



HAL
open science

Nonnegative Compression for Semi-Nonnegative Independent Component Analysis

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

► **To cite this version:**

Lu Wang, Amar Kachenoura, Laurent Albera, Ahmad Karfoul, Huazhong Shu, et al.. Nonnegative Compression for Semi-Nonnegative Independent Component Analysis. The eighth IEEE Sensor Array and Multi-Channel Signal Processing Workshop, Jun 2014, A Coruna, Spain. 4 p. hal-01012136v1

HAL Id: hal-01012136

<https://hal.science/hal-01012136v1>

Submitted on 25 Jun 2014 (v1), last revised 29 Oct 2014 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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: wangLyul986@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 (SN-ICA) 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.

I. INTRODUCTION AND PROBLEM FORMULATION

The Semi-Nonnegative Independent Component Analysis (SN-ICA) problem is defined as follows [1], [2]:

Problem 1. Given a real N -dimensional random vector process $\mathbf{x}[m]$, find an $(N \times P)$ mixing matrix \mathbf{A} and a P -dimensional source random process $\mathbf{s}[m]$, such that:

$$\mathbf{x}[m] = \mathbf{A}\mathbf{s}[m] + \mathbf{v}[m], \quad m \in \{1, 2, \dots, M\} \quad (1)$$

where \mathbf{A} has nonnegative components, $\mathbf{s}[m]$ has statistically independent components, and $\mathbf{v}[m]$ is an N -dimensional Gaussian noise vector, independent of $\mathbf{s}[m]$. M is the number of sample points.

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 and they become nonnegative after a complicated phase shift procedure.

In many ICA algorithms, in order to reduce the dimension of the observation space, common compression method truncates the N -dimensional vector $\mathbf{x}[m]$ into a vector $\bar{\mathbf{x}}[m]$ of dimension $P \ll N$. The estimate of the rank P is determined by the number of eigenvalues of the covariance matrix of $\mathbf{x}[m]$ not exceedingly close to zero. The compressed observation vector $\bar{\mathbf{x}}[m]$ is expressed as follows [3], [4]:

$$\bar{\mathbf{x}}[m] = \mathbf{W}^T \mathbf{x}[m] = (\mathbf{W}^T \mathbf{A}) \mathbf{s}[m] = \bar{\mathbf{A}} \mathbf{s}[m] \quad (2)$$

where the columns of $\mathbf{W} \in \mathbb{R}^{N \times P}$ are the scaled eigenvectors corresponding to the P largest eigenvalues of the covariance

matrix $E(\mathbf{x}[m]\mathbf{x}[m]^T)$, assuming $\mathbf{x}[m]$ being centered. This procedure is known as the spatial prewhitening and \mathbf{W} is called a prewhitening matrix. However, such a compression method can not guarantee the nonnegativity of the compressed mixing matrix $\bar{\mathbf{A}} = \mathbf{W}^T \mathbf{A}$, because generally \mathbf{W} is not nonnegative. In some practical situations, exploiting the nonnegativity property can improve the ICA result [5]–[7]. Fortunately, it is possible to transform \mathbf{W} into the nonnegative quadrant by column-pair rotations and shearing transformations, which can be achieved by multiplying \mathbf{W} by a series of Givens rotation matrices and elementary upper triangular matrices, respectively. Now let us recall the definitions of these two matrices:

Definition 1. A Givens rotation matrix $\mathbf{R}^{(i,j)}(\theta_{i,j})$, 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.

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

Then the nonnegative compression problem is defined as follows:

Problem 2. Given a prewhitening matrix $\mathbf{W} \in \mathbb{R}^{N \times P}$ of an N -dimensional random vector process $\mathbf{x}[m]$, find a sequence of Givens rotation matrices and elementary upper triangular matrices, such that their product:

$$\bar{\mathbf{W}} = \mathbf{W} \underbrace{\prod_{i=1}^P \prod_{j=i+1}^P \mathbf{R}^{(i,j)}(\theta_{i,j})}_{\stackrel{\text{def}}{=} \mathbf{V}} \prod_{i=1}^P \prod_{j=i+1}^P \mathbf{U}^{(i,j)}(u_{i,j}) \quad (3)$$

have nonnegative components.

