i th and j th columns of X are

$$X^i = S_{1i}A^1 + S_{2i}A^2, \quad (4.4)$$

$$X^j = S_{1j}A^1 + S_{2j}A^2, \quad (4.5)$$

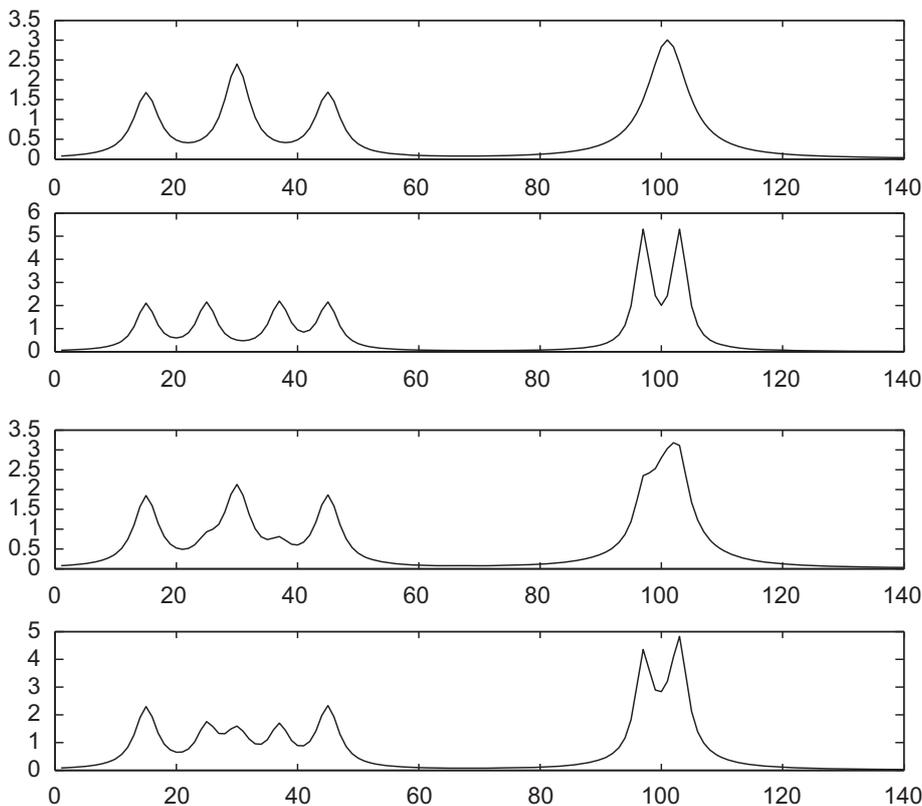


Fig. 5.1. Top two rows are the sources, bottom two rows are the mixtures.

where S_3 's contribution is ignored due to pairwise overlap assumption. Subtraction of the two equations recovers A^2 . Similarly, we can recover A^1 by looking at the two indices e and f with $S_{2e} = S_{2f}$. With S_3 being ignored, $X^e = S_{1e} A^1 + S_{2e} A^2$, $X^f = S_{1f} A^1 + S_{2f} A^2$. The same process can be used for the recovery of A^3 . The peak-based correction method can be generalized to n sources and n measurements. Under the POC condition, we have the following result:

Theorem (Peak correction). Consider $n \geq 3$ POC sources. Suppose that for any $k \in [1, n]$, $\exists j \neq k, j \in [1, n]$ such that $S_{j,m_1} = S_{j,m_2} \neq 0$ for some $m_1 < m_2$, and that $S_{k,m_1} \neq S_{k,m_2}$, where S is the NN output from data matrix X . Moreover, there exists $m_3 \in (m_1, m_2)$ such that $S_{j,m_3} = \max_{m \in [m_1, m_2]} S_{j,m}$. Then all column vectors of the mixing matrix A can be recovered from data matrix X by the peak correction procedure.

Proof. The local maximum inside $[m_1, m_2]$ implies that there is a reliable peak in the interval $[m_1, m_2]$. By the postprocessing assumption that peaks in NN output are reliable, we have the following equations:

$$X_{m_1} = S_{j,m_1} A^j + S_{k,m_1} A^k,$$

$$X_{m_2} = S_{j,m_2} A^j + S_{k,m_2} A^k$$

implying A^k is recovered up to scaling by differencing

$$X_{m_1} - X_{m_2} = (S_{k,m_1} - S_{k,m_2}) A^k. \quad \square$$

5. Numerical experiments

We report here the numerical results of the methods proposed in the last section. The data are either synthetic mixtures or real-world NMR spectra. In both cases the mixing coefficients and the matrix S are positive.

5.1. PBC method

To test the performance of PBC method, two synthetic examples and one real-world example are used where POC condition is valid. The synthetic mixtures are constructed as follows. First, the source spectra are composed of Lorentzian-shaped peaks which could be viewed as idealized NMR spectra, then the measured signals are constructed by the mixing model $X=AS$. In the second example of real-world data, we shall separate the NMR signals of camphor and quinine from their mixtures.

The first example is two mixtures and two sources (Fig. 5.1). The NN method is employed to obtain approximate separations with the results in Fig. 5.2. Following the PBC procedure described in Section 3, we first seek indices of reliable peaks for the estimation of A . The recovered A_{NN} , A_{PBC} and the true mixing matrix A are compared below (For the purpose of comparison, first rows of A_{NN} , A_{PBC} are scaled to be same as that of A). The plot of Fig. 5.3 shows the PBC results, which agree with the true sources very well due to

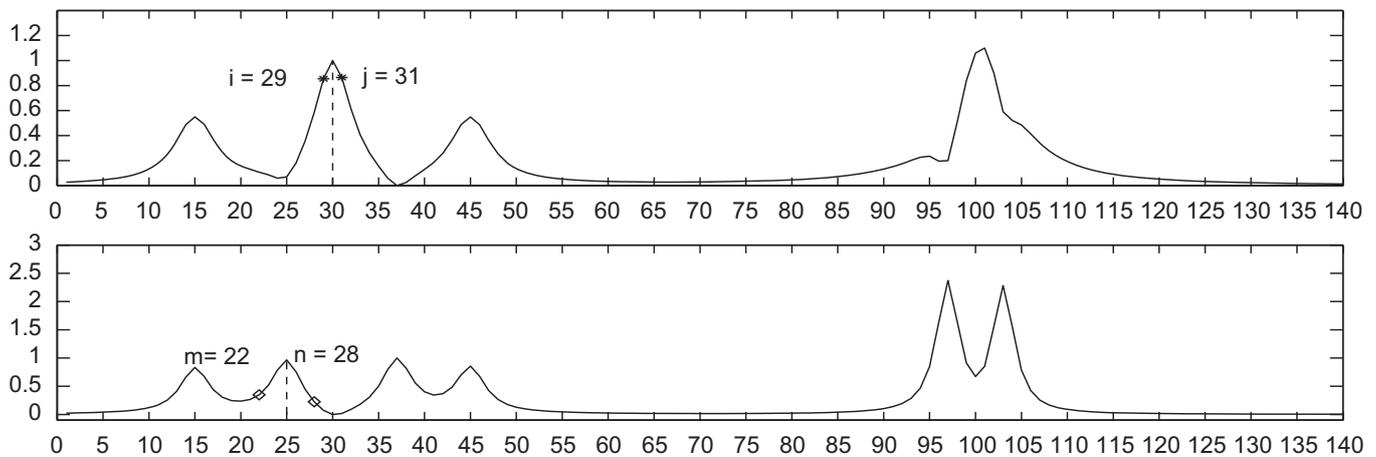


Fig. 5.2. Recovered source signals by NN method, where the marked indices are used for PBC.

