Nonnegative Compression for Semi-Nonnegative Independent Component Analysis

Lu Wang, Amar Kachenoura, Laurent Albera, Ahmad Karfoul, Hua Zhong Shu, Lotfi Senhadji

Nonnegative Compression for Semi-Nonnegative Independent Component Analysis

Lu Wang*†§, Amar Kachenoura*†§, Laurent Albera*†§¶, Ahmad Karfoul‖, Hua Zhong Shu‡§ and Lotfi Senhadji*†§

*INSERM, UMR 1099, Rennes, F-35000, France
†Université de Rennes 1, LTSI, Rennes, F-35000, France
‡LIST, Southeast University, 2 Sipailou, 210096, Nanjing, China
§Centre de Recherche en Information Biomédicale sino-français (CRIBs), Rennes, France
¶INRIA, Centre Inria Rennes - Bretagne Atlantique, 35042 Rennes, France.
‖AL-Baath University, Mechanical and Electrical Engineering, PB. 2244, Homs, Syria.
Email: wangLyu1986@hotmail.com

Abstract—**In many Independent Component Analysis (ICA) problems the mixing matrix is nonnegative while the sources are unconstrained, giving rise to what we call hereafter the Semi-Nonnegative ICA (SeNICA) problems. Exploiting the nonnegativity property can improve the ICA result. Besides, in some practical applications, the dimension of the observation space must be reduced. However, the classical dimension compression procedure, such as prewhitening, breaks the nonnegativity property of the compressed mixing matrix. In this paper, we introduce a new nonnegative compression method, which guarantees the nonnegativity of the compressed mixing matrix. Simulation results show its fast convergence property. An illustration of Blind Source Separation (BSS) of Magnetic Resonance Spectroscopy (MRS) data confirms the validity of the proposed method.**

Index Terms—**Blind source separation, independent component analysis, nonnegative compression.**

I. INTRODUCTION AND PROBLEM FORMULATION

The Semi-Nonnegative Independent Component Analysis (SeNICA) problem is defined as follows [1]–[3]:

Problem 1. *Given M realizations of a real N-dimensional random vector $\boldsymbol{x}$, find an $(N \times P)$ full column rank mixing matrix $\boldsymbol{A}$ and the corresponding M realizations of P-dimensional source random vector $\boldsymbol{s}$, such that:*

$$\boldsymbol{x} = \boldsymbol{A}\boldsymbol{s} \tag{1}$$

where $\boldsymbol{A}$ has nonnegative components and $\boldsymbol{s}$ has statistically independent components.

This problem is encountered in many Blind Source Separation (BSS) applications. For example, in Magnetic Resonance Spectroscopy (MRS), the mixing matrix contains the positive concentrations of the source metabolites, while the source spectra are not necessarily nonnegative (they can become nonnegative only after a complicated phase shift procedure).

In many ICA algorithms, in order to make the mixing matrix orthogonal but also to reduce the dimension of the observation space, the N-dimensional vector $\boldsymbol{x}$ is compressed into a vector $\bar{\boldsymbol{x}}$ of dimension $P \leq N$ as follows [4], [5]:

$$\bar{\boldsymbol{x}} = \boldsymbol{W}^\mathsf{T}\boldsymbol{x} = (\boldsymbol{W}^\mathsf{T}\boldsymbol{A})\boldsymbol{s} = \overline{\boldsymbol{A}}\boldsymbol{s} \tag{2}$$

where the columns of $\boldsymbol{W} \in \mathbb{R}^{N \times P}$ are the scaled eigenvectors corresponding to the P largest eigenvalues of the covariance matrix $\mathrm{E}[\boldsymbol{x}\boldsymbol{x}^\mathsf{T}] - \mathrm{E}[\boldsymbol{x}]\mathrm{E}[\boldsymbol{x}^\mathsf{T}]$. This procedure is known as the prewhitening of $\boldsymbol{x}$ and $\boldsymbol{W}^\mathsf{T}$ is called prewhitening matrix. However, such a compression method can not guarantee the nonnegativity of the compressed mixing matrix $\overline{\boldsymbol{A}} = \boldsymbol{W}^\mathsf{T}\boldsymbol{A}$, because generally the unitary matrix $\boldsymbol{W}$ is not nonnegative. Now, in some practical situations, exploiting the nonnegativity property of the mixing matrix can improve the ICA result [1]–[3], [6]–[8]. Therefore, the objective of this paper is to propose a nonnegative compression method that guarantees the nonnegativity of the compressed mixing matrix $\overline{\boldsymbol{A}}$. To this end, note that $\mathrm{span}(\boldsymbol{W}) = \mathrm{span}(\boldsymbol{A})$, where the columns of $\boldsymbol{A}$ form a basis of nonnegative vectors of the signal subspace. Hence, there exists a nonsingular matrix $\boldsymbol{B} \in \mathbb{R}^{P \times P}$ such that the columns of matrix $\overline{\boldsymbol{W}} = \boldsymbol{W}\boldsymbol{B}$ form a basis of nonnegative vectors of the signal subspace $\mathrm{span}(\boldsymbol{A})$. It means that we can linearly transform the unitary matrix $\boldsymbol{W}$ into a nonnegative matrix $\overline{\boldsymbol{W}}$ multiplying it by $\boldsymbol{B}$ on its right. Now, since $\boldsymbol{B}$ is full rank, it admits the following QR matrix factorization:

