Blind Source Separation by Fully Nonnegative Constrained Iterative Volume Maximization

Zuyuan Yang, Shuxue Ding, *Member, IEEE*, and Shengli Xie, *Senior Member, IEEE*

***Abstract*—Blind source separation (BSS) has been widely discussed in many real applications. Recently, under the assumption that both of the sources and the mixing matrix are nonnegative, Wang *et al.* develop an amazing BSS method by using volume maximization. However, the algorithm that they have proposed can guarantee the nonnegativities of the sources only, but cannot obtain a nonnegative mixing matrix necessarily. In this letter, by introducing additional constraints, a method for fully nonnegative constrained iterative volume maximization (FNCIVM) is proposed. The result is with more interpretation, while the algorithm is based on solving a single linear programming problem. Numerical experiments with synthetic signals and real-world images are performed, which show the effectiveness of the proposed method.**

***Index Terms*—Blind source separation, fully nonnegative constrained iterative volume maximization.**

I. Introduction

BLIND source separation (BSS) is an attractive technology which can separate the sources only from their mixtures, i.e., observations. Due to its wide applications, such as in spectral unmixing [1], biomedical image processing [2], etc., BSS is a commonly hot topic in the areas of signal processing, machine learning, etc. As many real-world sources are nonnegative, blind separation of nonnegative sources, i.e., nonnegative BSS (nBSS), has also been extensively discussed.

Among the various approaches to BSS, independent component analysis (ICA) plays an important role [3]–[5]. With merging the intrinsic nonnegativity to ICA based methods, some reasonable progresses have been made in these years. Oja and Plumbley exploited the nonnegativity and independence of the sources and analyzed the identifiability of well-grounded sources [6]. Astakhov *et al.* introduced a stochastic nonnegative ICA algorithm [7]. The methods above aim to recover the independent components from the mixtures, but may fail to solve BSS problem with statistically dependent sources.

Nonnegative matrix factorization (NMF) is a newly developed method which aims to Decompose a nonnegative matrix into the product of two nonnegative matrices or factors [8]. As it does not rely on the mutual independence between the sources, NMF (with some constraints) shows some potentials to perform nBSS no matter the sources are mutually dependent or independent [9]. However, a well known problem of NMF is that it may not directly yield a unique decomposition [10]. Although the uniqueness may be improved by imposing some constraints to the factors, it is still a challenging problem to uniquely identify the sources in general cases [11].

For separating the positive dependent sources, Wang *et al.* developed an amazing method, i.e., nonnegative least-correlated component analysis by iterative volume maximization (nLCA-IVM) [12]. Although it has been derived under the assumptions that both of the mixing matrix and the sources are nonnegative, problematically, developed algorithm, nLCA-IVM, cannot ensure the nonnegativity of the estimated mixing matrix. In this letter, a method of fully nonnegative constrained iterative volume maximization (FNCIVM) is proposed, and the results of which satisfy the assumptions completely. Because of this, the results become more interpretive. We show that the nBSS problem can be converted into a single linear programming (LP) problem under some constraints that are essence of nonnegativities of the mixing matrix and the sources.

II. BSS Using FNCIVM

Typical BSS mixing and unmixing models are as following, respectively (neglecting the background noise):

$$\mathbf{X} = \mathbf{AS} \tag{1}$$

$$\mathbf{Y} = \mathbf{WX} \tag{2}$$

where $\mathbf{X} \in \Re^{m\times N}$ denotes the observations, $\mathbf{A} \in \Re^{m\times n}$ denotes the mixing matrix; $\mathbf{S} \in \Re^{n\times N}$ denotes the latent sources; $\mathbf{Y} \in \Re^{n\times N}$ denotes the separated signals; $\mathbf{W} \in \Re^{n\times m}$ denotes the unmixing matrix; and m, n, N denote the number of the observations, sources, and samples, respectively.

Recently, due to the widely applications, the determined or over-determined BSS (i.e., $m \geq n$) has attracted sustaining attentions. For simplicity, we will mainly discuss the case of $m = n > 2$ (for $m > n$, the observations can be preprocessed using the rank reduction method in [12]). To be practical in imaging applications, it is assumed that the sources are nonnegative, the mixing matrix is nonnegative, with full rank, unit row sum (i.e., $a_{ij} \geq 0, \mathrm{rank}(A) = n, \sum_{t=1}^{n} a_{it} = 1, \forall i$) [12].