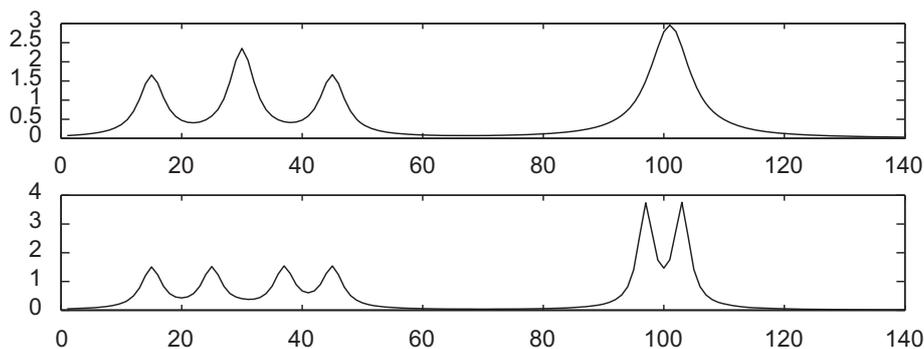


Fig. 5.3. Recovery after PBC processing.

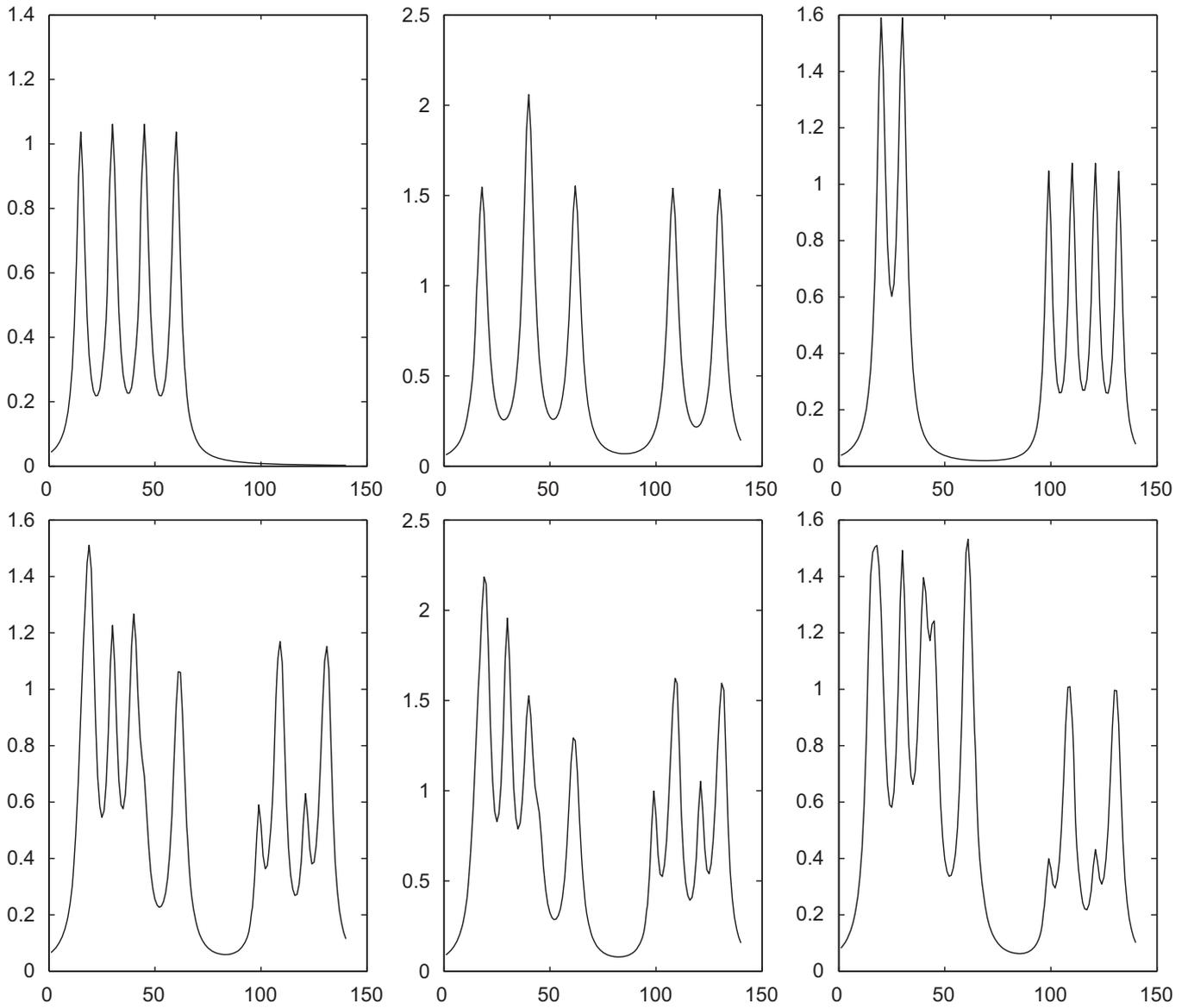


Fig. 5.4. Top two rows are the three sources, bottom two rows are the three mixtures.