$$\boldsymbol{B} = \boldsymbol{Q}\boldsymbol{R}\boldsymbol{\Lambda} \tag{3}$$

where $\boldsymbol{Q} \in \mathbb{R}^{P \times P}$ is a unitary matrix, where $\boldsymbol{R} \in \mathbb{R}^{P \times P}$ is a (unit) upper triangular matrix whose diagonal elements are equal to one and where $\boldsymbol{\Lambda} \in \mathbb{R}^{P \times P}$ is a nonsingular nonnegative diagonal matrix. Due to the fact that a multiplication of $\boldsymbol{W}\boldsymbol{Q}\boldsymbol{R}$ by $\boldsymbol{\Lambda}$ does not change the subspace $\mathrm{span}(\boldsymbol{W}\boldsymbol{Q}\boldsymbol{R})$ or the sign of the of the column vectors of $\mathrm{span}(\boldsymbol{W}\boldsymbol{Q}\boldsymbol{R})$, looking for $\boldsymbol{B}$ of the form $\boldsymbol{B} = \boldsymbol{Q}\boldsymbol{R}$ is sufficient. On the other hand, every unitary matrix can be factorized as a product of Givens matrices.

Definition 1. *A Givens matrix $\boldsymbol{Q}^{(i,j)}(\theta_{i,j})$, with $i < j$, is equal to an identity matrix except the (i,i)-th, (j,j)-th, (i,j)-th and (j,i)-th entries, which are equal to $\cos(\theta_{i,j})$, $\cos(\theta_{i,j})$, $-\sin(\theta_{i,j})$ and $\sin(\theta_{i,j})$, respectively.*

In addition, every unit upper triangular matrix can be factorized as a product of elementary upper triangular matrices.
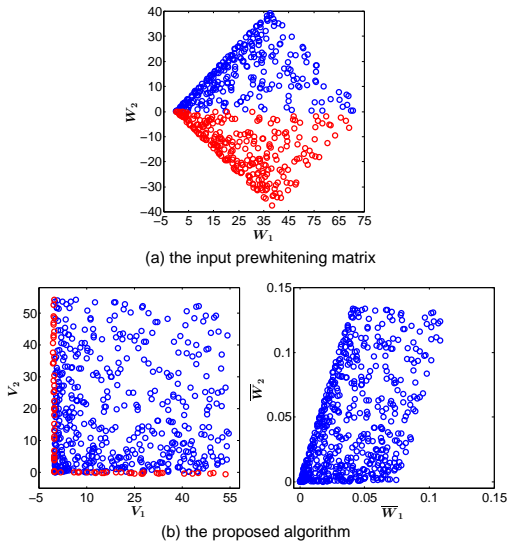
Fig. 1. An illustration of the proposed algorithm. (a) the input prewhitening matrix $\boldsymbol{W}$; (b) the proposed algorithm, left: $\boldsymbol{V}$ is obtained by multiplying $\boldsymbol{W}$ by a Givens rotation matrix, right: $\overline{\boldsymbol{W}}$ is obtained by multiplying $\boldsymbol{V}$ by an elementary upper triangular matrix. The blue circles denote the nonnegative values and the red circles denote the negative values.

Definition 2. *An elementary upper triangular matrix $\boldsymbol{R}^{(i,j)}(u_{i,j})$, with $i < j$, is equal to an identity matrix except the (i,j)-th entry, which is equal to $u_{i,j}$.*

Note that $\boldsymbol{R}^{(i,j)}(u_{i,j})$ is also known as the shear matrix where $u_{i,j}$ is called a shear factor.