Figure 1 illustrates, in the case of $P = 2$, the process of transforming a prewhitening matrix \mathbf{W} into the nonnegative quadrant by equation (3). Prewhitening makes the axes of the matrix \mathbf{W} orthogonal to each other (figure 1(a)). A Givens rotation matrix searches for a rotation angle that makes the outputs matrix \mathbf{V} as nonnegative as possible. However, sometimes it still remains some negative values near the quadrant boundaries (figure 1(b) left). That is because the columns of \mathbf{A} are neither well-grounded nor statistically independent

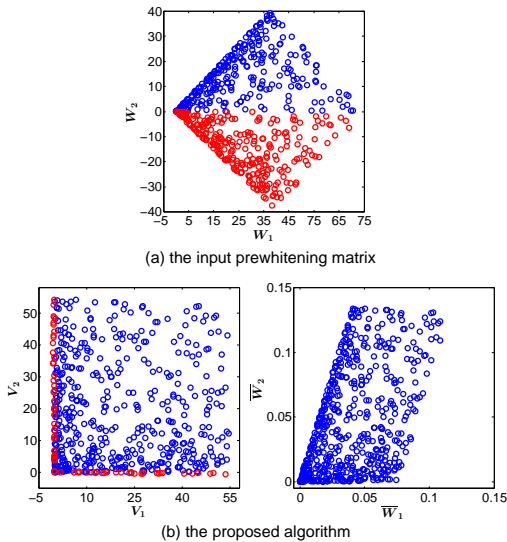


Fig. 1. An illustration of the proposed algorithm. (a) the input prewhitening matrix \mathbf{W} ; (b) the proposed algorithm, left: \mathbf{V} is obtained by multiplying \mathbf{W} by a Givens rotation matrix, right: $\bar{\mathbf{W}}$ is obtained by multiplying \mathbf{V} by an elementary upper triangular matrix. The blue circles denote the nonnegative values and the red circles denote the negative values.

[8]. We propose to use an elementary upper triangular matrix which projects the remaining negative values of \mathbf{V} into the nonnegative quadrant (figure 1(b) right). We thus obtain a nonnegative compression matrix $\bar{\mathbf{W}}$ and the $(P \times P)$ compressed mixing matrix $\bar{\mathbf{A}} = \bar{\mathbf{W}}^T \mathbf{A}$, which preserves the nonnegativity property of \mathbf{A} . Then $\bar{\mathbf{A}}$ and $\mathbf{s}[m]$ can be estimated by SN-ICA methods [1], [5], [6]. Therefore it is not necessary to estimate the original mixing matrix \mathbf{A} . It is noteworthy that the scaling and permutation ambiguities of the ICA problem are not removed by the compression step. Now the challenge is how to compute the Givens rotation and the elementary upper triangular matrices.

II. METHOD

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

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

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

$$[\mathbf{V}_i, \mathbf{V}_j] = [\mathbf{W}_i, \mathbf{W}_j] \begin{bmatrix} \cos(\theta_{i,j}) & -\sin(\theta_{i,j}) \\ \sin(\theta_{i,j}) & \cos(\theta_{i,j}) \end{bmatrix} \quad (4)$$

where \mathbf{V}_i and \mathbf{W}_i are the i -th columns of \mathbf{V} and \mathbf{W} , respectively. Now, let us consider the following negativity measure criterion defined in [8] and [9]:

$$J(\theta_{i,j}) = \frac{1}{2} \sum_{n=1}^N (V_{n,i}^2 \mathbb{1}_{V_{n,i} < 0} + V_{n,j}^2 \mathbb{1}_{V_{n,j} < 0}) \quad (5)$$