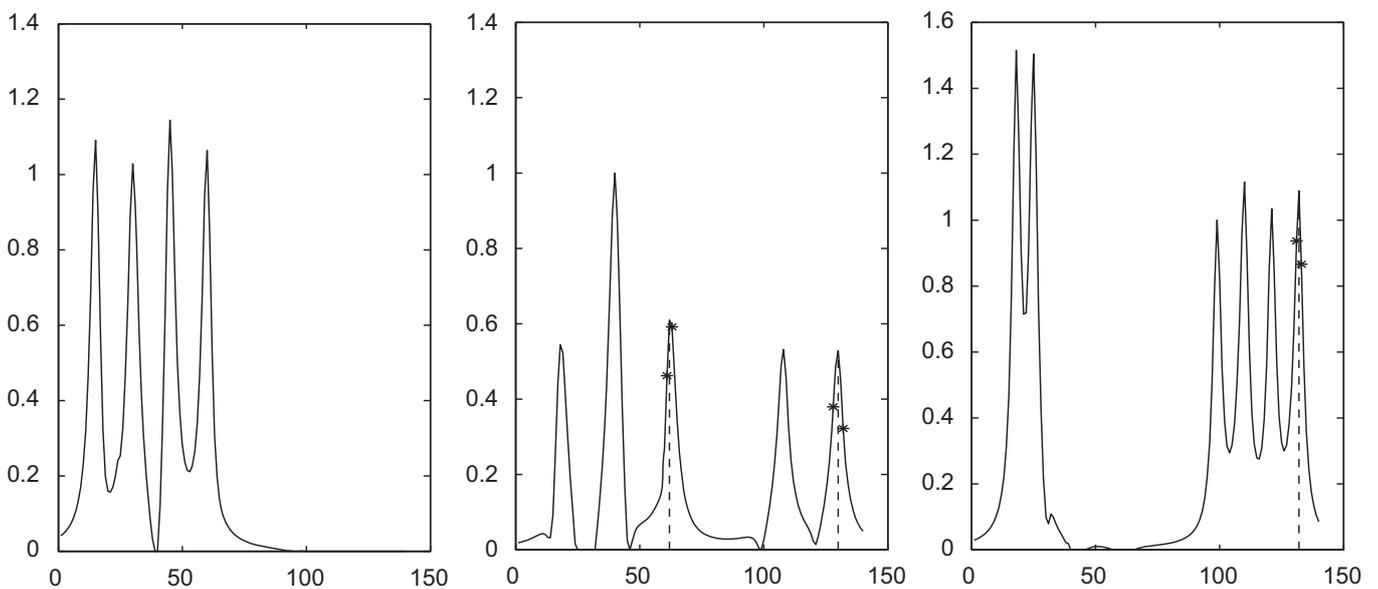


Fig. 5.5. Recovered source signals by NN method, where the marked indices are used for PBC.

an almost perfect recovery of mixing matrix:

$$A = \begin{pmatrix} 0.7142 & 0.4172 \\ 0.7000 & 0.9088 \end{pmatrix}, \quad A_{\text{NN}} = \begin{pmatrix} 0.4172 & 0.7142 \\ 0.7553 & 0.7719 \end{pmatrix},$$

$$A_{\text{PBC}} = \begin{pmatrix} 0.7142 & 0.4172 \\ 0.6886 & 0.9139 \end{pmatrix}.$$

The second example has three mixtures and three sources (see Fig. 5.4). The quality of the separation by PBC is clearly apparent by comparison of the spectra in Figs. 5.5 and 5.6.

To test the performance of PBC method, we compute its Comon's index [7]. 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} denoted by

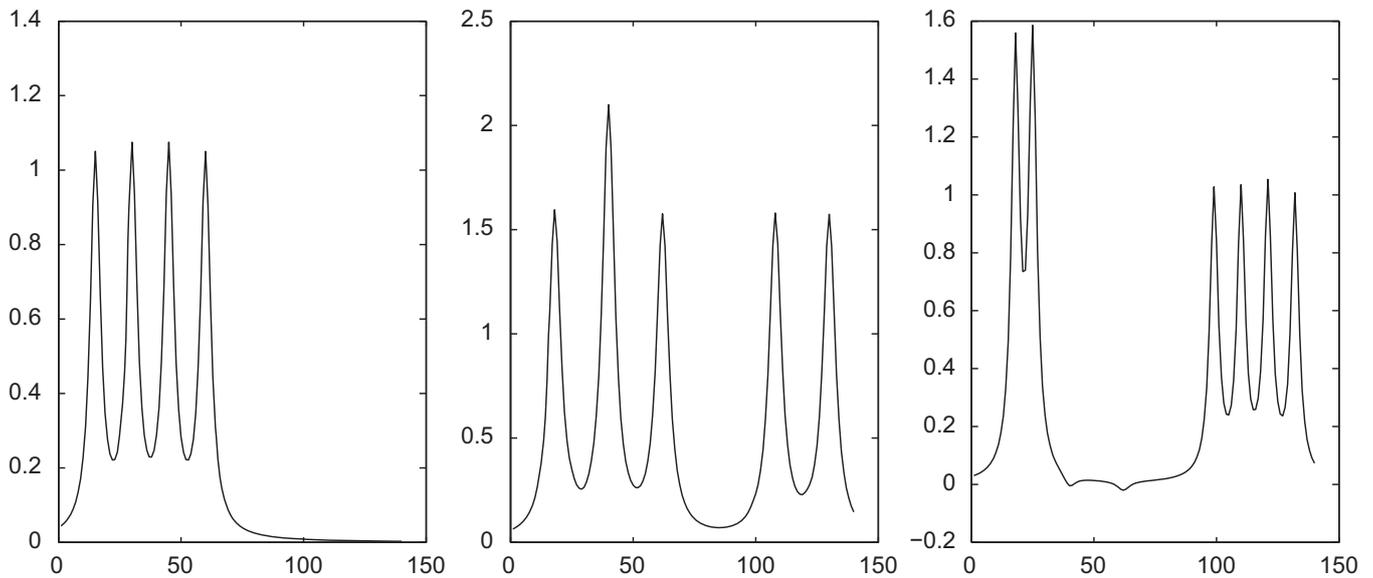


Fig. 5.6. Recovery results after PBC processing.

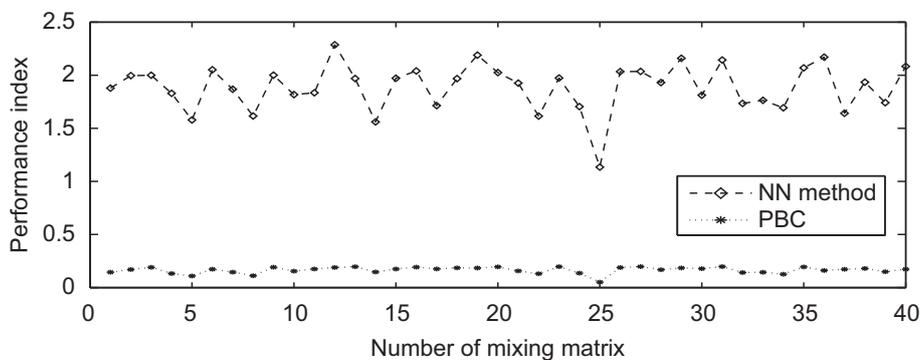


Fig. 5.7. Performance of NN and PBC method on random mixing matrix.