Then the nonnegative compression problem is defined as follows:

Problem 2. *Given a prewhitening matrix $\boldsymbol{W} \in \mathbb{R}^{N \times P}$ of an N-dimensional random vector $\boldsymbol{x}$, find a sequence of Givens rotation matrices $\boldsymbol{Q}^{(i,j)}(\theta_{i,j})$ and elementary upper triangular matrices $\boldsymbol{R}^{(i,j)}(u_{i,j})$, such that the following matrix:*

$$\overline{\boldsymbol{W}} = \underbrace{\boldsymbol{W} \prod_{i=1}^{P} \prod_{j=i+1}^{P} \boldsymbol{Q}^{(i,j)}(\theta_{i,j})}_{\overset{def}{=} \boldsymbol{V}} \prod_{i=1}^{P} \prod_{j=i+1}^{P} \boldsymbol{R}^{(i,j)}(u_{i,j}) \quad (4)$$

has nonnegative components.

Figure 1 illustrates, in the case of $P = 2$, the process of transforming the matrix $\boldsymbol{W}$ into a nonnegative one using equation (4). Prewhitening makes the axes of the matrix $\boldsymbol{W}$ orthogonal to each other (figure 1(a)). A Givens matrix searches for a rotation angle that makes the outputs matrix $\boldsymbol{V}$ as nonnegative as possible. It is noteworthy that if the column vectors of $\boldsymbol{W}$ is statistically independent and well-grounded, only the Givens transformation is sufficient to give a nonnegative $\boldsymbol{V}$ [9]. However, in practice, generally it is not the case. As we can see in figure 1(b) left, it still remains some negative values near the positive orthant boundaries. Then we use an elementary upper triangular matrix which projects the remaining negative values of $\boldsymbol{V}$ into the positive orthant (figure 1(b) right). We thus obtain a nonnegative compression

matrix $\overline{\boldsymbol{W}}$, which preserves the nonnegativity property of $\boldsymbol{A}$. Indeed, the $(P \times P)$ compressed mixing matrix $\overline{\boldsymbol{A}} = \overline{\boldsymbol{W}}^{\top} \boldsymbol{A}$ is yet nonnegative. Then $\overline{\boldsymbol{A}}$ and the M realizations of $\boldsymbol{s}$ can be estimated using a SeNICA method such as one of those described in [1]–[3], [6], [7]. Therefore it is not necessary to estimate the original mixing matrix $\boldsymbol{A}$. Now the challenge is how to compute the Givens and elementary upper triangular matrices.

II. METHOD

A new two-step NonNegative COMPression method (figure 1), called NN-COMP, is presented in this section.

A. Step 1: estimation of $\boldsymbol{Q}^{(i,j)}(\theta_{i,j})$

The Givens rotation matrix $\boldsymbol{Q}^{(i,j)}(\theta_{i,j})$ transforms each pair of columns (i,j) of $\boldsymbol{W}$ as follows:

$$[\boldsymbol{v}_i, \boldsymbol{v}_j] = [\boldsymbol{w}_i, \boldsymbol{w}_j] \begin{bmatrix} \cos(\theta_{i,j}) & -\sin(\theta_{i,j}) \\ \sin(\theta_{i,j}) & \cos(\theta_{i,j}) \end{bmatrix} \quad (5)$$

where $\boldsymbol{v}_i$ and $\boldsymbol{w}_i$ are the i-th columns of $\boldsymbol{V}$ and $\boldsymbol{W}$, respectively. Now, let us consider the following negativity measure criterion defined in [9] and [10]:

$$J(\theta_{i,j}) = \frac{1}{2} \sum_{n=1}^{N} (V_{n,i}^2 \mathbb{1}_{V_{n,i}<0} + V_{n,j}^2 \mathbb{1}_{V_{n,j}<0}) \quad (6)$$

where $\mathbb{1}_{\alpha<0} = \begin{cases} 1, & \text{if } \alpha < 0 \\ 0, & \text{otherwise} \end{cases}$, $v_{n,i}$ and $v_{n,j}$ are the (n,i)-th and (n,j)-th elements of $\boldsymbol{V}$, respectively. The purpose is to find an angle, $\theta_{i,j}$, which minimizes the total sum of squares of negative elements of $\boldsymbol{V}$. The global optimum of $J(\theta_{i,j})$ is difficult to obtain analytically due to the existence of the Heaviside-step-like function $\mathbb{1}_{\alpha<0}$. We propose to compute $\theta_{i,j}$ iteratively by a Newton's method. For a given iteration (it), let us consider the second order Taylor expansion, $J_T(\theta_{i,j})$, of $J(\theta_{i,j})$ around $\theta_{i,j}^{(it)}$:

$$J_T(\theta_{i,j}^{(it)}+\Delta\theta) = J(\theta_{i,j}^{(it)}) + \frac{dJ(\theta_{i,j}^{(it)})}{d\theta_{i,j}^{(it)}}\Delta\theta + \frac{d^2 J(\theta_{i,j}^{(it)})}{2d(\theta_{i,j}^{(it)})^2}(\Delta\theta)^2 \quad (7)$$