Based on the analysis in [12], recovering the sources can be posed as this optimization problem with respect to $\mathbf{W}$

Manuscript received April 20, 2010; revised June 16, 2010; accepted June 21, 2010. Date of publication July 01, 2010; date of current version July 22, 2010. This work was supported by the National 973 Project of China under Grant 2010CB731800, the Key Program of National Natural Science Foundation of China under Grant U0835003, the National Natural Science Foundation of China under Grant 60874061 and the 2008-2010 Grants-In-Aid for Scientific Research, Ministry of Education, Culture, Sports, Science and Technology, Japan (20500209). The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Jen-Tzung Chien.

Z. Y. Yang and S. L. Xie are with School of Electronic & Information Engineering, South China University of Technology, Guangzhou, Guangdong 510641, China (e-mail: yangzuyuan@yahoo.com.cn; eeoshlxie@scut.edu.cn).

S. Ding is with the School of Computer Science and Engineering, the University of Aizu, Aizu-Wakamatsu City, Fukushima 965-8580, Japan and also with the Brain Science Institute, RIKEN, Saitama 351-0198, Japan (e-mail: sding@u-aizu.ac.jp).

Digital Object Identifier 10.1109/LSP.2010.2055854

$$\max_{\mathbf{W}} \quad |\det(\mathbf{W})|$$
$$\text{s.t. } \mathbf{W}\mathbf{1}_m = \mathbf{1}_n;\ \ y_{ij} = \sum_{t=1}^{m} w_{it}x_{tj} \geq 0,\ \forall i,j \quad (3)$$

where $\mathbf{1}_n$ denotes the all-one column vector of n-dimension.

A problem is, however, that only the nonnegativity of the components is constrained in (3). It does not necessarily generate a nonnegative mixing matrix. In order to obtain more interpretative results, we will introduce additional constraints to $\mathbf{W}$ such that the estimation of $\mathbf{A}$ is also nonnegative. Let $\mathbf{M}^{kt}$ be the submatrix of $\mathbf{W}$ with the kth row and tth column removed and c_{kt} be the algebraic complement of w_{kt}, then $c_{kt} = (-1)^{k+t}\det(\mathbf{M}^{kt})$. Let $\mathbf{W}^*$ be the adjoint matrix of $\mathbf{W}$, then $\mathbf{A} = \mathbf{W}^{-1} = (1/\det(\mathbf{W}))\mathbf{W}^*$. Based on the analysis in [13], it holds that $w^*_{kt} = c_{kt}$. Therefore

$$a_{tk} = \frac{w^*_{tk}}{\det(\mathbf{W})} = \frac{c_{kt}}{\det(\mathbf{W})} = \frac{(-1)^{k+t}}{\det(\mathbf{W})}\det(\mathbf{M}^{kt}). \quad (4)$$

To present a more clear relationship between elements in $\mathbf{A}$ and $\mathbf{W}$, we factorize $\det(\mathbf{M}^{kt})$ in (4) by the co-factor expansion. For $\forall i, k \neq i$, let $i_1 = i-1$ if $k < i$ and $i_1 = i$ if $k > i$, then

$$\det(\mathbf{M}^{kt}) = \sum_{j_1=1}^{m-1} [\mathbf{M}^{kt}]_{i_1 j_1} d^{kt}_{i_1 j_1} \quad (5)$$

where $d^{kt}_{i_1 j_1}$ denotes the algebraic complement of $[\mathbf{M}^{kt}]_{i_1 j_1}$. Note that $\left\{[\mathbf{M}^{kt}]_{i_1 1}, \ldots, [\mathbf{M}^{kt}]_{i_1 (m-1)}\right\}$ is a subset of $\{w_{i1}, \ldots, w_{im}\}$. So, $d^{kt}_{i_1 j_1}$ is independent of $\{w_{i1}, \ldots, w_{im}\}$ [13]. Let

$$[\tilde{\mathbf{M}}^{kt}]_{i_1 j_1} = \begin{cases} [\mathbf{M}^{kt}]_{i_1 j_1}, & \forall j_1 \in \{1, \ldots, t-1\} \\ w_{i j_1}, & j_1 = t \\ [\mathbf{M}^{kt}]_{i_1 (j_1 - 1)}, & \forall j_1 \in \{t+1, \ldots, m\} \end{cases}. \quad (6)$$