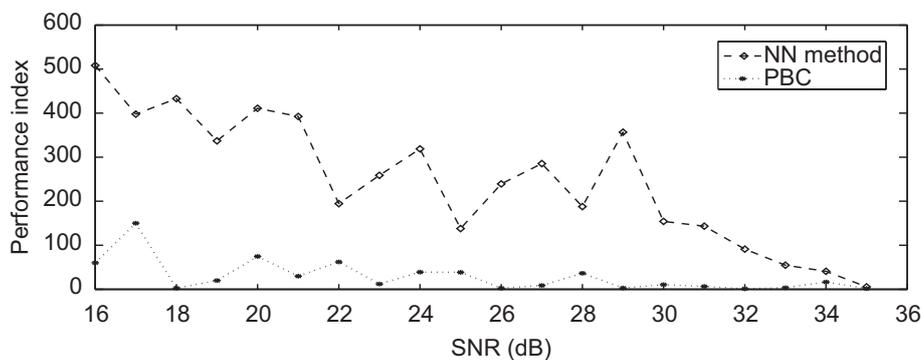


Fig. 5.8. Test of the reliability of NN and PBC method with the presence of noise.

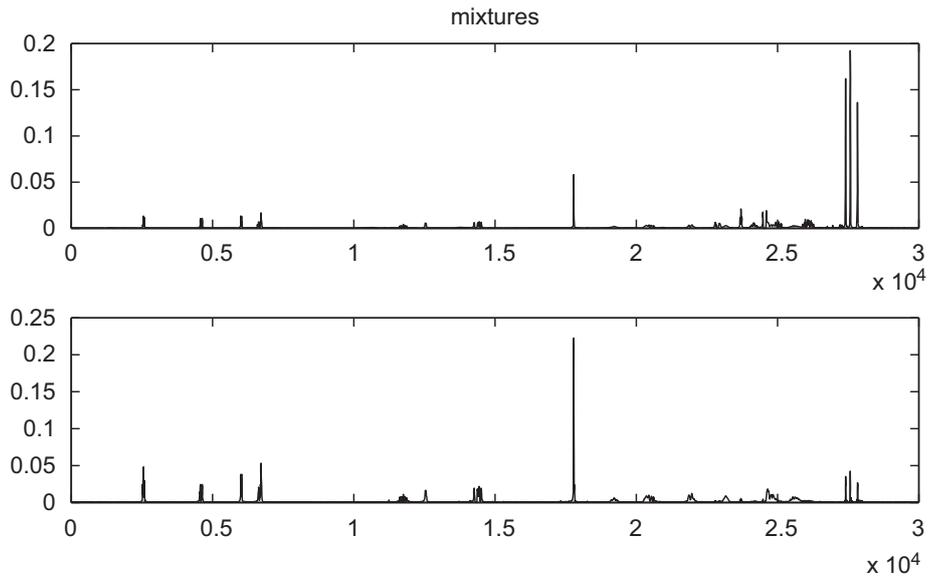


Fig. 5.9. Mixtures of camphor and quinine.

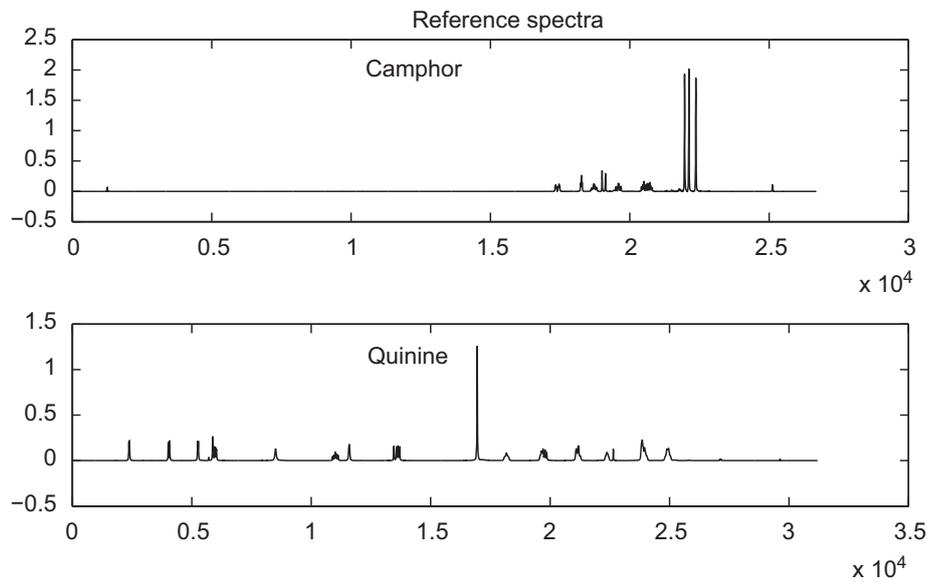


Fig. 5.10. Reference spectra of camphor and quinine.

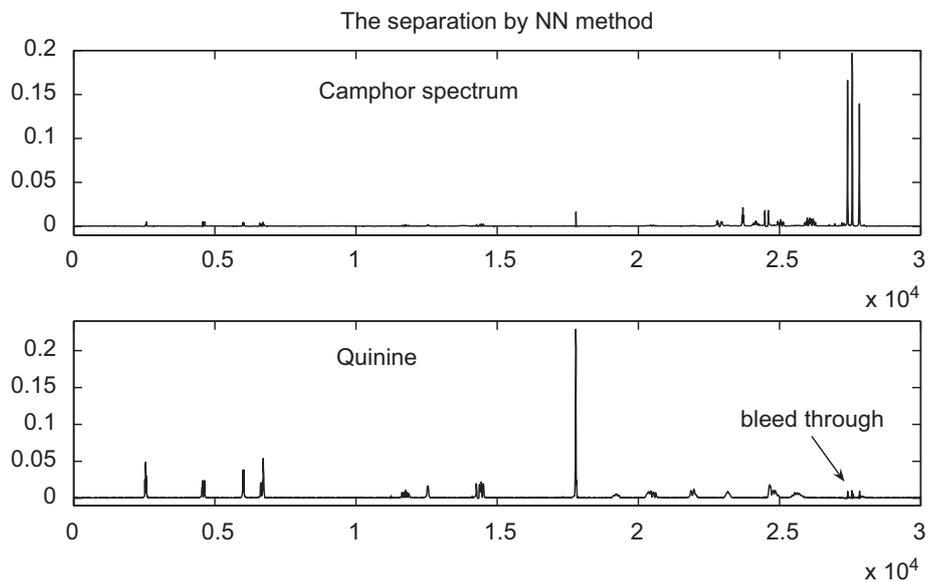


Fig. 5.11. Camphor and quinine spectra recovery by NN method. A noticeable bleed through in quinine spectrum can be seen.

$\varepsilon(A, \hat{A})$ which reads

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

where $D = A^{-1}\hat{A}$, and d_{ij} is the entry of D . In [7] Comon proved that A and \hat{A} are treated nearly equivalent in the sense of BSS (i.e., $\hat{A} = APA$) if $\varepsilon(A, \hat{A}) \approx 0$. Figs. 5.7 and 5.8 show Comon's index between the true mixing matrices and the matrices computed by NN and PBC methods. For the result in Fig. 5.7, we compute the performance indices for NN method and PBC method using 30 random mixing

matrices; while the plot of Fig. 5.8 is the comparison of the two methods with the presence of noises. Clearly a better performance is achieved in PBC method. But it should be pointed that NN's capability in extracting the major peaks is crucial to the success of PBC.