where $\theta_{i,j}^{(it)}$ is the solution at the it-th iteration, $\Delta\theta = \theta_{i,j} - \theta_{i,j}^{(it)}$, $dJ(\theta_{i,j}^{(it)})/d\theta_{i,j}^{(it)}$ and $d^2 J(\theta_{i,j}^{(it)})/d(\theta_{i,j}^{(it)})^2$ are the first and second order derivatives of (6) with respect to (w.r.t.) $\theta_{i,j}^{(it)}$, respectively, which are given by [10]:

$$dJ(\theta_{i,j}^{(it)})/d\theta_{i,j}^{(it)} = \sum_{n=1}^{N} V_{n,i}^{(it)} v_{n,j}^{(it)} \times$$
$$(\mathbb{1}_{v_{n,i}^{(it)}<0} \mathbb{1}_{v_{n,j}^{(it)}>0} - \mathbb{1}_{v_{n,i}^{(it)}>0} \mathbb{1}_{v_{n,j}^{(it)}<0}) \quad (8)$$

and:

$$d^2 J(\theta_{i,j}^{(it)})/d(\theta_{i,j}^{(it)})^2 = \sum_{n=1}^{N} \left((v_{n,j}^{(it)})^2 - (v_{n,i}^{(it)})^2 \right) \times$$
$$(\mathbb{1}_{v_{n,i}^{(it)}<0} \mathbb{1}_{v_{n,j}^{(it)}>0} - \mathbb{1}_{v_{n,i}^{(it)}>0} \mathbb{1}_{v_{n,j}^{(it)}<0}) \quad (9)$$

with:

$$[v_{n,i}^{(it)}, v_{n,j}^{(it)}] = [w_{n,i}, w_{n,j}] \begin{bmatrix} \cos(\theta_{i,j}^{(it)}) & -\sin(\theta_{i,j}^{(it)}) \\ \sin(\theta_{i,j}^{(it)}) & \cos(\theta_{i,j}^{(it)}) \end{bmatrix} \quad (10)$$

Then the minimum of (7) w.r.t. $\Delta\theta$ can be reached at:

$$\Delta\theta = -[d^2 J(\theta_{i,j}^{(it)})/d(\theta_{i,j}^{(it)})^2]^{-1}[dJ(\theta_{i,j}^{(it)})/d\theta_{i,j}^{(it)}] \quad (11)$$

$d^2 J(\theta_{i,j}^{(it)})/d(\theta_{i,j}^{(it)})^2 = 0$ means that all the elements of $\boldsymbol{v}_i^{(it)}$ and $\boldsymbol{v}_j^{(it)}$ have the same sign. In this situation it is not necessary to perform the optimization since it will not decrease the criterion [10]. The process of all the rotation angles $\theta_{i,j}$ is called a rotation sweep.

B. Step 2: estimation of $\boldsymbol{R}^{(i,j)}(u_{i,j})$

After the rotation step, the elementary upper triangular matrix $\boldsymbol{R}^{(i,j)}(u_{i,j})$ transforms each pair of columns (i,j) of $\boldsymbol{V}$ as follows:

$$[\overline{\boldsymbol{w}}_i, \overline{\boldsymbol{w}}_j] = [\boldsymbol{v}_i, \boldsymbol{v}_j] \begin{bmatrix} 1 & u_{i,j} \\ 0 & 1 \end{bmatrix} \quad (12)$$

where $\overline{\boldsymbol{w}}_i$ is the i-th column of $\overline{\boldsymbol{W}}$, and $u_{i,j}$ is the shear factor. From (12), $\overline{\boldsymbol{w}}_i$ remains unchanged, the negativity measure criterion is then defined as follows:

$$J(u_{i,j}) = \frac{1}{2} \sum_{n=1}^{N} \overline{W}_{n,j}^2 \mathbb{1}_{\overline{W}_{n,j} < 0} \quad (13)$$

Similarly, the second order Taylor expansion $J_T(u_{i,j})$ of $J(u_{i,j})$ around $u_{i,j}^{(it)}$ can be expressed as follows:

$$J_T(u_{i,j}^{(it)} + \Delta u) = J(u_{i,j}^{(it)}) + \frac{dJ(u_{i,j}^{(it)})}{du_{i,j}^{(it)}} \Delta u + \frac{d^2 J(u_{i,j}^{(it)})}{2d(u_{i,j}^{(it)})^2}(\Delta u)^2 \quad (14)$$

