Turbo Source Extraction Algorithm and Noncancellation Source Separation Algorithms by Kurtosis Maximization

Chong-Yung Chi, *Senior Member, IEEE*, and Chun-Hsien Peng

***Abstract*—The kurtosis maximization criterion has been effectively used for blind spatial extraction of one source from an instantaneous mixture of multiple non-Gaussian sources, such as the kurtosis maximization algorithm by Ding and Nguyen, and the fast kurtosis maximization algorithm (FKMA) by Chi and Chen. By empirical studies, we found that the smaller the normalized kurtosis magnitude of the extracted source signal, the worse the performance of these algorithms. In this paper, with the assumption that each source is a non-Gaussian linear process, a novel blind source extraction algorithm, called turbo source extraction algorithm (TSEA), is proposed. The ideas of the TSEA are to exploit signal temporal properties for increasing the normalized kurtosis magnitude, and to apply spatial and temporal processing in a cyclic fashion to improve the signal extraction performance. The proposed TSEA not only outperforms the FKMA, but also shares the convergence and computation advantages enjoyed by the latter. This paper also considers the extraction of multiple sources, also known as source separation, by incorporating the proposed TSEA into the widely used multistage successive cancellation (MSC) procedure. A problem with the MSC procedure is its susceptibility to error propagation accumulated at each stage. We propose two noncancellation multistage (NCMS) algorithms, referred to as NCMS-FKMA and NCMS-TSEA, that are free from the error propagation effects. Simulation results are presented to show that the NCMS-TSEA yields substantial performance gain compared with some existing blind separation algorithms, together with a computational complexity comparison. Finally, we draw some conclusions.**

***Index Terms*—Blind source separation (BSS), kurtosis maximization, multistage successive cancellation (MSC) procedure, noncancellation multistage (NCMS) algorithms, turbo source extraction algorithm (TSEA).**

I. INTRODUCTION

BLIND source separation (BSS) (or independent component analysis) has been a widely known problem in science and engineering areas such as array signal processing, wireless communications, and biomedical signal processing [1]–[5]. With a given set of P sensor measurements, denoted as a $P \times 1$ vector $\mathbf{x}[n] = [x_1[n], x_2[n], \ldots, x_P[n]]^{\mathrm{T}}$, the BSS problem is to extract K unknown source signals, denoted as a $K \times 1$ vector $\mathbf{s}[n] = [s_1[n], s_2[n], \ldots, s_K[n]]^{\mathrm{T}}$, based on the following instantaneous (or memoryless) multiple-input multiple-output (MIMO) model

$$\mathbf{x}[n] = \mathbf{A}\mathbf{s}[n] + \mathbf{w}[n] = \sum_{i=1}^{K} \mathbf{a}_i s_i[n] + \mathbf{w}[n] \tag{1}$$

where $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_K]$ is an unknown $P \times K$ mixing matrix and $\mathbf{w}[n] = [w_1[n], w_2[n], \ldots, w_P[n]]^{\mathrm{T}}$ is a $P \times 1$ noise vector. There have been a number of BSS algorithms reported in the open literature basically including algorithms using second-order statistics (SOS) (known as correlations) [3]–[5], [7], [8], algorithms using higher order statistics (HOS) (known as cumulants) [1], [2], [4], [6], [9]–[16], [18] and a variety of linear and nonlinear algorithms using principles such as maximum-likelihood method and maximum entropy [1], [2], [19], or using characteristics and features of either of source signals and the mixing matrix such as nonstationarity and nonnegativity or their combinations [20]–[24].

Some statistical assumptions about the sources $\mathbf{s}[n]$ and some conditions about the unknown mixing matrix $\mathbf{A}$ are required by most existing algorithms using either SOS or HOS. For instance, the source signals $\mathbf{s}[n]$ are required to be temporally colored and spatially uncorrelated with different power spectra by SOS based algorithms such as the algorithm for multiple unknown signals extraction (AMUSE) proposed by Tong *et al.* [4], [7], the second-order blind identification (SOBI) algorithm proposed by Belouchrani *et al.* [5], and the matrix-pencil approach proposed by Chang *et al.* [8]. The AMUSE and SOBI algorithm further require $P > K$ and the noise correlation matrix given or estimated in advance, while the matrix-pencil approach requires $P \geq K$ instead of $P > K$ without need of the noise correlation matrix in the mean time [8]. On the other hand, algorithms using HOS generally do not require temporal conditions or assumptions about the sources $\mathbf{s}[n]$. For instance, Cardoso's [10] fourth-order blind identification (FOBI) algorithm (which ignores the noise effect) and Tong, Liu, Soon, and Huang's [4] extended FOBI (EFOBI) algorithm (which takes the noise effect into account) using HOS require that source signals $\mathbf{s}[n]$ be mutually independent with distinct fourth-order moments, and that $P > K$. Hyvärinen and Oja's [2], [11] fast fixed-point algorithm (also called the FastICA) using kurtosis (a fourth-order

Manuscript received January 31, 2005; revised September 26, 2005. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Ercan E. Kuruoglu. This work was supported by the National Science Council, R.O.C., under Grant NSC 93-2213-E-007-079 and Industrial Technology Research Institute under Grant S5-93002. This work was partly presented at the Third IEEE International Symposium on Signal Processing and Information Technology, Darmstadt, Germany, December 14–17, 2003 and partly presented at the Third IEEE Sensor Array and Multichannel Signal Processing Workshop, Barcelona, Spain, July 18–21, 2004.

C.-Y. Chi is with the Department of Electrical Engineering and Institute of Communications Engineering, National Tsing Hua University, Hsinchu, Taiwan 30013, R.O.C. (e-mail: cychi@ee.nthu.edu.tw).

C.-H. Peng is with the Department of Electrical Engineering and Institute of Communications Engineering, National Tsing Hua University, Hsinchu, Taiwan 30013, R.O.C. (e-mail: d905610@oz.nthu.edu.tw).

Digital Object Identifier 10.1109/TSP.2006.877640

cumulant), and Ding and Nguyen's kurtosis maximization algorithm (KMA) [12], [14] require that source signals $\mathbf{s}[n]$ be non-Gaussian (either temporally independent identically distributed (i.i.d.) or colored) and mutually independent, and that $P \geq K$.

The KMA works well, but it is basically a gradient-type iterative algorithm which is not very computationally efficient. Chi and Chen [15], [16] then proposed a fast kurtosis maximization algorithm (FKMA) which is computationally efficient with a super-exponential convergence rate. Both KMA and FKMA only involve spatial processing, regardless of whether the source signals are colored or not. As reported in [12], we also empirically found that the smaller the normalized kurtosis magnitudes of source signals, the worse the performance of the KMA and FKMA. This observation motivates us to use an additional process, namely temporal process (linear filter), to transform the original source signals into temporally processed source signals with higher normalized kurtosis magnitudes, thereby improving the signal extraction performance.

In the paper, we propose a novel blind source extraction algorithm also by kurtosis maximization, referred to as turbo source extraction algorithm (TSEA), which extracts one source signal through a spatial processing (for source extraction), and a temporal processing (for conversion of the extracted source into a filtered source with larger normalized kurtosis magnitude) cyclically. For the extraction of all the unknown sources using either the FKMA or the proposed TSEA, the widely used multistage successive cancellation (MSC) procedure [15], [16], [25]–[27] has been an effective approach in spite of error propagation effects accumulated from stage to stage [27]. Furthermore, two new noncancellation multistage (NCMS) BSS algorithms, referred to as NCMS-FKMA and NCMS-TSEA, which outperform the corresponding BSS algorithms involving the MSC procedure, are also proposed together with some simulation results including a computational load comparison, in addition to a performance comparison with some existing BSS algorithms through simulation.

The remaining parts of this paper are organized as follows. A review of the FKMA and the MSC procedure is presented in Section II together with some performance observations of the FKMA. Section III presents the proposed TSEA. Then the proposed NCMS-FKMA and NCMS-TSEA are presented in Section IV together with some analytic results. Then, some simulation results are provided to verify the efficacy of the proposed algorithms in Section V. Finally, some conclusions are drawn in Section VI.

II. Review of FKMA and MSC Procedure

For ease of later use, let us define the following notations:

$*$	convolution operation of discrete-time (scalar, vector, or matrix) signals;
$E\{\cdot\}$	expectation operator;
$\|\| \cdot \|\|$	Euclidean norm of vectors or matrices;
$\mathbf{0}_P$	$P \times 1$ zero vector;
$\mathbf{I}_P$	$P \times P$ identity matrix;
Superscript "$*$"	complex conjugation;
Superscript "T"	transpose of vectors or matrices;
Superscript "H"	complex conjugate transpose (Hermitian) of vectors or matrices;
$\mathcal{R}(\mathbf{B})$	range space of matrix $\mathbf{B}$;
$\mathcal{N}(\mathbf{B})$	null space of matrix $\mathbf{B}$;
$\mathrm{rank}(\mathbf{B})$	dimension of $\mathcal{R}(\mathbf{B})$;
$\mathrm{cum}\{z_1, z_2, z_3, z_4\}$	fourth-order joint cumulant of random variables z_1, z_2, z_3, and z_4;
$C_4\{z\} = \mathrm{cum}\{z_1 = z, z_2 = z, z_3 = z^*, z_4 = z^*\}$	*kurtosis* of random variable z;
SNR $\triangleq E\{\|\| \mathbf{x}[n] - \mathbf{w}[n] \|\|^2\} / E\{\|\| \mathbf{w}[n] \|\|^2\}$	signal-to-noise ratio.