The third example concerns the blind separations of camphor and quinine signals from their mixtures (Fig. 5.9). Compared to the reference spectra (Fig. 5.10), the NN separation results (Fig. 5.11) are rather satisfactory, especially the spectrum of quinine. But we observe that there is a noticeable bleed through in its spectrum, which might come from the large peaks of camphor. By a peak-based correction, we are able to reduce the bleed through considerably. In fact, the bleed through can be barely seen in PBC recovery from Fig. 5.12.

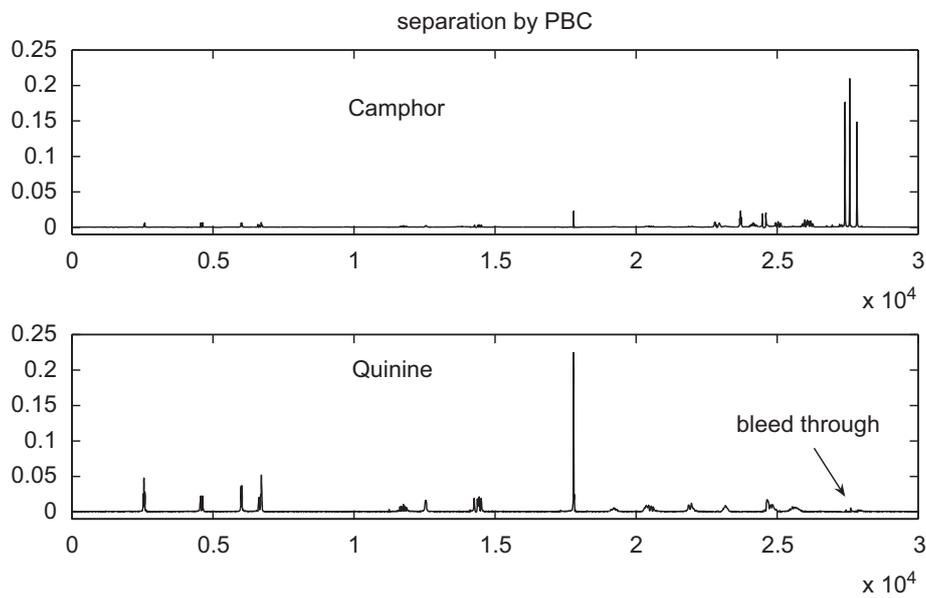


Fig. 5.12. Camphor and quinine spectra recovered by PBC. Note the bleed through is hard to see.

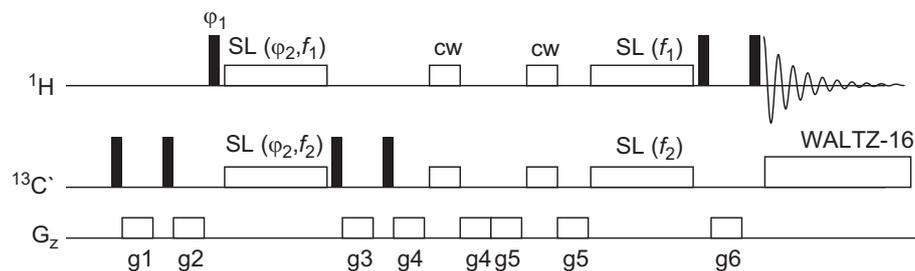


Fig. 5.13. Pulse sequence for exclusive Hartmann-Hahn magnetization transfer. Narrow filled icons are nonselective 90° pulses. Extended open rectangular icons (SL) are spin lock segments for selective magnetization transfer between carbon-13 and proton spins with chemical shifts near to the transmitter frequencies f_1 and f_2 . Each spin lock was of duration 240 ms to allow transfer through the weak 2 Hz long-range heteronuclear coupling between the acetyl group carbonyl carbon-13 and the proximal ring proton to this carbon. The first two 90° pulses and PFGs eliminate native carbon-13 magnetization. Then a selective Hartmann-Hahn magnetization transfer excites carbon-13 spins near to the chosen carbon-13 transmitter frequency, as long as the proton spin is also close to the chosen proton transmitter frequency. The magnetization is stored. The 1 ms z-gradients $g_1 - g_6$ had the following DAC values in the varian spectrometer software: 4500, 6300, 4090, 5000, 3000, and 3500. Each was followed by a stabilization delay of 300 μ s. The interscan relaxation delay was set to 4 s, and acquisition time for each FID was 0.8 s. Eight dummy scans were used to establish a steady state, and ensure accurate subtraction of signals not arising from the polarization scheme. Thirty-two transients comprised each data set, requiring 3.5 min of spectrometer time each. The phase cycle was $\phi_1 = 90^\circ, 270^\circ$ and $\phi_2 = 0^\circ, 180^\circ$. The two carbon 180° pulses were 25 Hz soft selective pulses while cw decoupling was applied to the proton methyl region (using a modest 200 Hz decoupling field) near 2.06 ppm to ensure that the small 7 Hz long-range coupling between the methyl protons and carbonyl carbon in the acetyl group is collapsed, thereby enhancing the selectivity of the excitation sculpting DPFGE sequence.