Then, it holds that $[\tilde{\mathbf{M}}^{kt}]_{i_1 j_1} = w_{i j_1}, \forall j_1 \in \{1, \ldots, m\}$. Let

$$\tilde{d}^{kt}_{i_1 j_1} = \begin{cases} d^{kt}_{i_1 j_1}, & \forall j_1 \in \{1, \ldots, t-1\} \\ 0, & j_1 = t \\ d^{kt}_{i_1 (j_1 - 1)}, & \forall j_1 \in \{t+1, \ldots, m\} \end{cases}. \quad (7)$$

Based on (6) and (7), we can rewrite (5) as

$$\det(\mathbf{M}^{kt}) = \sum_{j_1=1}^{m} [\tilde{\mathbf{M}}^{kt}]_{i_1 j_1} \tilde{d}^{kt}_{i_1 j_1} = \sum_{j_1=1}^{m} w_{i j_1} \tilde{d}^{kt}_{i_1 j_1}. \quad (8)$$

Thus, substituting (8) to (4), we obtain

$$a_{tk} = \frac{(-1)^{k+t}}{\det(\mathbf{W})} \sum_{j_1=1}^{m} w_{i j_1} \tilde{d}^{kt}_{i_1 j_1}. \quad (9)$$

On the other hand, by the co-factor expansion of $\det(\mathbf{W})$ with respect to the ith row of $\mathbf{W}$, it holds that

$$\det(\mathbf{W}) = \sum_{j=1}^{m} (-1)^{i+j} w_{ij} \det(\mathbf{M}^{ij}). \quad (10)$$

It is a linear function with respect to $w_{i1}, \ldots, w_{im}$. So, the objective function in (3) may be optimized line-by-line by using LP. For updating $w_{i1}, \ldots, w_{im}$, the nonnegativity constraint to $\mathbf{A}$ is:

$$a_{tk} = \begin{cases} \frac{(-1)^{k+t}}{\det(\mathbf{W})} \mathbf{w}_i \tilde{\mathbf{d}}^{kt}_{i-1} \geq 0 & k < i \\ \frac{(-1)^{k+t}}{\det(\mathbf{W})} \det(\mathbf{M}^{kt}) \geq 0 & k = i \\ \frac{(-1)^{k+t}}{\det(\mathbf{W})} \mathbf{w}_i \tilde{\mathbf{d}}^{kt}_{i} \geq 0 & k > i \end{cases} \quad (11)$$

where $\mathbf{w}_i = [w_{i1}, \ldots, w_{im}]$, $\tilde{\mathbf{d}}^{kt}_i = [\tilde{d}^{kt}_{i1}, \ldots, \tilde{d}^{kt}_{im}]^{\mathrm{T}}$.

Note that for $k = i$, a_{tk} is independent of $\mathbf{w}_i$ and $\det(\mathbf{M}^{kt})$ keeps constant when $\mathbf{w}_i$ is updated. Therefore, in order to ensure nonnegativity of a_{tk} in each iteration, the algorithm must keeps the sign of $\det(\mathbf{W})$ the same to that of a_{tk}. Based on the assumptions about $\mathbf{A}$, we can conclude that $|\det(\mathbf{W})| = 1/|\det(\mathbf{A})| \geq 1$, where the equality holds if and only if $\mathbf{W} = \mathbf{A} = \mathbf{I}$. Therefore, to obtain the maximization of $|\det(\mathbf{W})|$, it is reasonable to initialize $\mathbf{W} = \mathbf{I}$. If so, $\det(\mathbf{W}) > 0$ can be satisfied in every following iteration and (11) can be simplified as

$$\begin{cases} (-1)^{k+t} \mathbf{w}_i \tilde{\mathbf{d}}^{kt}_{i-1} \geq 0 & k < i \\ (-1)^{k+t} \mathbf{w}_i \tilde{\mathbf{d}}^{kt}_{i} \geq 0 & k > i \end{cases}. \quad (12)$$

So, under the fully nonnegativity constraints to both of $\mathbf{A}$ and $\mathbf{S}$, (3) can be reformulated as

$$\max_{\mathbf{w}_i} \quad \det(\mathbf{W}) = \sum_{j=1}^{m} (-1)^{i+j} w_{ij} \det(\mathbf{M}^{ij})$$
$$\text{s.t.} \begin{cases} (-1)^{k+t} \mathbf{w}_i \tilde{\mathbf{d}}^{kt}_{i-1} \geq 0 & \forall k < i, \forall t \\ (-1)^{k+t} \mathbf{w}_i \tilde{\mathbf{d}}^{kt}_{i} \geq 0 & \forall k > i, \forall t \\ \mathbf{w}_i \mathbf{1}_m = 1 \\ \mathbf{w}_i \mathbf{x}_q \geq 0, \forall q \end{cases} \quad (13)$$

where $\mathbf{x}_q$ denotes the qth column of $\mathbf{X}$.