Some general assumptions that are also made in the paper for the MIMO system model given by (1) are as follows:

A1) The unknown $P \times K$ mixing matrix $\mathbf{A}$ is of full column rank with $P \geq K$ (i.e., $\mathrm{rank}(\mathbf{A}) = K$), and K is known *a priori*.

A2) Each source signal $s_i[n], i \in \{1, 2, \ldots, K\}$, is a non-Gaussian linear process, i.e., $s_i[n]$ can be modeled as the output of a causal stable linear time-invariant system $b_i[n]$ as follows:

$$s_i[n] = b_i[n] * u_i[n] = \sum_{k=0}^{\infty} b_i[k] u_i[n-k] \tag{2}$$

where $u_i[n]$ is a stationary zero-mean non-Gaussian i.i.d. process with nonzero (real) kurtosis [9] given by

$$\begin{aligned} C_4\{u_i[n]\} &= \mathrm{cum}\{u_i[n], u_i[n], u_i^*[n], u_i^*[n]\} \\ &= E\{|u_i[n]|^4\} - 2(E\{|u_i[n]|^2\})^2 \\ &\quad - |E\{u_i^2[n]\}|^2 \end{aligned} \tag{3}$$

and $u_i[n]$ and $u_j[n]$ are statistically independent for all $i \neq j$.

A3) $\mathbf{w}[n]$ is zero-mean Gaussian, and statistically independent of $\mathbf{s}[n]$.

Let $\mathbf{v}$ be a $P \times 1$ *source extraction filter* (a spatial filter) with the input being $\mathbf{x}[n]$. Its output is then

$$e[n] = \mathbf{v}^{\mathrm{T}} \mathbf{x}[n]. \tag{4}$$

The source separation process includes the design of a set of K source extraction filters, each of which extracts a distinct source signal. Next, let us briefly review the FKMA proposed by Chi and Chen [15], [16], which is basically for the extraction of one of the K source signals, i.e., $e[n] = \hat{s}_k[n]$ where

$k \in \{1, 2, \ldots, K\}$, followed by some observations about its performance.

A. FKMA

The FKMA is an iterative algorithm for finding the optimum spatial filter $\mathbf{v}$ by maximizing the following objective function [12], [14]:

$$J(\mathbf{v}) = J(e[n]) = \frac{|C_4\{e[n]\}|}{E^2\{|e[n]|^2\}} \tag{5}$$

which is also the magnitude of normalized kurtosis of $e[n]$. Given $\mathbf{v}^{(i-1)}$ and $e^{(i-1)}[n]$ obtained at the $(i-1)$th iteration, $\mathbf{v}^{(i)}$ at the ith iteration is obtained through the following two steps.

Step 1) Update $\mathbf{v}^{(i)}$ by

$$\mathbf{v}^{(i)} = \frac{\mathbf{R}^{-1}\mathbf{d}\left(e^{(i-1)}[n]\right)}{\|\ \mathbf{R}^{-1}\mathbf{d}\left(e^{(i-1)}[n]\right)\ \|} \tag{6}$$

where

$$\mathbf{R} = \mathbf{R}_x^* = (E\{\mathbf{x}[n]\mathbf{x}^{\mathrm{H}}[n]\})^* \tag{7}$$

and

$$\begin{aligned}\mathbf{d}(e[n]) &= \mathrm{cum}\{e[n], e[n], e^*[n], \mathbf{x}^*[n]\} \\ &= E\{|e[n]|^2 e[n]\mathbf{x}^*[n]\} - 2E\{|e[n]|^2\}E\{e[n]\mathbf{x}^*[n]\} \\ &\quad - E\{e^2[n]\}E\{e^*[n]\mathbf{x}^*[n]\}.\end{aligned} \tag{8}$$

Then obtain the associated $e^{(i)}[n] = (\mathbf{v}^{(i)})^{\mathrm{T}}\mathbf{x}[n]$.

Step 2) If $J(\mathbf{v}^{(i)}) > J(\mathbf{v}^{(i-1)})$, go to the next iteration; otherwise update $\mathbf{v}^{(i)}$ through a gradient-type optimization algorithm, i.e.,

$$\mathbf{v}^{(i)} = \mathbf{v}^{(i-1)} + \rho \left.\frac{\partial J(\mathbf{v})}{\partial \mathbf{v}^*}\right|_{\mathbf{v}=\mathbf{v}^{(i-1)}} \tag{9}$$

such that $J(\mathbf{v}^{(i)}) > J(\mathbf{v}^{(i-1)})$ where ρ is a step size parameter, and then obtain the associated $e^{(i)}[n] = (\mathbf{v}^{(i)})^{\mathrm{T}}\mathbf{x}[n]$.

An initial condition $\mathbf{v}^{(0)}$ (which can be arbitrarily chosen with $\|\mathbf{v}^{(0)}\| = 1$) is needed to initialize the FKMA, and $\mathbf{v}^{(i)}$ is usually updated in Step 1) before convergence. It can be easily shown, from (5), that

$$\frac{\partial J(\mathbf{v})}{\partial \mathbf{v}^*} = 2J(\mathbf{v}) \cdot \left\{\frac{1}{C_4\{e[n]\}} \cdot \mathbf{d}(e[n]) - \frac{1}{E\{|e[n]|^2\}} \cdot \mathbf{R}\mathbf{v}\right\}. \tag{10}$$

Note, from (8), (9), and (10), that $\mathbf{d}(e^{(i-1)}[n])$ required for computing $\partial J(\mathbf{v})/\partial \mathbf{v}^*$ (see (10)) in Step 2) has been obtained in Step 1), and the correlation matrix $\mathbf{R}_x$ is the same at each iteration, indicating simple and straightforward computation for updating $\mathbf{v}^{(i)}$ in Step 2). The efficacy of the FKMA is supported by the following two facts.

Fact 1: With the assumptions A1) and A2), and the noise-free assumption, $J(e[n])$ given in (5) attains maximum (either locally or globally) if and only if

$$e[n] = \alpha_k s_k[n] \tag{11}$$

where α_k is an unknown nonzero constant and the unknown integer $k \in \{1, 2, \ldots, K\}$ [14].

Fact 2: The FKMA is also a hybrid algorithm using the super-exponential algorithm [17], [18] in Step 1 for fast convergence (basically with super-exponential convergence rate) and a gradient-type algorithm in Step 2) for the guaranteed convergence for finite data length N and finite SNR.

In practice, sample cumulants $\hat{C}_4\{e[n]\}$ and $\hat{\mathbf{d}}(e[n])$, and sample correlation matrix $\hat{\mathbf{R}}_x$, which are consistent estimates regardless of the SNR [28], are used by the FKMA. It is widely known that the performance of any source extraction (or separation) algorithms, including the FKMA is better either for higher SNR or for larger data length N. Furthermore, we empirically found that performance of the FKMA is also dependent upon temporal properties of source signals to be presented next.

It can be easily shown by A2) that

$$C_4\{s_i[n]\} = C_4\{u_i[n]\} \sum_{m=0}^{\infty} |b_i[m]|^4 \tag{12}$$

which leads to

$$J(s_i[n]) = \gamma(s_i[n]) \cdot J(u_i[n]) \quad \text{(by (5))} \tag{13}$$

where

$$\gamma(s_i[n]) = \frac{J(s_i[n])}{J(u_i[n])} = \frac{\sum_{m=0}^{\infty} |b_i[m]|^4}{\left(\sum_{m=0}^{\infty} |b_i[m]|^2\right)^2} \leq 1 \tag{14}$$

which is also an "entropy measure" of the stable sequence $b_i[n]$ [27] (with larger $\gamma(s_i[n])$ indicating smaller entropy of $b_i[n]$, and $\gamma(s_i[n]) = 1$ for $b_i[n] = \beta\delta[n-\tau]$, a unit sample sequence, for all $\beta \neq 0$ and integer τ). Note that the larger the value of $\gamma(s_i[n])$, the larger the value of $J(s_i[n])$. In other words, $\gamma(s_i[n])$ can be thought of as a measure of distance of $s_i[n]$ from a Gaussian process [29]. These characteristics of the non-Gaussian source signal $s_i[n]$ account for the following empirical observation [12].

Observation 1: The FKMA itself is an exclusive spatial processing algorithm. The smaller the value of $\gamma(s_i[n])$ (i.e., the smaller the distance of $s_i[n[$ from a Gaussian process), the worse the performance of the FKMA for finite SNR and finite data length N (i.e., the lower estimation accuracy of the extracted source $\hat{s}_i[n]$), although the asymptotic performance of the FKMA as $N \to \infty$ and SNR $\to \infty$ is independent of the value of $\gamma(s_i[n])$ by Fact 1.