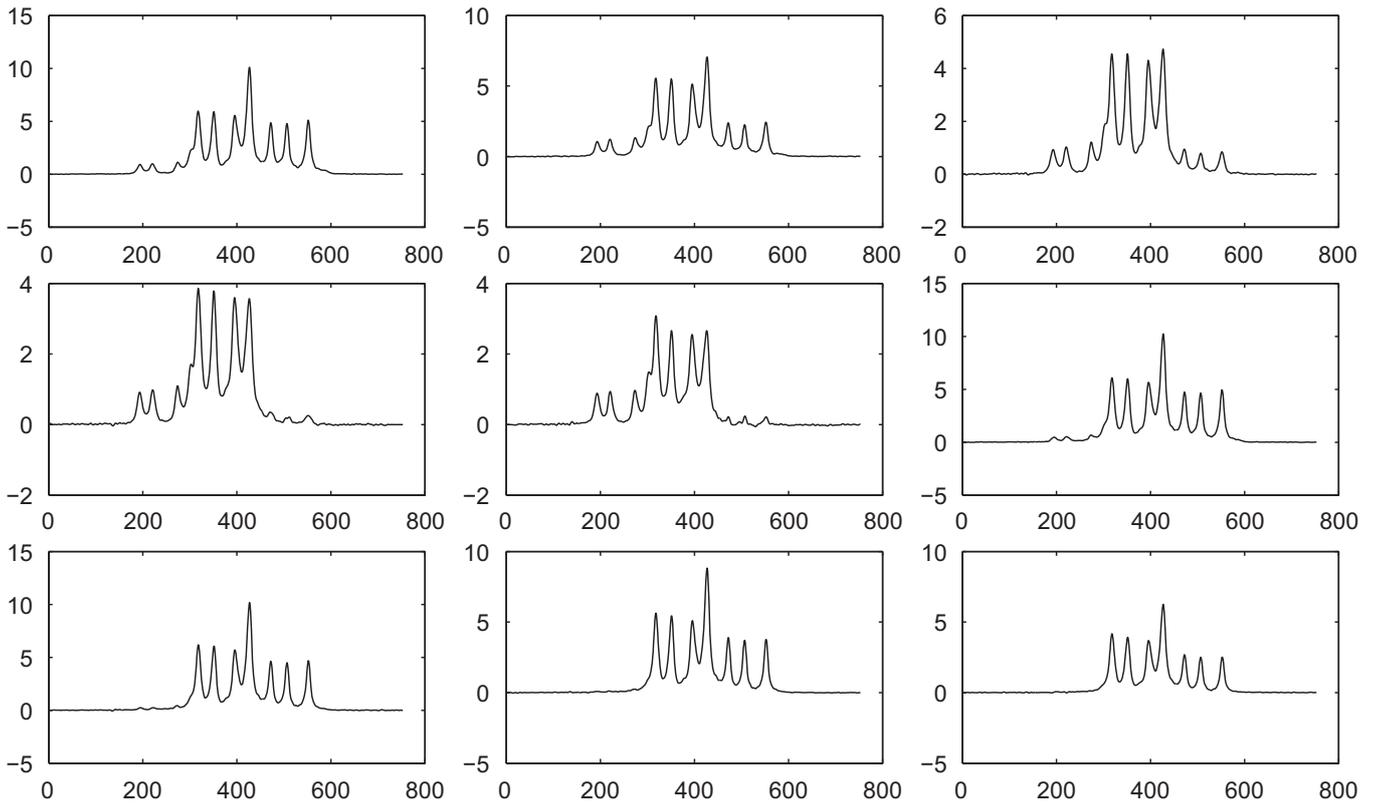


Fig. 5.14. Nine mixed NMR spectra of three sources.

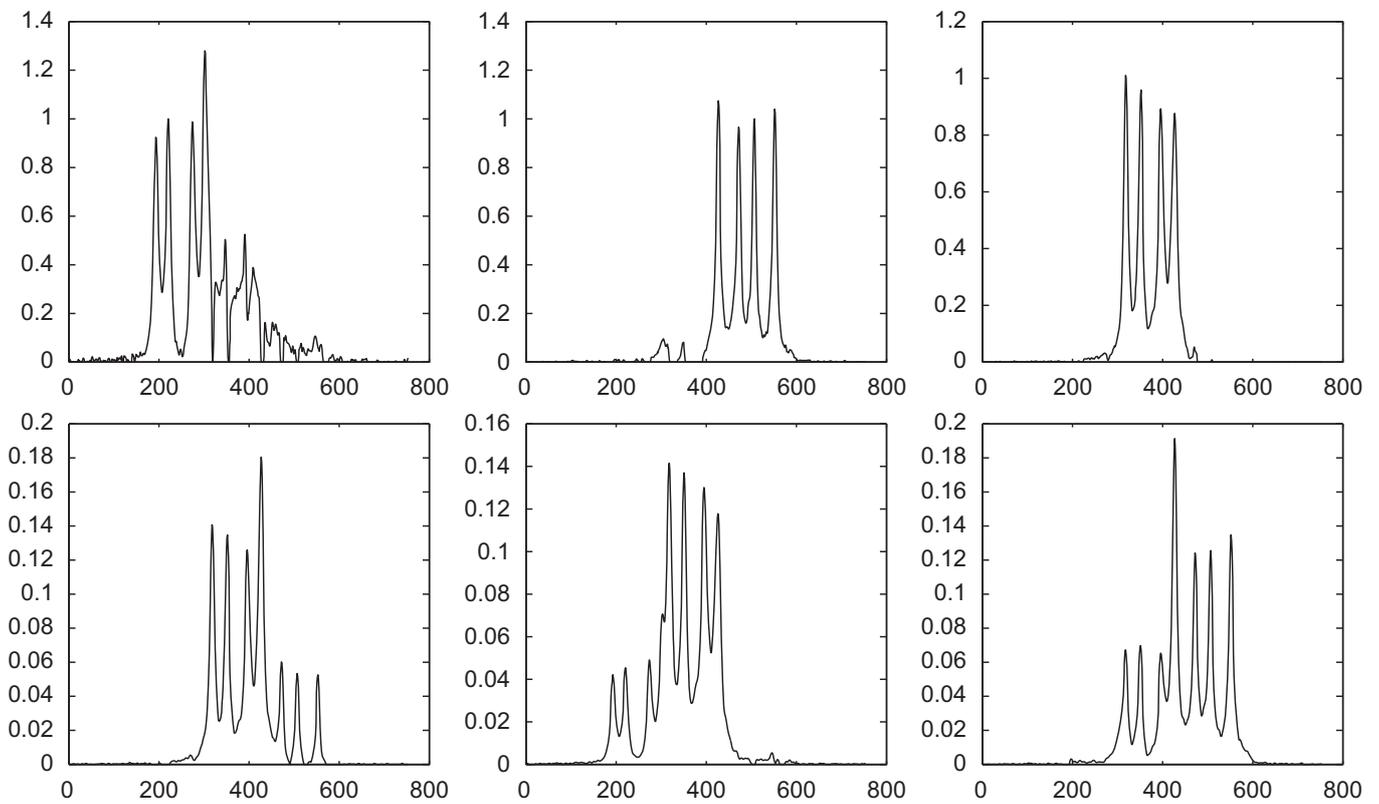


Fig. 5.15. The upper panel is the three sources recovered using nine mixtures via NN method; lower panel is the three sources recovered from nine mixtures via NMF algorithm.