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 \mathbf{V} , respectively. The purpose is to find an angle, $\theta_{i,j}$, which minimizes the total sum of squares of negative elements of \mathbf{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 $\theta_{i,j}^{(it)}$:

$$J_T(\theta_{i,j}^{(it)} + \Delta\theta) = J(\theta_{i,j}^{(it)}) + \frac{dJ(\theta_{i,j}^{(it)})}{d\theta_{i,j}^{(it)}} \Delta\theta + \frac{d^2 J(\theta_{i,j}^{(it)})}{2d(\theta_{i,j}^{(it)})^2} (\Delta\theta)^2 \quad (6)$$

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 (5) with respect to (w.r.t) $\theta_{i,j}^{(it)}$, respectively, which are given as follows [9]:

$$\begin{aligned} \frac{dJ(\theta_{i,j}^{(it)})}{d\theta_{i,j}^{(it)}} &= \sum_{n=1}^N V_{n,i}^{(it)} V_{n,j}^{(it)} \times \\ & \quad (\mathbb{1}_{V_{n,i}^{(it)} < 0} \mathbb{1}_{V_{n,j}^{(it)} > 0} - \mathbb{1}_{V_{n,i}^{(it)} > 0} \mathbb{1}_{V_{n,j}^{(it)} < 0}) \end{aligned} \quad (7)$$

and

$$\begin{aligned} \frac{d^2 J(\theta_{i,j}^{(it)})}{d(\theta_{i,j}^{(it)})^2} &= \sum_{n=1}^N \left((V_{n,j}^{(it)})^2 - (V_{n,i}^{(it)})^2 \right) \times \\ & \quad (\mathbb{1}_{V_{n,i}^{(it)} < 0} \mathbb{1}_{V_{n,j}^{(it)} > 0} - \mathbb{1}_{V_{n,i}^{(it)} > 0} \mathbb{1}_{V_{n,j}^{(it)} < 0}) \end{aligned} \quad (8)$$

with

$$[V_{n,i}^{(it)}, V_{n,j}^{(it)}] = [W_{n,i}, W_{n,j}] \begin{bmatrix} \cos(\theta_{i,j}^{(it)}) & -\sin(\theta_{i,j}^{(it)}) \\ \sin(\theta_{i,j}^{(it)}) & \cos(\theta_{i,j}^{(it)}) \end{bmatrix} \quad (9)$$

Then the minimum of (6) w.r.t $\Delta\theta$ can be reached at:

$$\Delta\theta = -[d^2 J(\theta_{i,j}^{(it)})/d(\theta_{i,j}^{(it)})^2]^{-1} [dJ(\theta_{i,j}^{(it)})/d\theta_{i,j}^{(it)}] \quad (10)$$

$d^2 J(\theta_{i,j}^{(it)})/d(\theta_{i,j}^{(it)})^2 = 0$ means that all the elements of $\mathbf{V}_i^{(it)}$ and $\mathbf{V}_j^{(it)}$ have the same sign. In this situation it is not necessary to perform the optimization since it will not decrease the criterion [9]. The process of all the rotation angles $\theta_{i,j}$ is called a rotation sweep.

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

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

$$[\bar{\mathbf{W}}_i, \bar{\mathbf{W}}_j] = [\mathbf{V}_i, \mathbf{V}_j] \begin{bmatrix} 1 & u_{i,j} \\ 0 & 1 \end{bmatrix} \quad (11)$$

where $\bar{\mathbf{W}}_i$ is the i -th column of $\bar{\mathbf{W}}$, and $u_{i,j}$ is also called a shearing factor. From (11), $\bar{\mathbf{W}}_i$ remains unchanged, the negativity measure criterion is then defined as follows:

$$J(u_{i,j}) = \frac{1}{2} \sum_{n=1}^N \bar{W}_{n,j}^2 \mathbb{1}_{\bar{W}_{n,j} < 0} \quad (12)$$