Some simulation results to justify Observation 1 will be presented in Section V later. Next, let us review the MSC procedure for the extraction of all the K source signals using the FKMA.

B. MSC Procedure

Estimates $\hat{s}_1[n], \hat{s}_2[n], \ldots, \hat{s}_K[n]$ (possibly in a nonsequential order) of all the source signals can be obtained using the FKMA through the MSC procedure [15], [16], [25], [27] that includes the following two steps at each stage:

T1) Obtain a source signal estimate $\hat{s}_k[n]$ (where k is unknown) using a source extraction algorithm, and then obtain the associated channel estimate $\hat{\mathbf{a}}_k$ (see (1)) given by

$$\hat{\mathbf{a}}_k = \frac{E\{\mathbf{x}[n]\hat{s}_k^*[n]\}}{E\{|\hat{s}_k[n]|^2\}}. \tag{15}$$

T2) Update $\mathbf{x}[n]$ by $\mathbf{x}[n] - \hat{\mathbf{a}}_k \hat{s}_k[n]$, namely, cancellation of the contribution of $\hat{s}_k[n]$ from $\mathbf{x}[n]$.

Note that the above MSC procedure is also equivalent to the Gram–Schmidt orthogonalization method reported [14] which, however, never involves the estimation of the mixing matrix $\mathbf{A}$. The resultant BSS algorithm that uses the FKMA in T1) is referred to as MSC-FKMA. Note that imperfect cancellation in T2) usually results in error propagation effects accumulated in the ensuing stages. Therefore, as any other source separation algorithms involving the MSC procedure, an inevitable disadvantage of the MSC-FKMA is given in the following remark:

R1) The estimates $\hat{s}_k[n]'$s obtained at later stages are usually less accurate due to error propagation effects from stage to stage.

III. Turbo Source Extraction Algorithm

By Observation 1, the FKMA, which only involves spatial processing, performs better for larger $\gamma(s_k[n])$, and meanwhile $1 - \gamma(s_k[n])$ characterizes the temporal property of the source signal $s_k[n]$ about its distance from an i.i.d. non-Gaussian process. Advisable temporal processing of $s_k[n]$ can lead to the processed source $s'_k[n]$ with $\gamma(s'_k[n]) > \gamma(s_k[n])$. This motivates that the source extraction processing should include not only spatial processing but also temporal processing, although the measurement vector $\mathbf{x}[n]$ given by (1) is merely a spatial mixture of non-Gaussian sources.

Instead of the source extraction filter (spatial filter) used by the FKMA, the following $P \times 1$ *source separation filter*:

$$\mathbf{v}_{\mathrm{TSEA}}[n] = \mathbf{v}v[n] \tag{16}$$

is considered where $\mathbf{v}$ is a $P \times 1$ spatial filter (source extraction filter) and $v[n]$ is a single-input single-output (SISO) temporal filter of order equal to L. Then the output $\varepsilon[n]$ of the source separation filter $\mathbf{v}_{\mathrm{TSEA}}[n]$ is

$$\begin{aligned}\varepsilon[n] &= \mathbf{v}_{\mathrm{TSEA}}^{\mathrm{T}}[n] * \mathbf{x}[n] = \mathbf{v}^{\mathrm{T}}\mathbf{y}[n] \\ &= \hat{s}_k[n] * v[n] = \sum_{l=0}^{L} \hat{s}_k[n-l]v[l]\end{aligned} \tag{17}$$

in which

$$\mathbf{y}[n] = \mathbf{x}[n] * v[n] \tag{18}$$
$$\hat{s}_k[n] = \mathbf{v}^{\mathrm{T}}\mathbf{x}[n]. \tag{19}$$

Note from (17) and (19) that the spatial filter $\mathbf{v}$ is used to extract a colored source signal $\hat{s}_k[n]$ (i.e., removing the spatial interference due to the mixing matrix $\mathbf{A}$), and $v[n]$ is a filter to further process $\hat{s}_k[n]$ such that its output $\varepsilon[n]$ is a better approximation to an i.i.d. non-Gaussian process than $\hat{s}_k[n]$ (i.e., $\gamma(\varepsilon[n]) > \gamma(\hat{s}_k[n])$).

Again, the criterion $J(\varepsilon[n])$ given by (5) is used for the design of the optimum source separation filter $\mathbf{v}_{\mathrm{TSEA}}[n] = \mathbf{v}v[n]$. Let $\boldsymbol{\theta} = [\mathbf{v}^{\mathrm{T}}, v[0], v[1], \ldots, v[L]]^{\mathrm{T}}$. However, it is almost formidable to find a closed-form solution for the optimum $\boldsymbol{\theta}$ by solving $\partial J(\varepsilon[n])/\partial\boldsymbol{\theta} = \mathbf{0}_{P+L+1}$. Surely, one can resort to gradient-type iterative algorithms (which are not very computationally efficient) for finding the optimum $\boldsymbol{\theta}$. Instead, the proposed TSEA is a cyclically iterative algorithm as shown in Fig. 1, which basically makes use of the computationally efficient FKMA for the design of $\mathbf{v}$ and $v[n]$ (denoted as $\mathrm{FKMA}^{(\mathrm{s})}$ and $\mathrm{FKMA}^{(\mathrm{t})}$ in Fig. 1). As shown in Fig. 1, the proposed TSEA consists of two steps at the ith cycle as follows.

S1)

a) Temporal Processing: Compute

$$\mathbf{y}^{(i-1)}[n] = \mathbf{x}[n] * v^{(i-1)}[n] \quad \text{(by (18))}. \tag{20}$$

b) Spatial Processing: Process $\mathbf{y}^{(i-1)}[n]$ using the $\mathrm{FKMA}^{(\mathrm{s})}$ (with $\mathbf{v}^{(i-1)}$ as the initial condition for $\mathbf{v}$) to obtain the optimum $\mathbf{v}^{(i)}$ and

$$\varepsilon[n] = (\mathbf{v}^{(i)})^{\mathrm{T}}\mathbf{y}^{(i-1)}[n] \quad \text{(by the first line of (17))}. \tag{21}$$

S2)

a) Spatial Processing: Compute

$$e[n] = \hat{s}_k[n] = (\mathbf{v}^{(i)})^{\mathrm{T}}\mathbf{x}[n] \quad \text{(by (19))} \tag{22}$$

where $k \in \{1, 2, \ldots, K\}$.

b) Temporal Processing: Find the optimum $\boldsymbol{v}_i = [v^{(i)}[0], v^{(i)}[1], \ldots, v^{(i)}[L]]^{\mathrm{T}}$ and

$$\varepsilon[n] = e[n] * v^{(i)}[n] = \boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{e}[n] \tag{23}$$

(by (22) and the second line of (17)) using the $\mathrm{FKMA}^{(\mathrm{t})}$ (with $\boldsymbol{v}_{i-1}$ as the initial condition for $\boldsymbol{v}$), where $\boldsymbol{e}[n] = [e[n], e[n-1], \ldots, e[n-L]]^{\mathrm{T}}$.

An initial condition for $\mathbf{v}^{(0)}$ (spatial processing) and one for $\boldsymbol{v}_0$ (temporal processing), which can be arbitrarily chosen with

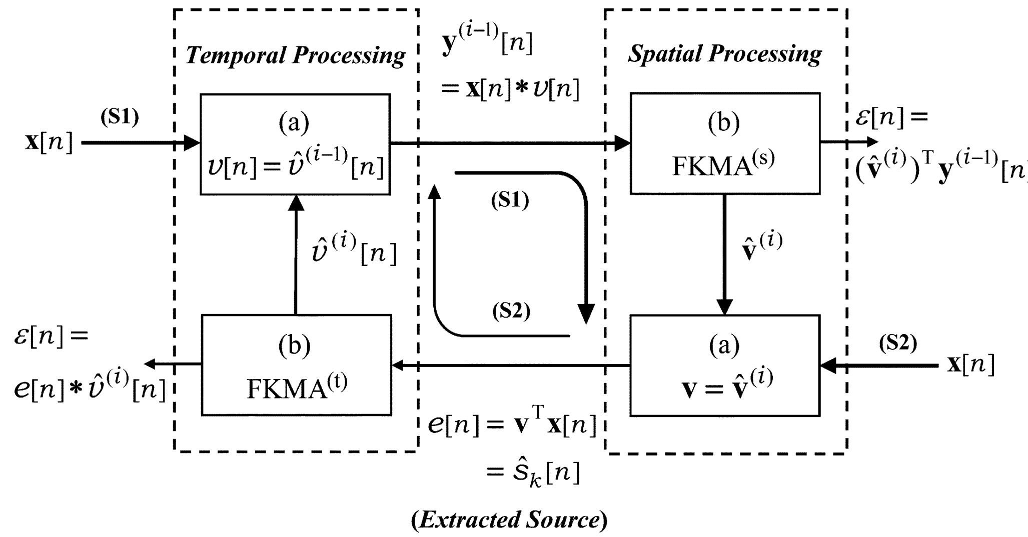

Fig. 1. Signal processing procedure of the proposed TSEA.

unity norm, are needed to initialize the proposed TSEA. Two worthy remarks regarding the proposed TSEA are as follows:

R2) The FKMA is employed in a cyclic fashion like a turbine to process the measurement vector $\mathbf{x}[n]$ for finding both of the optimum spatial filter $\mathbf{v}$ in S1) and temporal filter $v[n]$ in S2), and meanwhile the objective function $J(\varepsilon[n])$ increases whenever either the spatial filter $\mathbf{v}$ or the temporal filter $v[n]$ is updated. Therefore, for each cycle, both of the FKMA$^{(s)}$ and FKMA$^{(t)}$ also converge at super-exponential rate by Fact 2.