where $u_{i,j}^{(it)}$ is the solution at the it-th iteration, $\Delta u = u_{i,j} - u_{i,j}^{(it)}$, $dJ(u_{i,j}^{(it)})/du_{i,j}^{(it)}$ and $d^2 J(u_{i,j}^{(it)})/d(u_{i,j}^{(it)})^2$ are the first and second order derivatives of (13) w.r.t. $u_{i,j}^{(it)}$, respectively, which are given by:

$$\frac{dJ(u_{i,j}^{(it)})}{du_{i,j}^{(it)}} = \sum_{n=1}^{N} \overline{W}_{n,i}^{(it)} \overline{W}_{n,j}^{(it)} \mathbb{1}_{\overline{W}_{n,j}^{(it)} < 0} \quad (15)$$

and:

$$\frac{d^2 J(u_{i,j}^{(it)})}{d(u_{i,j}^{(it)})^2} = \sum_{n=1}^{N} (\overline{W}_{n,i}^{(it)})^2 \mathbb{1}_{\overline{W}_{n,j}^{(it)} < 0} \quad (16)$$

with:

$$[\overline{W}_{n,i}^{(it)}, \overline{W}_{n,j}^{(it)}] = [v_{n,i}, v_{n,j}] \begin{bmatrix} 1 & u_{i,j}^{(it)} \\ 0 & 1 \end{bmatrix} \quad (17)$$

Then the minimum of (14) w.r.t. Δu can be reached at:

$$\Delta u = -[d^2 J(u_{i,j}^{(it)})/d(u_{i,j}^{(it)})^2]^{-1}[dJ(u_{i,j}^{(it)})/du_{i,j}^{(it)}] \quad (18)$$

When $d^2 J(u_{i,j}^{(it)})/d(u_{i,j}^{(it)})^2 = 0$, it means that all the elements of $\overline{\boldsymbol{w}}_j^{(it)}$ have the same sign and (18) does not need to be computed. The process of all the shear factors $u_{i,j}$ is called a shear sweep.

It is noteworthy that the second order derivatives (9) and (16) may not exists since the first order derivatives of the criteria (6) and (13) are continuous but non-differentiable at any point where $w_{n,j} = 0$ for some n, j [9]. In this case, we skip the estimation of the current parameter and continue to estimate the next one. There may exist a few very small negative values (less than $-1e^{-20}$) in the final result, which can be forced to zeros without affect the effectiveness of the algorithm. In practice, the new NN-COMP algorithm needs several rotation sweeps followed by several shear sweeps, in order to ensure the convergence.

III. SIMULATION RESULTS

In this section, the performance of the proposed NN-COMP algorithm is evaluated following the SeNICA model. Firstly, we aim at studying the convergence property of NN-COMP. Secondly, we test the usefulness of NN-COMP as a preprocessing step of a recent published SeNICA method based on a Nonnegative Joint Diagonalization by Congruence (NJDC) method NNLUJ1D [7]. For all the following experiments, a noisy SeNICA model $\boldsymbol{x} = \boldsymbol{As} + \boldsymbol{\nu}$ is considered and the additive noise $\boldsymbol{\nu}$ is modeled as a zero-mean unit-variance Gaussian vector. In addition, the Signal to Noise Ratio (SNR) is defined by: $\text{SNR} = 20 \log_{10}(\|\boldsymbol{As}\|_F / \|\boldsymbol{\nu}\|_F)$. We repeat the experiments with 200 independent Monte Carlo (MC) trials.
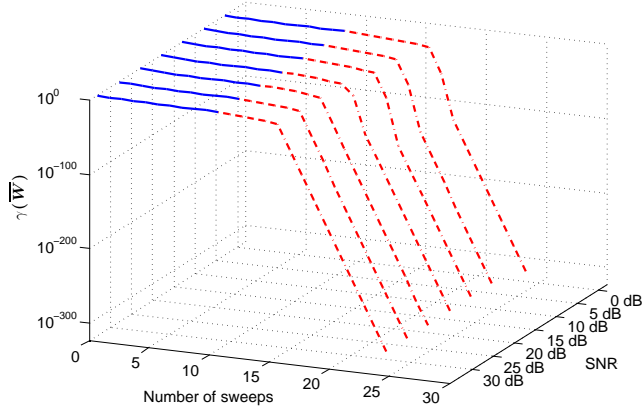
A. Convergence test