In (13), both of the objective function and the constraints are linear functions with respect to $w_{i1}, \ldots, w_{im}$. Therefore, LP based methods can be used to solve this problem directly [14]. In this letter, the LP toolbox in Matlab software is invoked (Here is the construction of the posed LP problem for references: $x = \mathbf{w}_i^{\mathrm{T}}$, $f = -[(-1)^{i+1}\det(\mathbf{M}^{i1}), \ldots, (-1)^{i+m}\det(\mathbf{M}^{im})]^{\mathrm{T}}$, $\mathbf{A}\mathrm{eq} = \mathbf{1}_m^{\mathrm{T}}$, beq $= 1$, $\mathbf{A} = -[\mathbf{A}_{11}, \ldots, \mathbf{A}_{1m}, \mathbf{A}_{21}, \ldots, \mathbf{A}_{2m}, \mathbf{X}]^{\mathrm{T}}$, b $= \mathbf{0}_{m(m-1)+N}$, where for $t = 1, \ldots, m$, $\mathbf{A}_{1t} = [(-1)^{1+t}\tilde{\mathbf{d}}^{1t}_{i-1}, \ldots, (-1)^{i-1+t}\tilde{\mathbf{d}}^{(i-1)t}_{i-1}]$, and $\mathbf{A}_{2t} = [(-1)^{i+1+t}\tilde{\mathbf{d}}^{(i+1)t}_{i}, \ldots, (-1)^{m+t}\tilde{\mathbf{d}}^{mt}_{i}]$). Note that the number of the constraints can be reduced significantly by eliminating the redundant inequalities in prior by using extreme point finding algorithms [12].

III. Numerical Results

Two numerical experiments, in which one is with synthetic signals and another is with real-world images, are performed to evaluate the proposed FNCIVM algorithm. And its performance is compared with that of several conventional algorithms such as, the volume based nLCA-IVM [12], the ICA based algorithms: KernelICA [5], FastICA [3], and nonnegative ICA (NICA) [6], the NMF based algorithms: normal NMF [15] and FlexNMF [16]. The signal-to-interference ratio (SIR) index ($SIR_j = 10\log_{10}(E[h_j(t)]^2/E[h_j(t) - \hat{h}_j(t)]^2)$, $h_j(t)$ and $\hat{h}_j(t)$ denotes the source and the corresponding recovered source, respectively, and they are normalized with zero-mean

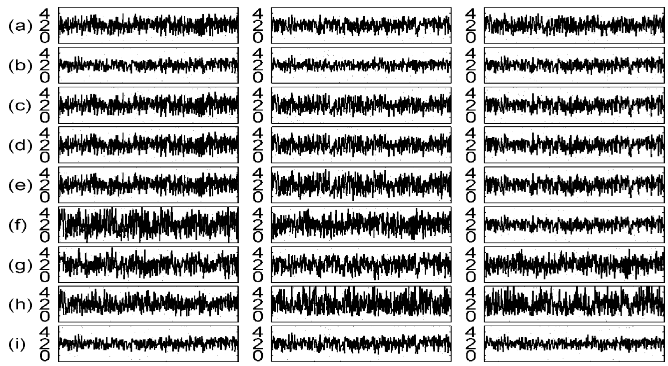

Fig. 1. Synthetic source signals, mixtures, and recoveries by FNCIVM, nLCA-IVM, KernelICA, FastICA, NICA, NMF and FlexNMF, respectively. (a) Sources. (b) Mixtures. (c) Recoveries by FNCIVM. (d) Recoveries by nLCA-IVM. (e) Recoveries by KernelICA. (f) Recoveries by FastICA. (g) Recoveries by NICA. (h) Recoveries by NMF. (i) Recoveries by FlexNMF.