R3) The obtained signal estimate $\hat{s}_k[n]$ given by (22), the output of the spatial filter $\mathbf{v}$ in S2), is the signal of interest rather than the output $\varepsilon[n]$ of the source separation filter $\mathbf{v}_{\mathrm{TSEA}}[n]$. Moreover, the proposed TSEA reduces to the FKMA (i.e., FKMA$^{(s)}$ in S1)) as $v[n] = \delta[n]$ and S2) is removed.

Next, let us discuss why the cyclic operation of spatial–temporal processing of the proposed TSEA performs better than the FKMA. The output of the optimum source separation filter $\mathbf{v}_{\mathrm{TSEA}}[n]$ obtained by the proposed TSEA at the ith cycle can be approximated as follows:

$$\begin{aligned}\varepsilon[n] &= e[n] * v^{(i)}[n] = \hat{s}_k[n] * v^{(i)}[n] \quad \text{(by (22) and (23))}\\ &\simeq (\alpha_k u_k[n] * b_k[n]) * v^{(i)}[n] \quad \text{(by (2) and (11))}\\ &= \alpha_k u_k[n] * g_k[n] \qquad (24)\end{aligned}$$

where α_k is an unknown nonzero constant and

$$g_k[n] = b_k[n] * v^{(i)}[n] \qquad (25)$$

is the overall signal model associated with $s_k[n]$. Increasing $J(\varepsilon[n])$ in (S2) is equivalent to increasing

$$\gamma(\varepsilon[n]) \simeq \frac{\sum_m |g_k[m]|^4}{\left(\sum_m |g_k[m]|^2\right)^2} \quad \text{(by (24) and (14))}.$$

Therefore, the SISO temporal filter $v[n]$ performs like a "higher-order whitening" filter because the larger the value of $\gamma(\varepsilon[n])$, the better approximation of $\varepsilon[n]$ to an i.i.d. non-Gaussian signal (or the better approximation of $g_k[n]$ to $\alpha\delta[n-\tau]$). As L is sufficiently large, $v[n]$ becomes a deconvolution filter or a linear equalizer (which tries to remove the effects of "channel" $b_k[n]$), and its output $\varepsilon[n]$ turns out to be an estimate of the i.i.d. signal $u_k[n]$ (see (2)).

Another interpretation why the performance of the proposed TSEA is superior to that of the FKMA is as follows. Note that $\mathbf{y}[n]$ given by (18) is actually a mixture of filtered sources $s'_i[n] = s_i[n] * v[n], i = 1, 2, \ldots, K$ with the same mixing matrix $\mathbf{A}$ associated with $\mathbf{x}[n]$ (see (1)). In the spatial processing of S1) (part b) of S1)), the extraction of $\varepsilon[n] \simeq s'_k[n]$ by processing $\mathbf{y}[n]$ is much more effective than the extraction of $s_k[n]$ by processing $\mathbf{x}[n]$ using the FKMA because of $\gamma(\varepsilon[n]) > \gamma(s_k[n])$, thus leading to more accurate estimate $\hat{s}_k[n]$ in S2). Let us conclude this section with the following two remarks:

R4) The performance gain of the TSEA reaches the maximum as long as the order L (a parameter under our choice) of the temporal filter is sufficiently large. The chosen value for L is sufficient as $J(\varepsilon[n])$ does not increase significantly because the larger the value of $J(\varepsilon[n])$, the better approximation of $\varepsilon[n]$ to an i.i.d. non-Gaussian signal. On the other hand, the asymptotic performance of FKMA approaches that of the TSEA as $N \to \infty$ and SNR $\to \infty$ by Fact 1.

R5) The proposed TSEA can surely be employed to extract all the source signals through the MSC procedure. The resultant BSS algorithm that uses the TSEA in T1) of the MSC procedure (see Section II-B) is referred to as MSC-TSEA, and it also outperforms the MSC-FKMA, at the extra expense of the temporal processing at each stage.

IV. Noncancellation Multistage Source Separation Algorithms

In view of the performance degradation of BSS algorithms involving the MSC procedure due to error propagation effects as mentioned in R1), this section presents two multistage source separation algorithms without involving cancellation, referred to as NCMS-FKMA and NCMS-TSEA, respectively. Next let us concentrate on the NCMS-FKMA.

A. Proposed NCMS-FKMA

At each stage, the proposed NCMS-FKMA basically comprises a *preprocessing* of constrained source extraction and a processing of unconstrained source extraction using the FKMA. The former (preprocessing) designs the source extraction filter, denoted $\mathbf{v}_\ell$, to extract a distinct source by imposing some constraint on $\mathbf{v}_\ell$ instead of cancellation processing where the subscript "ℓ" denotes the stage number. The latter is nothing but the source extraction using the FKMA with $\mathbf{v}_\ell$ as the initial condition of the source extraction filter. Next, let us present the preprocessing.

Assume that at the end of stage $\ell-1$, we have extracted $\ell-1$ source signals and obtained the associated $\ell-1$ column estimates of $\mathbf{A}$ using (15), denoted as $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{\ell-1}$. Let $\mathbf{C}_\ell$ be a $P \times (\ell-1)$ matrix defined as

$$\mathbf{C}_\ell = [\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{\ell-1}], \quad \ell = 2, 3, \ldots, K \tag{26}$$

and $\mathbf{C}_\ell^\perp$ be a $P \times P$ projection matrix for which $\mathcal{R}(\mathbf{C}_\ell^\perp)$ is orthogonal to $\mathcal{R}(\mathbf{C}_\ell)$, and $\mathbf{C}_1^\perp = \mathbf{I}_P$. Assume that $\mathbf{C}_\ell$ is of full column rank, i.e., $\mathrm{rank}(\mathbf{C}_\ell) = \ell - 1$. By singular value decomposition (SVD), $\mathbf{C}_\ell$ can be expressed as [30]

$$\mathbf{C}_\ell = \mathbf{U}\boldsymbol{\Sigma}\bar{\mathbf{V}}^{\mathrm{H}} = \sum_{i=1}^{\ell-1} \sigma_i \mathbf{u}_i \bar{\mathbf{v}}_i^{\mathrm{H}} \tag{27}$$

where $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_P]$ is a $P \times P$ unitary matrix, $\bar{\mathbf{V}} = [\bar{\mathbf{v}}_1, \bar{\mathbf{v}}_2, \ldots, \bar{\mathbf{v}}_{\ell-1}]$ is an $(\ell-1) \times (\ell-1)$ unitary matrix, $\boldsymbol{\Sigma}$ is a $P \times (\ell-1)$ matrix with the (i,j)th entry $\{\boldsymbol{\Sigma}\}_{ij} = 0, i \neq j$, and $\{\boldsymbol{\Sigma}\}_{ii} = \sigma_i > 0, i = 1, 2, \ldots, \ell-1$ (singular values of $\mathbf{C}_\ell$). The projection matrix $\mathbf{C}_\ell^\perp$ is known as [30]

$$\mathbf{C}_\ell^\perp = (\mathbf{C}_\ell^\perp)^{\mathrm{H}} = \mathbf{U}_\ell \mathbf{U}_\ell^{\mathrm{H}} \tag{28}$$

where

$$\mathbf{U}_\ell = [\mathbf{u}_\ell, \mathbf{u}_{\ell+1}, \ldots, \mathbf{u}_P] \tag{29}$$

which is also a full-rank $P \times (P-\ell+1)$ matrix (i.e., $\mathrm{rank}(\mathbf{U}_\ell) = P - \ell + 1$).