5.2. Experimental NMR pulse sequence details

All spectra were obtained on a Varian UnityPlus 500 MHz NMR spectrometer using either simple pulse-acquire for

camphor and quinine, or the pulse sequence in Fig. 5.13, in which case the sample was a disaccharide sugar alcohol, maltitol, that had been “isotagged” with ^{13}C acetyl groups using a simple procedure described in the literature [23],

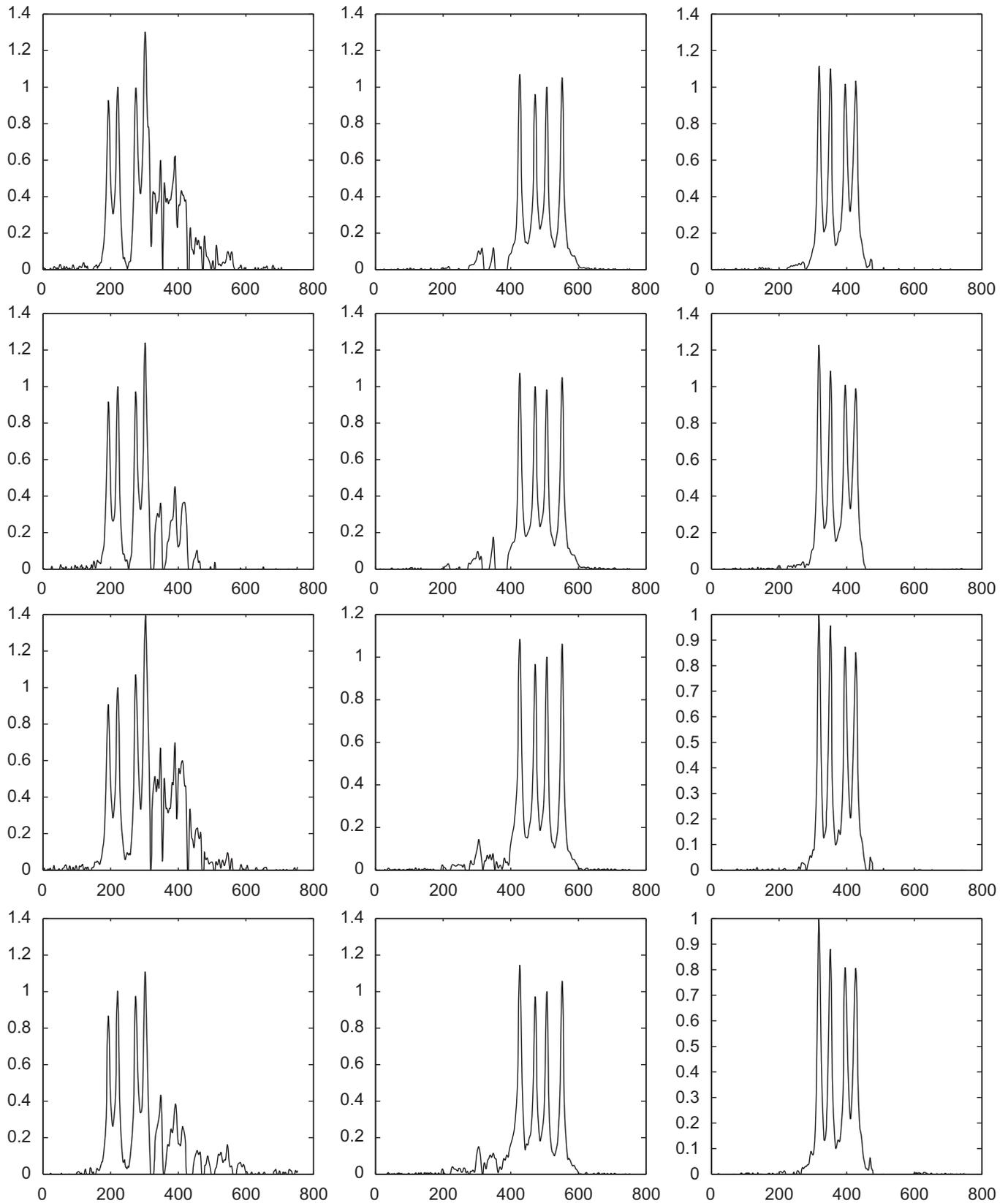


Fig. 5.16. NN's results from different groups of mixtures. First row: (2,4,7), second row: (1,5,9), third row: (3,6,8), last row: (2,6,9).

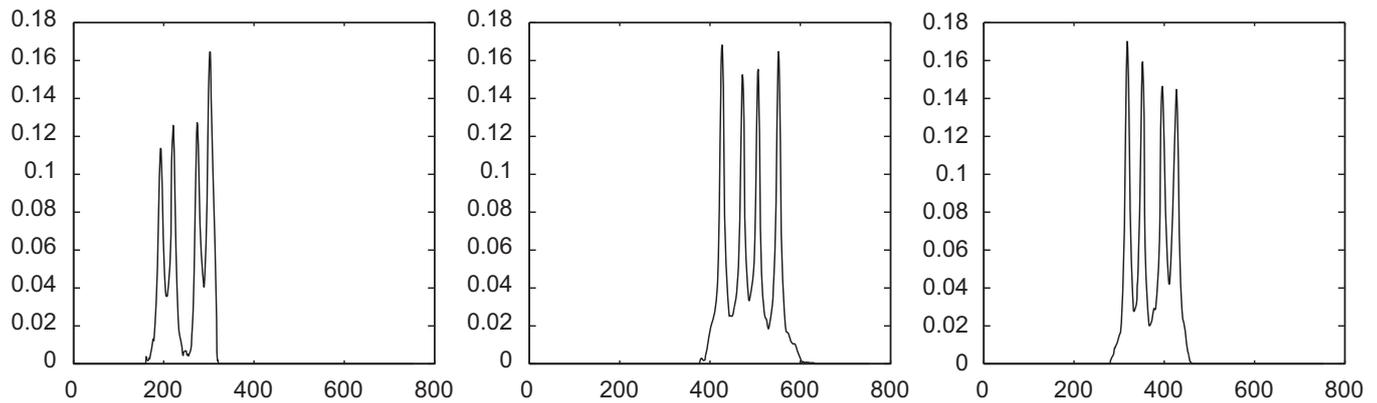


Fig. 5.17. Enhanced signals obtained by extracting the coherent parts of NN output.

employing 100% ^{13}C -1 acetic anhydride and a scandium triflate catalyst [12]. The pulse sequence of Fig. 5.13 combines two selective polarization transfer steps with an excitation sculpting double pulsed field gradient spin echo [25] to improve the selection of a particular carbonyl carbon resonance, all of which reside within a tight bandwidth of 0.5 ppm or so. For many of the sites, the proximal ring proton and the ^{13}C carbonyl carbon resonances are sufficiently isolated from other spin pairs that this exclusive heteronuclear Hartmann–Hahn pulse sequence results in just one in-phase proton spin multiplet, from which proton–proton coupling constants can be measured, giving important insight into the stereochemistry of the sugar rings. The 1D experiment uses the frequency selectivity of the transfers, and the low-power ($\gamma B_1/2\pi = 25$ Hz) 20 ms 180° carbonyl pulses used in the excitation sculpting [25] segment to try to isolate a particular proton multiplet by exciting only the proton on the molecule that is nearest to the carbonyl label whose frequency is chosen. However, in the case illustrated here, three proton multiplets with close chemical shifts also had isotag carbonyl carbons with close chemical shifts. It thus proved impossible, even with selective low-power Hartmann–Hahn magnetization transfer and selective 180° pulses to isolate a single multiplet, making the current separation of the three overlapping signals of great interest.