TABLE I
TRUE VALUE AND THE ESTIMATIONS OF THE MIXING MATRIX IN THE EXPERIMENTS OF SYNTHETIC SIGNALS

	True value and the estimations of $\mathbf{A}$								
	The first row			The second row			The third row		
True	0.3448	0.7636	0.0856	0.5762	0.4592	0.7095	0.0840	0.3812	0.0151
FNCIVM	0.3726	0.7557	0.0658	0.5539	0.5128	0.6782	0.1070	0.3733	0.0000
nLCA-IVM	0.4313	0.7453	0.0173	0.5938	0.5058	0.6453	0.1360	0.3682	-0.0239
KernelICA	0.0908	0.0370	0.1776	0.1606	0.2219	0.1251	0.0251	0.0101	0.0868
FastICA	0.0045	0.0856	0.1936	0.1355	0.1998	0.1703	-0.0026	0.0145	0.0904
NICA	0.2183	0.7635	0.2114	0.4866	0.4492	0.5698	0.1021	0.3600	0.0383
NMF	0.2420	0.6079	0.3769	0.6945	0.7391	0.2554	0.0684	0.2570	0.1779
FlexNMF	0.7686	0.0000	0.0000	0.1182	0.0000	0.8102	0.1182	1.6041	0.0000

TABLE II
SIRS OF DIFFERENT ALGORITHMS IN THE EXPERIMENTS OF SYNTHETIC SIGNALS

Algorithms	SIR_1(dB)	SIR_2(dB)	SIR_3(dB)	Mean(dB)
FNCIVM	19.5915	18.8911	21.4939	19.9922
nLCA-IVM	19.5915	13.1525	21.4939	18.0793
KernelICA	24.2549	20.8886	13.9000	19.6812
FastICA	12.3610	18.4885	10.7110	13.8535
NICA	9.0399	-	-	-
NMF	-	-	-	-
FlexNMF	-	13.1076	-	-

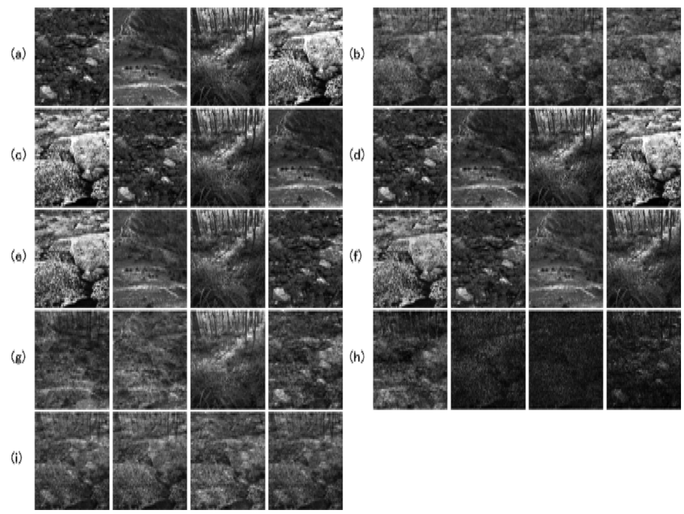

Fig. 2. Real-world source images, mixtures, and recoveries by FNCIVM, nLCA-IVM, KernelICA, FastICA, NICA, NMF and FlexNMF, respectively. (a) Sources. (b) Mixtures. (c) Recoveries by FNCIVM. (d) Recoveries by nLCA-IVM. (e) Recoveries by KernelICA. (f) Recoveries by FastICA. (g) Recoveries by NICA. (h) Recoveries by NMF. (i) Recoveries by FlexNMF.

and unit variance) is used to evaluate the performance on source recoveries [17]. And the Amari performance index (PI) in [18] is used to evaluate the performance on the mixing matrix estimation.

A. Synthetic Signals

Three random nonnegative signals are used to evaluate the proposed FNCIVM in this experiment [see Fig. 1(a)]. We have made 50 Monte Carlo experiments with various, random nonnegative mixing matrices, and the averaged elapsed times (s) of the above algorithms FNCIVM, nLCA-IVM, KernelICA, FastICA, NICA, NMF, FlexNMF are 0.9016, 1.4094, 3.9859, 0.0969, 0.7703, 0.1397, 0.1422, respectively. The proposed FNCIVM is faster than nLCA-IVM, although there exist slightly more constraints. The reason may be that only one LP problem needs to be solved in each iteration for FNCIVM. For simplicity, here we only analyze the result with one rand nonnegative matrix for example (the row labeled "True" in Table I shows this mixing matrix and Fig. 1(b) shows the corresponding mixtures).