The optimum spatial filter $\mathbf{v}_\ell$ to be designed in the preprocessing is

$$\begin{aligned}\mathbf{v}_\ell = \arg\max_{\mathbf{v}} \{ J(\mathbf{v}) = J(e[n]) : e[n] \\ = \mathbf{v}^{\mathrm{T}}\mathbf{x}[n], \mathbf{v}^{\mathrm{T}}\mathbf{C}_\ell = \mathbf{0}_{\ell-1}^{\mathrm{T}} \}\end{aligned} \tag{30}$$

where $J(\mathbf{v})$ and $\mathbf{C}_\ell$ were defined by (5) and (26), respectively. Finding the constrained optimal $\mathbf{v}_\ell$ given by (30) can be converted into an unconstrained optimization problem by the following transformation:

$$\mathbf{v}_\ell = (\mathbf{C}_\ell^\perp)^* \boldsymbol{v}_\ell \tag{31}$$

where

$$\boldsymbol{v}_\ell = \arg\max_{\boldsymbol{v}} \{ J(\boldsymbol{v}) = J(e[n]) : e[n] = \boldsymbol{v}^{\mathrm{T}} \mathbf{x}_\ell[n] \} \tag{32}$$

in which

$$\mathbf{x}_\ell[n] = \mathbf{C}_\ell^\perp \mathbf{x}[n]. \tag{33}$$

The associated optimum $e_\ell[n]$ is given by

$$e_\ell[n] = \boldsymbol{v}_\ell^{\mathrm{T}} \mathbf{x}_\ell[n] = \mathbf{v}_\ell^{\mathrm{T}} \mathbf{x}[n]. \tag{34}$$

The optimum $e_\ell[n]$ given by (34) obtained in the preprocessing is supported by the following theorem.

Theorem 1: Let S_ℓ be the set of all the extracted source signals up to stage $\ell-1$. With the assumptions $\mathcal{A}1$), $\mathcal{A}2$), and the noise-free assumption, the optimum $e_\ell[n]$ given by (34) is

$$e_\ell[n] = \alpha_k s_k[n] \tag{35}$$

where α_k is an unknown nonzero constant and $s_k[n] \notin S_\ell$. ■

The proof of Theorem 1 is presented in Appendix I. The proposed NCMS-FKMA is summarized as follows.

NCMS-FKMA:

F1) Set $\ell = 0$ and $\mathbf{C}_1^\perp = \mathbf{I}_P$.

F2) Update ℓ by $\ell + 1$. If $\ell \geq 2$, obtain $\mathbf{C}_\ell$ by (26) followed by its SVD, and then obtain $\mathbf{C}_\ell^\perp$ and $\mathbf{x}_\ell[n]$ by (28) and (33), respectively.

F3) a) Preprocessing (constrained source extraction): Obtain $\boldsymbol{v}_\ell$ and $e_\ell[n] = \boldsymbol{v}_\ell^{\mathrm{T}} \mathbf{x}_\ell[n]$ using the FKMA. Then obtain $\boldsymbol{a}_\ell$ using (15) (with $\hat{s}_k[n]$ replaced by $e_\ell[n]$ in (15)), and obtain $\mathbf{v}_\ell$ using (31).

b) Unconstrained source extraction: Obtain the optimum $\mathbf{v}$ and $e_\ell[n] = \mathbf{v}^{\mathrm{T}}\mathbf{x}[n]$ using the FKMA (with $\mathbf{v}^{(0)} = \mathbf{v}_\ell$ as the initial condition for $\mathbf{v}$).

F4) If $\ell < K$, go to $\mathcal{F}2$); otherwise, all the source signal estimates $\{e_\ell[n], \ell = 1, 2, \ldots, K\}$ have been obtained.

Three remarks about the proposed NCMS-FKMA are worth mentioning as follows.

R6) The FastICA using kurtosis [2], [11] obtains the prewhitened signal $\mathbf{z}[n] = \mathbf{D}\mathbf{x}[n]$ through eigendecomposition (or SVD) of $\mathbf{R}_x = E[\mathbf{x}[n]\mathbf{x}^{\mathrm{H}}[n]]$ such that $E\{\mathbf{z}[n]\mathbf{z}^{\mathrm{H}}[n]\} = \mathbf{I}_K$ and then performs source separation by kurtosis maximization. Contrasting with the FastICA using kurtosis, the SVD in F2) is for obtaining the projected data $\mathbf{x}_\ell[n]$ rather than any prewhitening process.

R7) The proposed NCMS-FKMA with only preprocessing F3a) alone (i.e., removal of F3b)) itself is also a BSS algorithm which can extract a distinct source (by Theorem 1). However, its performance depends on the estimation accuracies of the channel estimates $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{\ell-1}$ (used in the constraint of $\mathbf{v}^{\mathrm{T}}\mathbf{C}_\ell = \mathbf{0}_{\ell-1}^{\mathrm{T}}$ (see (30))). Although these channel estimates are never perfect for finite SNR and data length, F3a) provides a well-designed initial condition $\mathbf{v}_\ell$ for the unconstrained source extraction filter $\mathbf{v}$ in F3b).

R8) The unconstrained source extraction filter $\mathbf{v}$ in F3b) with a suitable initial condition $\mathbf{v}_\ell$ obtained by F3a) accordingly leads to one distinct source estimate $e_\ell[n]$ obtained at each stage neither involving cancellation nor imposing any constraints on the source extraction filter, as well as faster convergence of F3b) than F3a). Therefore, the proposed NCMS-FKMA outperforms the MSC-FKMA due to no error propagation effects at each stage, but moderately extra computational expense for the constrained source extraction F3a) is required. Moreover, only FKMA is employed for the source extraction in F3a) and F3b) at each stage, the proposed NCMS-FKMA is also a fast multistage BSS algorithm.

B. Proposed NCMS-TSEA

The NCMS-TSEA, which is basically obtained by replacing the FKMA used in $\mathcal{F}$3-a) and $\mathcal{F}$3-b) of the proposed NCMS-FKMA by the proposed TSEA, is given as follows.

NCMS-TSEA:

T1) Set $\ell = 0$ and $\mathbf{C}_1^{\perp} = \mathbf{I}_P$.

T2) Update ℓ by $\ell + 1$. If $\ell \geq 2$, obtain $\mathbf{C}_\ell$ by (26) followed by its SVD, and then obtain $\mathbf{C}_\ell^{\perp}$ and $\mathbf{x}_\ell[n]$ by (28) and (33), respectively.

T3) a) Preprocessing (constrained source extraction): Obtain $\mathbf{v}_{\mathrm{TSEA}}[n] = \boldsymbol{v}_\ell(\nu_\ell[n]), \varepsilon_\ell[n] = \mathbf{v}_{\mathrm{TSEA}}^{\mathrm{T}}[n] * \mathbf{x}_\ell[n]$, and $e_\ell[n] = \boldsymbol{v}_\ell^{\mathrm{T}}\mathbf{x}_\ell[n]$ using the proposed TSEA. Then obtain $\boldsymbol{a}_\ell$ using (15) (with $\hat{s}_k[n]$ replaced by $e_\ell[n]$ in (15)), and obtain $\mathbf{v}_\ell$ using (31).

b) Unconstrained source extraction: Obtain the optimum $\mathbf{v}_{\mathrm{TSEA}}[n] = \mathbf{v}v[n], \varepsilon_\ell[n] = \mathbf{v}_{\mathrm{TSEA}}^{\mathrm{T}}[n] * \mathbf{x}[n]$, and $e_\ell[n] = \mathbf{v}^{\mathrm{T}}\mathbf{x}[n]$ using the proposed TSEA (with $\mathbf{v}^{(0)} = \mathbf{v}_\ell$ and $v^{(0)}[n] = \nu_\ell[n]$ as the initial conditions for $\mathbf{v}$ and $v[n]$, respectively).

T4) If $\ell < K$, go to $\mathcal{T}$2); otherwise, all the source signal estimates $\{e_\ell[n], \ell = 1, 2, \ldots, K\}$ have been obtained.

Two worthy remarks regarding the proposed NCMS-TSEA are as follows.

R9) Remarks R7) and R8) regarding the constrained source extraction F3a) and unconstrained source extraction F3b) of the proposed NCMS-FKMA also apply to T3a) and T3b) of the proposed NCMS-TSEA which also shares the fast convergence of the proposed TSEA at each stage as presented in R2), so the performance of the proposed NCMS-TSEA is also superior to the MSC-TSEA at the extra computational expense for the constrained source extraction T3a).

R10) As the proposed MSC-TSEA performs better than the MSC-FKMA (see R5)), the proposed NCMS-TSEA also performs better than the proposed NCMS-FKMA for the same reasons as presented in Section III at the moderate expense of extra computational load for the temporal processing of the TSEA.

V. Simulation Results

To justify the efficacy of the proposed BSS algorithms, MSC-TSEA, NCMS-FKMA, and NCMS-TSEA, two parts of simulation results are to be presented. Section V-A focuses on the performance of the proposed NCMS-FKMA and NCMS-TSEA to manifest the superiority of the proposed TSEA to the FKMA, and Section V-B focuses on performance comparison of the proposed BSS algorithms with the existing MSC-FKMA, FastICA using kurtosis, SOBI algorithm, and AMUSE.