In this experiment, let us generate: the $(N \times P)$ mixing matrix $\boldsymbol{A}$ with elements independently drawn from a uniform distribution between 0 and 1, and the source vector $\boldsymbol{s}$ with $M = 2000$ sample points, which are independent and uniformly distributed between $-\sqrt{3}$ and $\sqrt{3}$. More precisely, we consider two situations: i) in the first one the number of observations N equals 20 and the number of sources P is set to 3, and ii) in the second one $N = 150$ and $P = 20$. For both situations, we vary the SNR from 0 dB to 30 dB with a step of 5 dB. To evaluate the convergence performance of the NN-COMP, we define a negativity criterion as a ratio between the sum of the squared negative entries and the sum of all the squared components of $\overline{\boldsymbol{W}}$: $\gamma(\overline{\boldsymbol{W}}) = \|\overline{\boldsymbol{W}}_-\|_F^2 / \|\overline{\boldsymbol{W}}\|_F^2$. Figure 2 depicts the mean values of the negativity criterion over the MC trials for different experiments. More precisely, the $\gamma(\overline{\boldsymbol{W}})$ values calculated for each rotation sweep are plotted with blue solid line and those calculated for each shear sweep are plotted with red dash-dot line. It is interesting to show that the negativity ratio dramatically decreases thanks to the second step of NN-COMP: multiplying $\boldsymbol{V}$ with the elementary upper triangular matrices. The proposed method seems to converge within a small number of sweeps, about 25 sweeps, whatever the number of the observations N, the number of sources P and the SNR values.

B. BSS performance on MRS data

We have recently presented a new NJDC method in [7], called NNLUJ1D. The SeNICA method based on NNLUJ1D, namely NNLUJ1D-ICA, outperforms some classical ICA
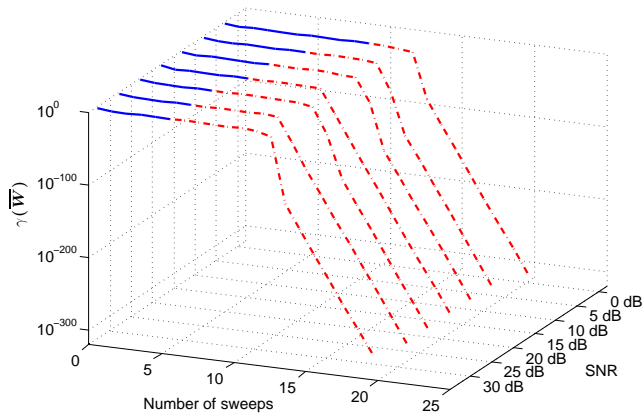
(a) $N = 20$ and $P = 3$

(b) $N = 150$ and $P = 20$

Fig. 2. Convergence toward nonnegativity as a function of the number of sweeps and the SNR values. The blue solid line denotes the rotation sweep and the red dash-dot line denotes the shear sweep.

methods on SeNICA problem. However, its numerical complexity, which is of order $\mathcal{O}(N^4(P + M))$, significantly increases as the number of observation N increases. Here, we propose to compare the behavior of the NNLUJ1D-ICA with and without using NN-COMP as a preprocessing step. For this purpose, an experiment is carried out on simulated MRS data, where the nonnegativity constraint on the mixing matrix $\boldsymbol{A}$ is verified. Two realistic MRS metabolites, namely the Choline and Creatine, are generated by Lorentzian and Gaussian functions [11]. The mixing matrix $\boldsymbol{A}$ is similarly generated as in the previous section. The performances of the two algorithms (NNLUJ1D-ICA and NN-COMP + NNLUJ1D-ICA) are studied as a function of the number of observations N, by varying N from 2 to 30 with a step of 2. We test two SNR values: 5 dB and 15 dB. The performance criterion is defined as the error between the source $\boldsymbol{s}$ and its estimate $\tilde{\boldsymbol{s}}$, as follows [12]:

$$\alpha(\boldsymbol{s}, \tilde{\boldsymbol{s}}) = (^1/_P) \sum_{p=1}^{P} \min_{(p,p') \in I_p^2} d(s_p, \tilde{s}_{p'}) \qquad (19)$$

where s_p and $\tilde{s}_{p'}$ are the p-th and p'-th components of $\boldsymbol{s}$ and $\tilde{\boldsymbol{s}}$, respectively. I_p^2 is defined recursively by: $I_1^2 = \{1, \cdots, P\} \times \{1, \cdots, P\}$, and $I_{p+1}^2 = I_p^2 - J_p^2$, where $J_p^2 = \arg\min_{(p,p') \in I_p^2} d(s_p, \tilde{s}_{p'})$. In addition, $d(s_p, \tilde{s}_{p'})$ is defined as