At first, the estimated precisions of the mixing matrix are compared. Table I shows the estimated mixing matrices by the proposed FNCIVM, the mentioned nLCA-IVM, KernelICA, FastICA, NICA, NMF, and FlexNMF, respectively (neglecting the indeterminacies of scale and permutation). And the corresponding PI indices are 17.4700, 13.9627, 15.3921, 9.3100, 3.8181, -0.4672, 4.1385, respectively. From these values and the results in Table I, one can see that FNCIVM indeed gives a nonnegative solution, i.e., the condition must be satisfied by the assumption. NMF, FlexNMF, NICA, and KernelICA also generate nonnegative results, but the corresponding PI indices are lower than that of FNCIVM. The estimated mixing matrices of nLCA-IVM and FastICA have negative elements, although their PI indices are comparable.

Then, the estimated precisions of the sources are compared. Table II shows the SIR indices of the mentioned algorithms above, where "Mean" denotes the averaged SIRs and "-" denotes that the SIR index is below 8 dB (in general, SIR levels below 8–12 dB thresholds are indicative a failure in obtaining the desired source separation [19], [20]). One can see that the proposed FNCIVM has the highest SIR index. Fig. 1(c)–(i) shows the estimated sources by these algorithms. By visual comparison, the proposed FNCIVM, together with nLCA-IVM, NICA, NMF, and FlexNMF, generates nonnegative recoveries. However, KernelICA and FastICA obtain the recoveries with some negative elements.

B. Real-World Images

Four real-world images [21] that are mixed by a nonnegative mixing matrix are as the sources for the evaluation of the proposed FNCIVM in this experiment. We mainly compare the precisions of the estimated sources. Fig. 2(a)–(i) show the sources, mixtures, and the recoveries by FNCIVM, nLCA-IVM, KernelICA, FastICA, NICA, NMF, FlexNMF, respectively. By visual comparison, the proposed FNCIVM performs the best.

TABLE III
SIRs OF DIFFERENT ALGORITHMS IN THE EXPERIMENTS OF REAL-WORLD IMAGES

Algorithms	SIR_1(dB)	SIR_2(dB)	SIR_3(dB)	SIR_4(dB)	Mean(dB)
FNCIVM	30.3965	15.8260	39.3318	36.4479	30.5005
nLCA-IVM	30.3965	12.0448	39.3318	36.4479	29.5553
KernelICA	21.2298	11.0855	15.7633	23.3414	17.8550
FastICA	18.3517	11.2013	20.1692	14.4022	16.0311
NICA	-	-	11.0429	-	-
NMF	-	-	-	-	-
FlexNMF	-	-	-	-	-

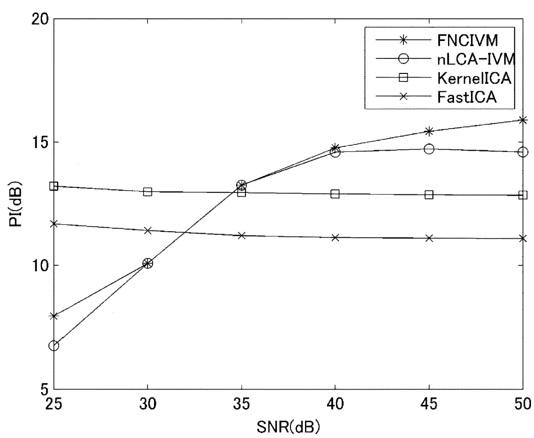

Fig. 3. PIs of different algorithms under different local noise levels.

Table III shows the corresponding SIR and the averaged SIR indices. From these values, the proposed FNCIVM is superior to the rest algorithms. The reason that FNCIVM is better than nLCA-IVM is that the former one obtains a *nonnegative* estimated mixing matrix, which satisfies the assumptions completely.

Furthermore, the cases of additive noises with different signal-noise-ratio (SNR) levels are tested. As in many scenarios, only local parts of the images may be corrupted by the noise, this case is mainly discussed in this Section. Fig. 3 shows the PIs of different algorithms under different Gaussian noise levels (about 25% of the pixels are corrupted by the noise). For better visual comparison, only the algorithms whose PI is above 5 dB are displayed. From Fig. 3, one can see that the volume based algorithms (FNCIVM and nLCA-IVM) tend to be affected more seriously than that of the ICA based algorithms (KernelICA and FastICA). The reason may be that the additive noise is Gaussian and it degrades the extreme point finding process in the volume based algorithms. However, in the case of lower noise, the volume based algorithms show some superiorities and FNCIVM performs better.