5.3. Error detection and reduction

Let us consider an over-determined problem of separating three source signals from nine experimental NMR mixtures in Fig. 5.14. We first apply NN method to the nine mixtures, the result is shown in Fig. 5.15. Even though there are small spurious noisy looking peaks around, the quality of separated sources 2 and 3 is good. As a comparison, we present the result of the known NMF algorithms [5] (the second plot in Fig. 5.15). Clearly the NMF result is not a good separation in that the sources are not well separated. In fact, due to the *nonconvexity* of the problem, the NMF algorithms may or may not converge to the same solutions on each run, depending on the initial conditions and the kind of the algorithm we use. To improve the separation result, we shall take advantage of multiple mixtures. Four groups of mixtures with three mixtures in each group are selected to perform BSS via

NN's method. The separation results are shown in Fig. 5.16. The selection process is then performed on the NN output to gather similar signals and plot them along columns.

It can be seen that NN method recovers the four major peaks of source 1, but more noticeable noises and artifacts are present around them. Applying the correlation-based decomposition technique, we are able to extract the coherent part from column one corresponding to source 1. The postprocessed signals are shown in Fig. 5.17, where the uncertain parts have been removed and the coherent components are retained.

6. Conclusion

This paper presented two postprocessing techniques of sparse blind source separation of positive and overlapping source data. These two approaches handle data that do not meet NN's sparseness requirements. With the NN results as inputs, the decomposition technique based on multiple mixtures extracted meaningful signal components via similarity measures. On the other hand, the peak-based correction method improves the accuracy of the mixing matrix estimation based on partial knowledge of the sources and NN's partially correct recoveries. The effectiveness is validated by NMR spectra data.

It should be noted that the regime we addressed in the paper also has analogues in hyperspectral imaging of geoscience. It happens when minerals being imaged overlap or when ground resolution is low. Dependent component analysis (DCA) is recently proposed [18] based on statistical assumptions of the sources, yet the resulting problem is nonconvex (same as ICA). It is worthwhile to pursue this line of inquiry for NMR data in the future, however, the DCA method is much more complex than ours.

Acknowledgements

This work was partially supported by NSF-ADT Grants DMS-0911277, DMS-0712881, and CHE-07013182. F. del Rio acknowledges the support of the UC MEXUS-CONACYT Visiting Scholar Fellowship programme.

References

- [1] 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, 1993, pp. 11–14.
- [2] P. Bofill, M. Zibulevsky, Underdetermined blind sources separation using sparse representations, *Signal Process.* 81 (2005) 2353–2362.
- [3] G. Brown, D.-L. Wang, Separation of speech by computational auditory scene analysis, in: J. Benesty, S. Makino, J. Chen (Eds.), *Speech Enhancement*, Springer, New York, 2005, pp. 371–402.
- [4] C.-I. Chang (Ed.), *Hyperspectral Data Exploitation: Theory and Applications*, Wiley-Interscience, 2007.
- [5] S. Choi, A. Cichocki, H. Park, S. Lee, Blind source separation and independent component analysis: a review, *Neural Inf. Process. Lett. Rev.* 6 (2005) 1–57.
- [6] A. Cichocki, S. Amari, *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*, John Wiley and Sons, New York, 2005.
- [7] P. Comon, Independent component analysis—a new concept?, *Signal Process* 36 (1994) 287–314.
- [8] P. Comon, C. Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applications*, Academic Press, 2010.
- [9] J.H. Dulà, R.V. Helgason, A new procedure for identifying the frame of the convex hull of a finite collection of points in multidimensional space, *Eur. J. Oper. Res.* 92 (1996) 352–367.
- [10] P. Georgiev, F. Theis, A. Cichocki, Sparse component analysis and blind source separation of underdetermined mixtures, *IEEE Trans. Neural Networks* 16 (4) (2005) 992–996.
- [11] X. Guo, C. Chang, Y. Lam, Blind separation of electron paramagnetic resonance signals using diversity minimization, *J. Magn. Reson.* 204 (2010) 26–36.
- [12] K. Ishihara, M. Kubota, H. Kurihara, H. Yamamoto, Scandium trifluoromethanesulfonate as an extremely active Lewis acid catalyst in acylation of alcohols with acid anhydrides and mixed anhydrides, *J. Org. Chem.* 61 (1996) 4560–4567.
- [13] J. Kolba, 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.
- [14] D. Lee, H. Seung, Learning of the parts of objects by non-negative matrix factorization, *Nature* 401 (1999) 788–791.
- [15] J. Liu, J. Xin, Y.-Y. Qi, A soft-constrained dynamic iterative method of blind source separation, *SIAM J. Multiscale Modeling Simulations* 7 (4) (2009) 1795–1810.
- [16] W. Naanaa, J.-M. Nuzillard, Blind source separation of positive and partially correlated data, *Signal Process.* 85 (9) (2005) 1711–1722.
- [17] M. Naceur, M. Loghmari, M. Boussema, The contribution of the sources separation method in the decomposition of mixed pixels, *IEEE Trans. Geosci. Remote Sensing* 42 (2004) 2642–2653.
- [18] J. Nascimento, J. Bioucas-Dias, Unmixing hyperspectral data: independent and dependent component analysis, in: C.-I. Chang (Ed.), *Hyperspectral Data Exploitation: Theory and Applications*, Wiley-Interscience, 2007, pp. 149–178.
- [19] D. Nuzillard, A. Bijaoui, Blind source separation and analysis of multispectral astronomical images, *Astron. Astrophys. Suppl. Ser.* 147 (2000) 129–138.
- [20] D. Nuzillard, S. Bourgb, J.-M. Nuzillard, Model-free analysis of mixtures by NMR using blind source separation, *J. Magn. Reson.* 133 (1998) 358–363.
- [21] M. Plumbley, Conditions for non-negative independent component analysis, *IEEE Signal Process. Lett.* 9 (2002) 177–180.
- [22] M. Plumbley, Algorithms for nonnegative independent component analysis, *IEEE Trans. Neural Networks* 4 (3) (2003) 534–543.
- [23] X. Meng, W.H. Nguyen, J.S. Nowick, A.J. Shaka, Selective heteronuclear Hartmann–Hahn: a multiple-pulse sequence for selective magnetization transfer in the structural elucidation of “isotagged” oligosaccharides, *J. Magn. Reson.* 203 (2010) 73–80.
- [24] K. Stadthanner, A. Tom, F. Theis, W. Gronwald, H.-R. Kalbitzer, E. Lang, On the use of independent analysis to remove water artifacts of 2D NMR protein spectra, in: *Proceedings of Bioeng'2003*, 2003.
- [25] K. Stott, J. Stonehouse, J. Keeler, T.L. Hwang, A.J. Shaka, Excitation Sculpting in high-resolution nuclear magnetic resonance spectroscopy: application to selective NOE experiments, *J. Am. Chem. Soc.* 117 (14) (1995) 4199–4200.
- [26] M.E. Winter, N-findr: an algorithm for fast autonomous spectral end-member determination in hyperspectral data, in: *Proceedings of the SPIE*, vol. 3753, 1999, pp. 266–275.