(a) Choline and Creatine source metabolites

(b) Two of the observations

(c) Estimates by NNLUJ1D-ICA

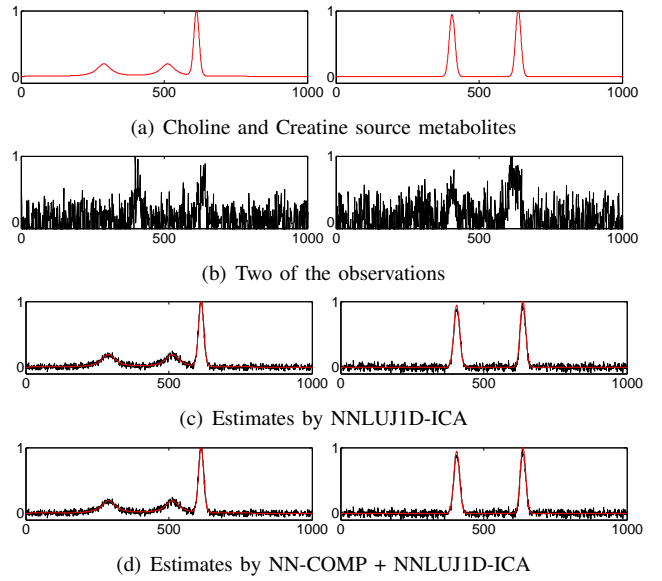(d) Estimates by NN-COMP + NNLUJ1D-ICA

Fig. 3. An example of MRS metabolites separation with 20 observations and a SNR of 5 dB. The red lines indicate the original sources, and the dark lines designate the estimated sources in figures (c) and (d).

the pseudo-distance between two random variables [13]:

$$d(s_p, \tilde{s}_{p'}) = 1 - (\|s_p \tilde{s}_{p'}^{\mathsf{T}}\|^2)/(\|s_p\|^2 \|\tilde{s}_{p'}\|^2)$$

The estimation accuracy of the mixing matrix is not provided, since we just estimate the compressed mixing matrix.

Figure 3 displays an example of the separation results of the two methods (NNLUJ1D-ICA and NN-COMP + NNLUJ1D-ICA) with N=20 observations and a SNR value of 5 dB. Both methods separate the sources quasi-perfectly. We can see that the NN-COMP algorithm does not deteriorate the separation quality ($\alpha(\boldsymbol{s}, \tilde{\boldsymbol{s}}) = 0.0543$ and 0.0544 for NNLUJ1D-ICA and NN-COMP + NNLUJ1D-ICA, respectively). The average curves of error $\alpha(\boldsymbol{s}, \tilde{\boldsymbol{s}})$, as well as that of the numerical complexities [14] in terms of floating point operations (flops) as a function of N are shown in figure 4. It shows that for both methods, the increase of N yields a better estimation of $\boldsymbol{s}$ regardless of the tested values of SNR. The two methods perform quite similarly in terms of the source separation quality. For each N value and SNR level, the errors $\alpha(\boldsymbol{s}, \tilde{\boldsymbol{s}})$ generated by the two methods are approximately the same. However, the numerical complexity of NNLUJ1D-ICA grows exponentially up to 10^{10} flops as a function of N. Fortunately, the burdensome numerical complexity can be reduced to less than 10^6 flops, thanks to the NN-COMP method. It confirms the validity and the effectiveness of the proposed NN-COMP method as a preprocessing step of SeNICA algorithms.

IV. CONCLUSION

In this paper, we introduced a nonnegative compression method, called NN-COMP, in order to solve large scale SeNICA problems. It transforms the classical prewhitening matrix into a nonnegative matrix, which then preserves the
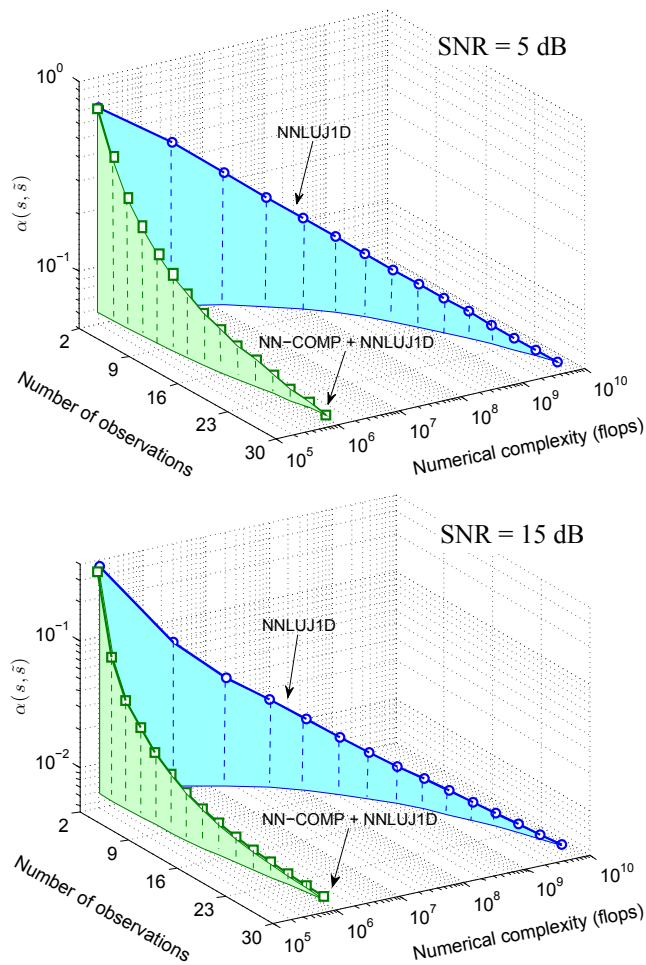
Fig. 4. Performance evolution of two different methods as a function of the number of observations for separating 2 MRS metabolites.