IV. CONCLUSION

In this letter, a BSS method is presented that has been posed as the fully nonnegative constrained iterative volume maximization problem. It can give solutions that satisfy the nonnegative assumptions both for the sources and for the mixing matrix, while the later nonnegativity cannot be satisfied in the conventional nLCA-IVM algorithm. Since this reason, the proposed can improve the nLCA-IVM algorithm for recovering the nonnegative sources from the nonnegative mixtures. Based on the numerical experiments with synthetic signals and with real-world images, the proposed FNCIVM are more interpretative than nLCA-IVM, and the source estimation precisions have also become considerably higher. Compared with the other conventional algorithms, such as KernelICA, FastICA, NICA, NMF, and FlexNMF, the proposed FNCIVM performs much better, especially for the case of weak noise.

REFERENCES

[1] A. Plaza *et al.*, "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, vol. 113, pp. S110–S122, 2009.
[2] W. H. Yang, D. Q. Dai, and H. Yan, "Feature extraction and uncorrelated discriminant analysis for high-dimensional data," *IEEE Trans. Knowl. Data Eng.*, vol. 20, pp. 601–614, 2008.
[3] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. Neural Netw.*, vol. 10, pp. 626–634, 1999.
[4] J. V. Stone, "Independent component analysis: An introduction," *Trends Cogn. Sci.*, vol. 6, pp. 59–64, 2002.
[5] F. R. Bach and M. I. Jordan, "Kernel independent component analysis," *J. Mach. Learn. Res.*, vol. 3, pp. 1–48, 2003.
[6] E. Oja and M. Plumbley, "Blind separation of positive sources by globally convergent gradient search," *Neural Comput.*, vol. 16, pp. 1811–1825, 2004.
[7] S. A. Astakhov, H. Stogbauer, A. Kraskov, and P. Grassberger, "Monte Carlo algorithm for least dependent non-negative mixture decomposition," *Anal. Chem.*, vol. 78, pp. 1620–1627, 2006.
[8] D. D. Lee and H. S. Seung, "Learning the parts of objects by nonnegative matrix factorization," *Nature*, vol. 401, pp. 788–791, 1999.
[9] A. Cichocki, R. Zdunek, and S. Amari, "New algorithm for non-negative matrix factorization in applications to blind source separation," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process.*, 2006.
[10] D. Donoho and V. Stodden, "When does non-negative matrix factorization give a correct decomposition into parts?," in *Adv. Neural Inf. Process. Syst.*, 2004.
[11] H. Laurberg, M. G. Christensen, M. D. Plumbley, L. K. Hansen, and S. H. Jensen, "Theorems on positive data: On the uniqueness of NMF," *Comput. Intell. Neurosci.*, vol. 2008, 2008.
[12] F. Y. Wang, C. Y. Chi, T. H. Chan, and Y. Wang, "Nonnegative least-correlated component analysis for separation of dependent sources by volume maximization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, pp. 875–888, 2010.
[13] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge Univ. Press, 1985.
[14] S. Mehrotra, "On the implementation of a Primal-Dual interior point method," *SIAM J. Optim.*, vol. 2, pp. 575–601, 1992.
[15] D. D. Lee and H. S. Seung, "Algorithms for nonnegative matrix factorization," in *Adv. Neural Inf. Process. Syst.*, 2001.
[16] A. Cichocki, A. H. Phan, R. Zdunek, and L. Q. Zhang, "Flexible component analysis for sparse, smooth, nonnegative coding or representation," *Lecture Notes in Computer Science*, vol. 4984, pp. 811–820, 2008.
[17] A. Cichocki and S. Amari, *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*. New York: Wiley, 2003.
[18] F. Yin, T. Mei, and J. Wang, "Blind-source separation based on decorrelation and nonstationarity," *IEEE Trans. Circuits Syst. I*, vol. 54, pp. 1150–1158, 2007.
[19] C. F. Caiafa and A. N. Proto, "Separation of statistically dependent sources using an L2-distance non-Gaussianity measure," *Signal Process.*, vol. 86, pp. 3404–3420, 2006.
[20] R. Boscolo, H. Pan, and V. P. Roychowdhury, "Independent component analysis based on nonparametric density estimation," *IEEE Trans. Neural Netw.*, vol. 15, pp. 55–65, 2004.
[21] [Online]. Available: http://www.cs.helsinki.fi/u/phoyer/