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:

$$J_T(u_{i,j}^{(it)} + \Delta u) = J(u_{i,j}^{(it)}) + \frac{dJ(u_{i,j}^{(it)})}{du_{i,j}^{(it)}} \Delta u + \frac{d^2 J(u_{i,j}^{(it)})}{2d(u_{i,j}^{(it)})^2} (\Delta u)^2 \quad (13)$$

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 (12) w.r.t $u_{i,j}^{(it)}$, respectively, which are given as follows:

$$\frac{dJ(u_{i,j}^{(it)})}{du_{i,j}^{(it)}} = \sum_{n=1}^N \bar{W}_{n,i}^{(it)} \bar{W}_{n,j}^{(it)} \mathbb{1}_{\bar{W}_{n,j}^{(it)} < 0} \quad (14)$$

and

$$\frac{d^2 J(u_{i,j}^{(it)})}{d(u_{i,j}^{(it)})^2} = \sum_{n=1}^N (\bar{W}_{n,i}^{(it)})^2 \mathbb{1}_{\bar{W}_{n,j}^{(it)} < 0} \quad (15)$$

with

$$[\bar{W}_{n,i}^{(it)}, \bar{W}_{n,j}^{(it)}] = [V_{n,i}, V_{n,j}] \begin{bmatrix} 1 & u_{i,j}^{(it)} \\ 0 & 1 \end{bmatrix} \quad (16)$$

Then the minimum of (13) w.r.t Δu can be reached at:

$$\Delta u = -[d^2 J(u_{i,j}^{(it)})/d(u_{i,j}^{(it)})^2]^{-1} [dJ(u_{i,j}^{(it)})/du_{i,j}^{(it)}] \quad (17)$$

When $d^2 J(u_{i,j}^{(it)})/d(u_{i,j}^{(it)})^2 = 0$, it means that all the elements of $\bar{W}_j^{(it)}$ have the same sign and (17) does not need to be computed. The process of all the shearing factors $u_{i,j}$ is called a shearing sweep.

The criterions (5) and (12), and their first order derivatives are continuous but non-differentiable at any point where $W_{n,j} = 0$ for some n, j [8]. In this case, there may exist a few very small negative values (less than $-1e^{-20}$) in the result, which can be forced to zeros without affect the effectiveness of the algorithm. In practice, the new NN-COMP algorithm contains some rotation sweeps followed by several shearing 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 SN-ICA model presented in equation (1). Firstly, we aim to study the convergence property of NN-COMP. Secondly, we test the usefulness of the NN-COMP as a preprocessing step of a recent published SN-ICA method, namely NNLUJ1D [6]. For all the following experiments, the additive noise $\nu[m]$ is modeled as a zero-mean unit-variance Gaussian vector process, and the Signal to Noise Ratio (SNR) is defined by: $\text{SNR} = -20 \log_{10}(\|\nu[m]\|_F / \|\mathbf{A}\mathbf{s}[m]\|_F)$. 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 \mathbf{A} with elements independently drawn from a uniform distribution between 0 and 1, and the source vector $\mathbf{s}[m]$ 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 $\bar{\mathbf{W}}$: $\gamma(\bar{\mathbf{W}}) = \|\bar{\mathbf{W}}_-\|_F^2 / \|\bar{\mathbf{W}}\|_F^2$. Figure 2 depicts the mean values of the negativity criterion over the