The i.i.d. $u_i[n]$'s used for generating $s_i[n]$'s were equiprobable random binary sequences of ± 1 and the noise vector $\mathbf{w}[n]$ was real, zero-mean, spatially independent, and white Gaussian with covariance $\sigma^2\mathbf{I}_P$. The synthetic $\mathbf{x}[n]$ was generated according to (1) and then processed by the BSS algorithms under test. The convergence criterion [32] used was

$$2 - 2\left|(\mathbf{v}^{(i-1)})^{\mathrm{T}} \cdot (\mathbf{v}^{(i)})^*\right| < 10^{-4}$$

whenever the FKMA was employed, where $\mathbf{v}^{(i)}$ was the spatial filter obtained at the ith iteration, and that for the TSEA was

$$\left|J\left(\varepsilon_i^{(\mathrm{S2})}[n]\right) - J\left(\varepsilon_i^{(\mathrm{S1})}[n]\right)\right| \Big/ J\left(\varepsilon_i^{(\mathrm{S1})}[n]\right) < 10^{-4}$$

where $\varepsilon_i^{(\mathrm{S1})}[n]$ and $\varepsilon_i^{(\mathrm{S2})}[n]$ were the $\varepsilon[n]$ obtained in S1) and S2), respectively, at the ith cycle. Moreover, the initial conditions for the spatial filter and the temporal filter used were $\mathbf{v}^{(0)} = [1, 1, \ldots, 1]^{\mathrm{T}}/\sqrt{P}$, and $v^{(0)}[n] = \delta[n]$, respectively, whenever they were not specified by the BSS algorithm under test.

Fifty independent runs were conducted for performance evaluation of each BSS algorithm. At the ith independent run, an optimum spatial filter $\hat{\mathbf{v}}$ and the associated source estimate $\hat{s}_k[n]$ were obtained. The estimate $\hat{s}_k[n]$ can be expressed as

$$\hat{s}_k[n] = \mathbf{f}_i^{\mathrm{T}}\mathbf{s}[n] + \varpi_i[n] \tag{36}$$

where $\varpi_i[n] = \hat{\mathbf{v}}^{\mathrm{T}}\mathbf{w}[n]$ and $\mathbf{f}_i = \mathbf{A}^{\mathrm{T}}\hat{\mathbf{v}}$ for the BSS algorithms without involving the MSC procedure, and that involving the MSC procedure can be found in [27, App. B]. The average output signal-to-interference-plus-noise ratio (SINR) associated with $\hat{s}_k[n]$ over the 50 independent runs can be calculated as

$$\text{output SINR}_k = \frac{1}{50}\sum_{i=1}^{50} \frac{|f[i,k]|^2 E\{|s_k[n]|^2\}}{\sum_{j=1,j\neq k}^{K} |f[i,j]|^2 E\{|s_j[n]|^2\} + E\{|\varpi_i[n]|^2\}} \tag{37}$$

where $f[i,j]$ is the jth entry of the $K \times 1$ vector $\mathbf{f}_i$. The total averaged output SINR defined as

$$\text{output SINR} = \frac{1}{K}\sum_{k=1}^{K} \text{output SINR}_k \tag{38}$$

is used as the performance index of each BSS algorithm.

A. Performance of NCMS-FKMA and NCMS-TSEA

The performance comparison of the proposed TSEA and the FKMA was conducted by virtue of the proposed NCMS-TSEA and NCMS-FKMA without involving any constraints and cancellation processes. In the simulation, a 5×4 real mixing matrix $\mathbf{A}$ and four MA models $b_i[n]$'s taken from [8] were used which are given as follows:

$$\mathbf{A} = \begin{bmatrix} 0.2380 & 0.2887 & -0.7120 & 0.4914 \\ 0.3397 & -0.7494 & -0.1157 & 0.2097 \\ 0.6107 & 0.4959 & 0.2661 & 0.2504 \\ 0.3558 & 0.2644 & -0.4216 & -0.6640 \\ -0.5731 & -0.1983 & -0.4807 & 0.4593 \end{bmatrix} \tag{39}$$

$$b_i[n] = \exp\left(-\frac{n+1}{10\cdot\lambda_i}\right), \quad n = 0, 1, \ldots, 5. \tag{40}$$

Four cases are considered as follows:

Case 1) Output SINR versus SNR for different data length ($N = 500, 1000$, and 1500), $\lambda_i = 0.5$ (or $\gamma(s_i[n]) = 0.2368$) for all i, and $L = 5$ (the order of the temporal filter $v[n]$).

Case 2) Output SINR versus different data length N for SNR $= 30$ dB, $\lambda_i = 1$ (or $\gamma(s_i[n]) = 0.1856$) for all i, and $L = 5$.

Case 3) Output SINR versus $\gamma(s_i[n]) = \gamma$ (or $\lambda_i = \lambda$) for all i, for SNR $= 30$ dB, $N = 2000$, and $L = 5$.

Case 4) Output SINR versus L, and the average of $J(\varepsilon_k[n])$ $((1/K)\sum_{k=1}^{K} J(\varepsilon_k[n]))$ versus L, for SNR $= 30$ dB, $N = 2000$, and $\lambda_i = \lambda = 1$ as well as 0.5 (i.e., $\gamma(s_i[n]) = \gamma = 0.1856$ as well as 0.2368) for all i.

The simulation results for the four cases are shown in Figs. 2 through 5, respectively. One can observe, from Figs. 2 and 3, that the proposed NCMS-TSEA (□) performs better than the proposed NCMS-FKMA (△) for different values of SNR and data length N. Their performance differences are larger for either higher SNR (by Fig. 2) or smaller N (by Fig. 3), and the performance of the proposed NCMS-TSEA has saturated for $N > 2 \times 10^3$, and the asymptotic performance of the proposed NCMS-FKMA approaches that of the proposed NCMS-TSEA as $N \to \infty$ for high SNR (SNR $= 30$ dB) (by Fig. 3). These results are consistent with Observation 1 and R4). On the other hand, one can observe, from Fig. 4, that the proposed NCMS-TSEA outperforms the proposed NCMS-FKMA for all the values of γ, and their performance differences are larger for smaller γ. Moreover, the performance of the former is robust to the value of γ, while that of the latter is sensitive to the value of γ as mentioned in Observation 1. Finally, one can see, from Fig. 5(a), that the performance of the proposed NCMS-TSEA has basically reached a plateau for $L \geq 2$ implying that the optimum performance has been attained as long as L is sufficiently large, and from Fig. 5(b), that the average of $J(\varepsilon_k[n])$ increases with L while its increase is nonsignificant

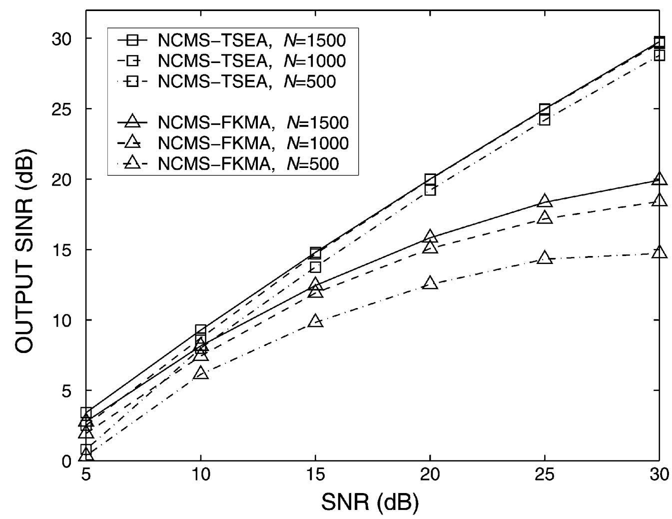

Fig. 2. Simulation results (output SINR versus SNR) of Case 1).

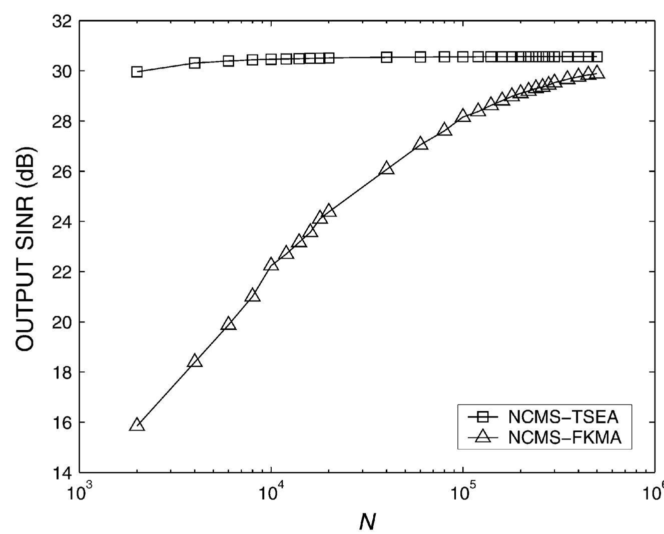

Fig. 3. Simulation results (output SINR versus data length N) of Case 2).