nonnegativity of the original mixing matrix. Simulation results show its fast convergence property. An illustration of a BSS application on MRS data confirms the validity and the efficiency of the proposed method.

REFERENCES

[1] J. Coloigner, A. Karfoul, L. Albera, and P. Comon, "Line search and trust region strategies for canonical decomposition of semi-nonnegative semi-symmetric 3rd order tensors," *Linear Algebra and its Applications*, vol. 450, no. 1, pp. 334–374, 2014.

[2] J. Coloigner, A. Fargeas, A. Kachenoura, L. Wang, L. Senhadji, R. De Crevoisier, O. Acosta, and L. Albera, "A novel classification method for prediction of rectal bleeding in prostate cancer radiotherapy based on a semi-nonnegative ICA of 3D planned dose distributions," 2014, accepted for publication in *IEEE Journal of Biomedical and Health Informatics*. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6846265

[3] J. Coloigner, L. Albera, A. Kachenoura, F. Noury, and L. Senhadji, "Semi-nonnegative joint diagonalization by congruence and semi-nonnegative ICA," *Signal Processing*, vol. 105, pp. 185–197, December, 2014 2014.

[4] P. Comon, "Independent component analysis, a new concept?" *Signal Processing*, vol. 36, no. 3, pp. 287–314, 1994.

[5] A. Belouchrani, K. Abed-Meraim, J. F. Cardoso, and E. Moulines, "A blind source separation technique using second-order statistics," *IEEE Transactions on Signal Processing*, vol. 45, no. 2, pp. 434–444, 1997.

[6] L. Wang, L. Albera, A. Kachenoura, H. Z. Shu, and L. Senhadji, "Nonnegative joint diagonalization by congruence based on LU matrix factorization," *IEEE Signal Processing Letters*, vol. 20, no. 8, pp. 807–810, 2013.

[7] L. Wang, L. Albera, H. Z. Shu, and L. Senhadji, "A new Jacobi-like nonnegative joint diagonalization by congruence," in *EUSIPCO'13, Proceedings of the XXI European Signal Processing Conference*, Marrakech, Morocco, September 9–13 2013, pp. 1–5.

[8] W. S. B. Ouedraogo, A. Souloumiac, M. Jaidane, and C. Jutten, "Non-negative blind source separation algorithm based on minimum aperture simplicial cone," *IEEE Transactions on Signal Processing*, vol. 62, no. 2, pp. 376–389, 2014.

[9] M. D. Plumbley, "Algorithms for nonnegative independent component analysis," *IEEE Transactions on Neural Networks*, vol. 14, no. 3, pp. 534–543, 2003.

[10] W. S. B. Ouedraogo, A. Souloumiac, and C. Jutten, "Non-negative independent component analysis algorithm based on 2D Givens rotations and a Newton optimization," in *Latent Variable Analysis and Signal Separation, Lecture Notes in Computer Science*, vol. 6365, 2010, pp. 522–529.

[11] S. Moussaoui, D. Brie, and C. Carteret, "Non-negative source separation using the maximum likelihood approach," in *IEEE Workshop on Statistical Signal Processing (SSP'05)*, 2005, pp. 1114–1119.

[12] J. Coloigner, "Line search and trust region strategies for canonical decomposition of semi-nonnegative semi-symmetric tensors," Ph.D. dissertation, Université de Rennes 1, 2012.

[13] L. Albera, A. Ferréol, P. Comon, and P. Chevalier, "Blind identification of overcomplete mixtures of sources (BIOME)," *Linear Algebra and its Applications*, vol. 391, pp. 3–30, 2004.

[14] L. Albera, P. Comon, L. Parra, A. Karfoul, A. Kachenoura, and L. Senhadji, "Biomedical applications," in *Handbook of Blind Source Separation*, P. Comon and C. Jutten, Eds. Oxford, UK: Elsevier, 2010, ch. 18, pp. 737–777.