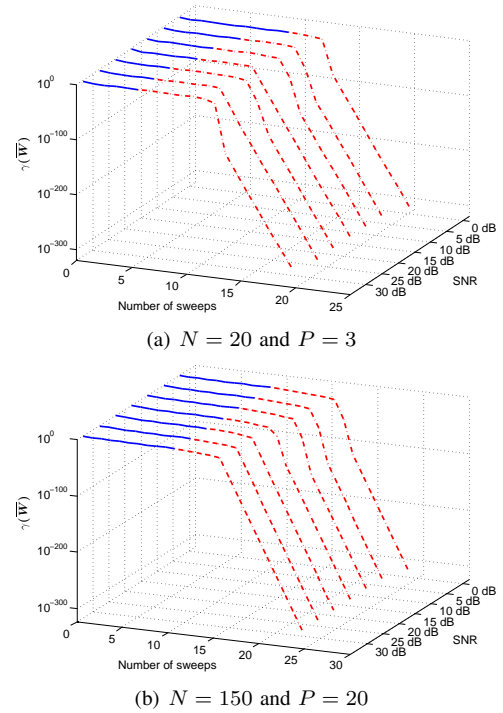


Fig. 2. Convergence toward nonnegativity as a function of the number of sweeps and the SNR values. The blue solid line denotes the rotation sweep and the red dash-dot line denotes the shearing sweep.

MC trials for different experiments. More precisely, the $\gamma(\bar{\mathbf{W}})$ values calculated for each rotation sweep are plotted with blue solid line and those calculated for each shearing 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 \mathbf{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 method in [6], called NNLUJ1D, that outperforms some classical ICA methods on SN-ICA 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 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 \mathbf{A} is verified. Two realistic MRS metabolites, namely the Choline and Creatine, are generated by Lorentzian and Gaussian functions [10]. The mixing matrix \mathbf{A} is similarly generated as in the previous section. The performances of the two algorithms (NNLUJ1D and NN-COMP + NNLUJ1D) 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 $\mathbf{s}[m]$ and its estimate $\tilde{\mathbf{s}}[m]$, as follows [2]:

$$\alpha(\mathbf{s}[m], \tilde{\mathbf{s}}[m]) = (1/P) \sum_{p=1}^P \min_{(p,p') \in I_p^2} d(s_p[m], \tilde{s}_{p'}[m]) \quad (18)$$

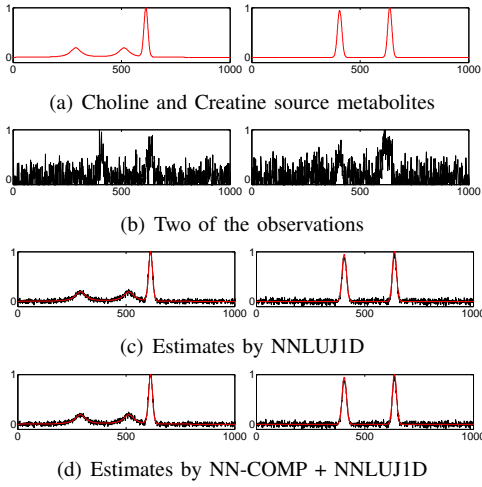


Fig. 3. An example of MRS metabolites separation with 20 observations and a SNR of 5 dB. The red lines indicate the original sources, and the dark lines designate the estimated sources in figures (c) and (d).

where $s_p[m]$ and $\tilde{s}_{p'}[m]$ are the p -th and p' -th components of $\mathbf{s}[m]$ and $\tilde{\mathbf{s}}[m]$, respectively. I_p^2 is defined recursively by: $I_1^2 = \{1, \dots, P\} \times \{1, \dots, P\}$, and $I_{p+1}^2 = I_p^2 - J_p^2$, where $J_p^2 = \operatorname{argmin}_{(p,p') \in I_p^2} d(s_p[m], \tilde{s}_{p'}[m])$. In addition, $d(s_p[m], \tilde{s}_{p'}[m])$ is defined as the pseudo-distance between two random variables [11]:

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

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 and NN-COMP + NNLUJ1D) 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 ($\alpha(\mathbf{s}[m], \tilde{\mathbf{s}}[m]) = 0.0543$ and 0.0544 for NNLUJ1D and NN-COMP + NNLUJ1D, respectively). The average curves of error $\alpha(\mathbf{s}[m], \tilde{\mathbf{s}}[m])$, as well as that of the numerical complexities [12] 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 $\mathbf{s}[m]$ 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 $\alpha(\mathbf{s}[m], \tilde{\mathbf{s}}[m])$ generated by the two methods are approximately the same. However, the numerical complexity of NNLUJ1D 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 SN-ICA algorithms.