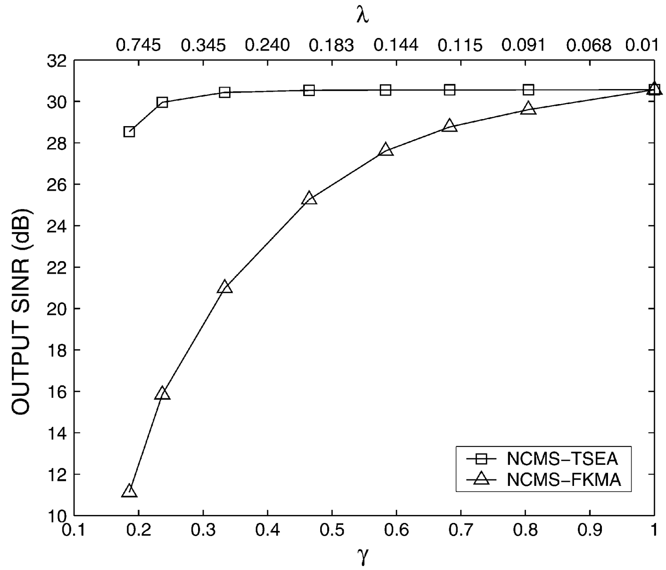

Fig. 4. Simulation results (output SINR versus γ) of Case 3).

for $L \geq 1$, implying that $L \geq 1$ is sufficient (see R4)) for all the simulations in Section V-A. Therefore, the above simulation results verify that the performance of the proposed TSEA is superior to that of the FKMA. As a remark, in spite of the same source power spectra (i.e., the same $\gamma(s_i[n])$ for all i) for each of the above four cases, the proposed TSEA works well with better performance than the FKMA.

B. Performance Comparison

The simulation results to be presented include performance tests to the proposed MSC-TSEA, NCMS-FKMA, NCMS-TSEA, and performance comparison with the MSC-FKMA, FastICA using kurtosis [11], AMUSE [7], and SOBI algorithm [5].

Both of the AMUSE and SOBI algorithm design the $P \times K$ source separation matrix (demixing matrix) $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_K] = \mathbf{D}^{\mathrm{H}}\mathbf{U}$ using SOS, where $\mathbf{D}$ is a $K \times P$ whitening matrix and $\mathbf{U}$ is a $K \times K$ unitary matrix. All the sources are extracted simultaneously without involving temporal processing as follows:

$$\hat{\mathbf{s}}[n] = \mathbf{V}^{\mathrm{H}}\mathbf{x}[n] = \mathbf{U}^{\mathrm{H}}\mathbf{D}\mathbf{x}[n] = \mathbf{U}^{\mathrm{H}}\mathbf{z}[n] \tag{41}$$

where $\mathbf{z}[n] = \mathbf{D}\mathbf{x}[n] = \mathbf{U}\mathbf{s}[n] + \mathbf{D}\mathbf{w}[n]$ is composed of the prewhitened source signals. The whitening matrix $\mathbf{D}$ is obtained through eigen-decomposition of the correlation matrix $\mathbf{R}_x = E[\mathbf{x}[n]\mathbf{x}^{\mathrm{H}}[n]]$. On the other hand, the unitary matrix $\mathbf{U}$ obtained by the AMUSE is through eigen-decomposition of the correlation matrix $\mathbf{R}_{\mathbf{z}}[\tau] = E[\mathbf{z}[n]\mathbf{z}^{\mathrm{H}}[n-\tau]]$ of the prewhitened signal $\mathbf{z}[n]$ for a chosen τ [7], while that obtained by the SOBI algorithm is through joint diagonalization of a set of $\mathbf{R}_{\mathbf{z}}[\tau_j]$ [5]. The simulation results to be presented below were obtained with $\tau = 1$ for the AMUSE and with $(\tau_1, \tau_2, \tau_3) = (1, 2, 3)$ for the SOBI algorithm.

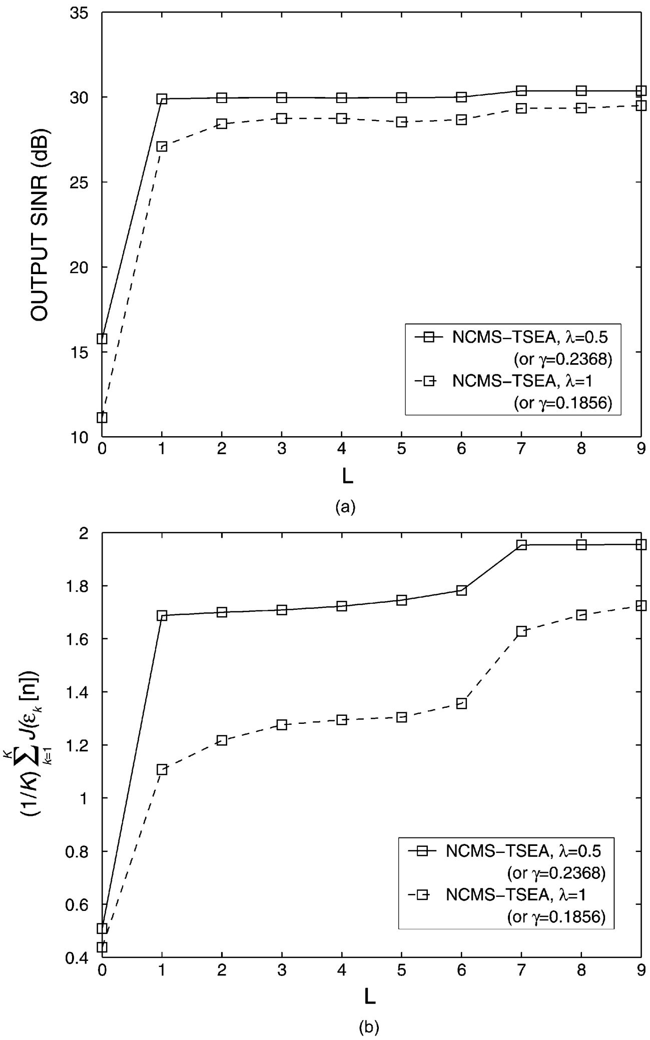

Fig. 5. Simulation results of Case 4). (a) Output SINR versus the order of the temporal filter L and (b) $(1/K)\sum_{k=1}^{K} J(\varepsilon_k[n])$ versus L.

The same 5×4 mixing matrix $\mathbf{A}$ given by (39) and the four MA models $b_i[n]$'s given by (40) with $(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (1, 0.4, 0.3, 0.2)$ (or $(\gamma(s_1[n]), \gamma(s_2[n]), \gamma(s_3[n]), \gamma(s_4[n])) = (0.1856, 0.2706, 0.3335, 0.4644)$) were used in the simulation. The following three cases are considered.

Case A) Output SINR_1 versus SNR for $N = 2000$ (data length) and $L = 5$ (the order of the temporal filter $v[n]$).

Case B) Output SINR versus SNR for $N = 2000$ and $L = 5$.

Case C) Output SINR versus N for SNR $= 20$ dB and $L = 5$.

Fig. 6 shows some simulation results for Case A), where only source 1 is of interest. Note that source 1 was extracted by both of the proposed MSC-TSEA and the MSC-FKMA at stage $\ell \geq 3$ in most of the 50 independent runs. One can see from this figure that the proposed NCMS-TSEA (□) performs best (highest SINR_1), the proposed MSC-TSEA (■) second, the proposed NCMS-FKMA (△) third and their performances are

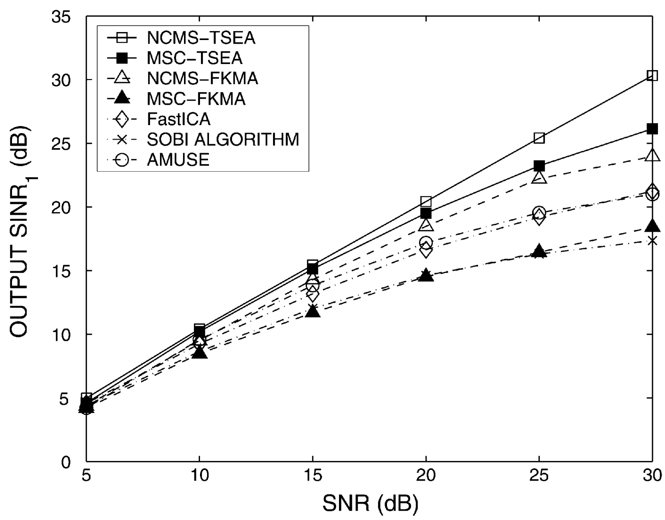

Fig. 6. Simulation results (output SINR_1 versus SNR) of Case A).

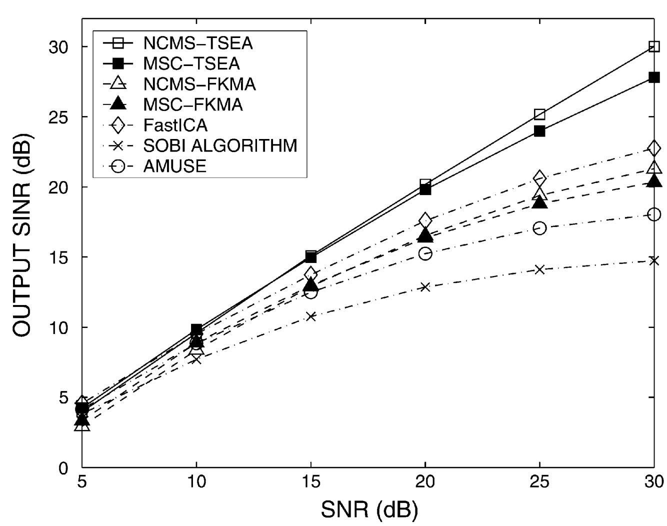

Fig. 7. Simulation results (output SINR versus SNR) of Case B).

superior to those of the MSC-FKMA (▲), FastICA using kurtosis (◊), AMUSE (◯), and SOBI algorithm (×).