IV. CONCLUSION

In this paper, we introduced a nonnegative compression method, called NN-COMP, for SN-ICA problem. It transforms the classical prewhitening matrix into a nonnegative matrix by using a Newton's method. Simulation results show its fast convergence property. An illustration of BSS application on MRS data confirms the validity and the efficiency of the

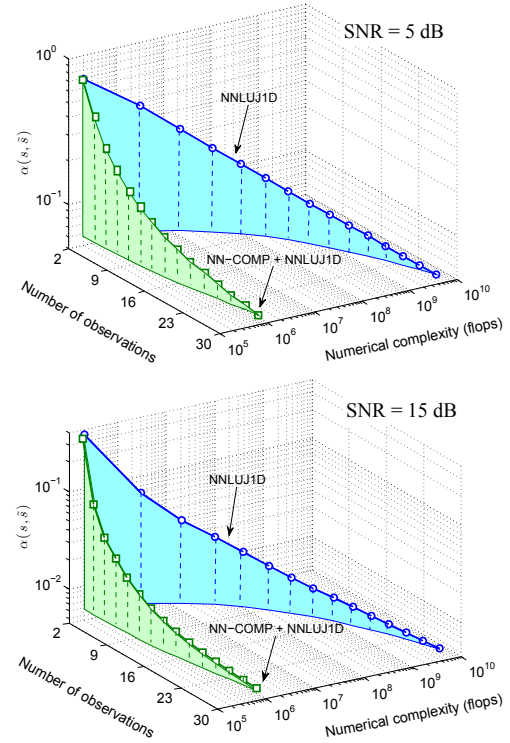


Fig. 4. Performance evolution of two different methods as a function of the number of observations for separating 2 MRS metabolites.

proposed method.

REFERENCES

- [1] J. Coloigner, L. Albera, A. Karfoul, A. Kachenoura, P. Comon, and L. Senhadji, "Semi-nonnegative independent component analysis: the (3,4)-SENICA^{EXP} method," in *LVA/ICA'10*, 2010, pp. 612–619.
- [2] 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.
- [3] P. Comon, "Independent component analysis, a new concept?" *Signal Process.*, vol. 36, no. 3, pp. 287–314, 1994.
- [4] A. Belouchrani, K. Abed-Meraim, J. F. Cardoso, and E. Moulines, "A blind source separation technique using second-order statistics," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 434–444, 1997.
- [5] L. Wang, L. Albera, A. Kachenoura, H. Shu, and L. Senhadji, "Non-negative joint diagonalization by congruence based on LU matrix factorization," *IEEE Signal Process. Lett.*, vol. 20, no. 8, pp. 807–810, 2013.
- [6] L. Wang, L. Albera, H. Shu, and L. Senhadji, "A new Jacobi-like nonnegative joint diagonalization by congruence," in *Proceedings of the XXI European Signal Processing Conference (EUSIPCO'13)*, 2013, paper id: 1569744591.
- [7] W. S. B. Ouedraogo, A. Souloumiac, M. Jaidane, and C. Jutten, "Non-negative blind source separation algorithm based on minimum aperture simplicial cone," *IEEE Trans. Signal Process.*, vol. 62, no. 2, pp. 376–389, 2014.
- [8] M. D. Plumbley, "Algorithms for nonnegative independent component analysis," *IEEE Trans. Neural Netw.*, vol. 14, no. 3, pp. 534–543, 2003.
- [9] 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 *LVA/ICA 2010*, 2010, pp. 522–529.
- [10] 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.
- [11] L. Albera, A. Ferréol, P. Comon, and P. Chevalier, "Blind identification of overcomplete mixtures of sources (BIOME)," *Linear Algebra Appl.*, vol. 391, pp. 3–30, 2004.
- [12] L. Albera, P. Comon, L. Parra, A. Karfoul, A. Kachenoura, and L. Senhadji, *Handbook of Blind Source Separation*. Elsevier, 2010, ch. 18: Biomedical applications, pp. 737–777.