Some more observations from Fig. 6 are as follows. The proposed NCMS-TSEA (□) performs much better than the proposed NCMS-FKMA (△) (without cancellation processes for both of them). The proposed MSC-TSEA (■) also performs much better than the MSC-FKMA (▲) (through cancellation processes for both of them) (as stated in R5)). In other words, the proposed TSEA outperforms the FKMA no matter whether the cancellation procedure is involved or not as mentioned in R10). Moreover, the performance of the proposed NCMS-TSEA (□) outperforms the proposed MSC-TSEA (■) and the same observation happens to the proposed NCMS-FKMA (△) and the MSC-FKMA (▲). In other words, the BSS algorithm (using either TSEA or FKMA) without involving the cancellation procedure performs better than involving the cancellation procedure as stated in R8) and R9).

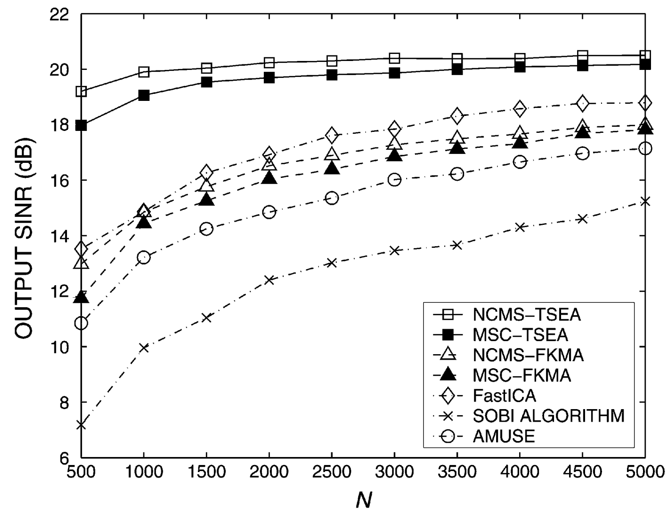

Fig. 8. Simulation results (output SINR versus data length N) of Case C).

Fig. 7 shows some simulation results (output SINR versus SNR) for Case B). One can see from this figure that the proposed NCMS-TSEA (□) performs best (highest SINR), the proposed MSC-TSEA (■) second, FastICA using kurtosis (◊) third, the proposed NCMS-FKMA (△) fourth and their performances are superior to those of the MSC-FKMA (▲), AMUSE (◯), and SOBI algorithm (×). All the other observations from Fig. 6 also apply to Fig. 7, except for smaller performance differences between the proposed NCMS-TSEA (□) and MSC-TSEA (■) for this case because all the sources except source 1 were extracted by the latter at stage $\ell \leq 2$ in most of the 50 independent runs (i.e., output $\mathrm{SINR}_i >$ output SINR_1 for $i \neq 1$). Due to the same reason, performance differences between the proposed NCMS-FKMA (△) and MSC-FKMA (▲) for Case B) are also smaller than those for Case A).

Fig. 8 shows some simulation results (output SINR versus data length N for Case C)). From Fig. 8, one can observe that all the seven BSS algorithms perform better for larger N. Both of the proposed NCMS-TSEA and MSC-TSEA significantly outperform the other BSS algorithms, while the FastICA using kurtosis (◊) performs third. The performance of the proposed NCMS-FKMA is slightly superior to the MSC-FKMA and much better than that of the AMUSE, while the SOBI algorithm performs worst.

Next, let us show some simulation results for source power spectra with different overlaps. The 3×2 mixing matrix $\mathbf{A}$ used was the one by removing the last two rows and last two columns of the 5×4 mixing matrix given by (39), and the transfer functions of the two MA models $b_i[n]$'s used were as follows:

$$\begin{aligned} B_1(z) &= (1+0.5z^{-1})(1-0.8z^{-1})(1+4z^{-1}) \\ B_2(z) &= [1+(0.5+\eta)z^{-1}][1-(0.8+\eta)z^{-1}] \\ &\quad \times [1+(4+\eta)z^{-1}]. \end{aligned}$$

Note that the larger the parameter η, the smaller the spectral overlap between the two source power spectra. The simulation results for the following case are shown in Fig. 9.

TABLE I
COMPLEXITY AND PERFORMANCE COMPARISON OF THE MSC-FKMA, MSC-TSEA, NCMS-FKMA, AND NCMS-TSEA WITH THE SIMULATION RESULTS IN SECTION V-B

Source Separation Algorithm	Spatial Processing No. of Iterations / Source (Spent by FKMA$^{(s)}$) N_s	Temporal Processing No. of Iterations / Source (Spent by FKMA$^{(t)}$) N_t	MSC	SVD of $\mathbf{C}_\ell$ (28)	Performance Rank (Output SINR)
MSC-FKMA	7.35	0	Yes	No	4
NCMS-FKMA	8.89 (Step F3a))+ 6.04 (Step F3b))=14.93	0	No	Yes	3
MSC-TSEA	4.18 / cycle × 2.80 cycles =11.70	3.09 / cycle × 2.80 cycles =8.65	Yes	No	2
NCMS-TSEA	$\overbrace{\text{4.51 / cycle × 3.12 cycles}}^{\text{S1) of TSEA in T3a)}}$ + $\underbrace{\text{4.05 / cycle × 2.86 cycles}}_{\text{S1) of TSEA in T3b)}}$ =25.65	$\overbrace{\text{3.26 / cycle × 3.12 cycles}}^{\text{S2) of TSEA in T3a)}}$ + $\underbrace{\text{1.66 / cycle × 2.86 cycles}}_{\text{S2) of TSEA in T3b)}}$ =14.92	No	Yes	1

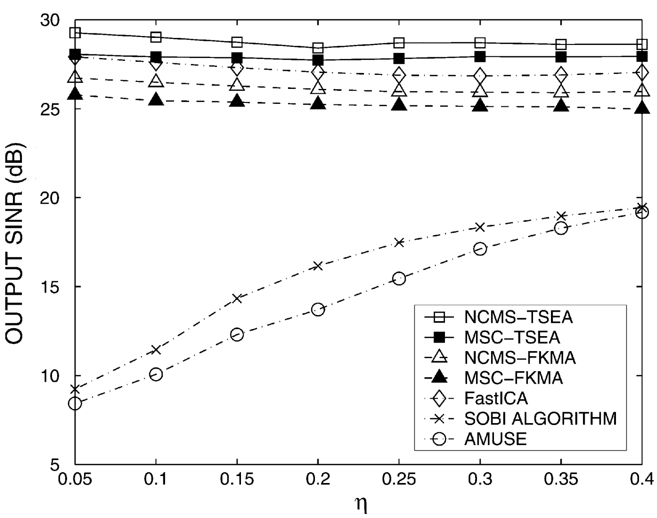

Fig. 9. Simulation results (output SINR versus spectral shift parameter η) of Case D).

Case D) Output SINR versus η for SNR $= 30$ dB, $N = 1000$, and $L = 3$.

From Fig. 9, we can observe that the five BSS algorithms using kurtosis perform well for all η and significantly outperform the AMUSE (◯) and SOBI algorithm ($\times$) whose performance degradation is larger for smaller η, and meanwhile the proposed NCMS-TSEA and MSC-TSEA perform best.

In order to compare the computational complexities of the four kurtosis maximization based BSS algorithms, namely, MSC-FKMA, MSC-TSEA, NCMS-FKMA, and NCMS-TSEA, averages of the total number of iterations per source N_s spent by the FKMA$^{(s)}$ and N_t spent by FKMA$^{(t)}$ for the simulation results in Section V-B are shown in Table I together with whether the MSC and SVD are involved, and performance rank. Some observations from the results shown in Table I are as follows. The value of N_s associated with the MSC-FKMA is 7.35 and that associated with the NCMS-FKMA is 14.93 (about double of the former) which is the sum of $N_s^{(a)} = 8.89$ for F3a) and $N_s^{(b)} = 6.04 < N_s^{(a)}$ for F3b), implying that the preprocessing of constrained source extraction F3a) is of benefit to the convergence speed of the unconstrained source extraction F3b). On the other hand, the values of N_s/cycle and N_t/cycle associated with the MSC-TSEA are 4.18/cycle and 3.09/cycle, respectively, and those associated with the NCMS-TSEA are $N_s^{(a)} = 4.51/\text{cycle}$ and $N_t^{(a)} = 3.26/\text{cycle}$ for T3a), and $N_s^{(b)} = 4.05/\text{cycle}$ ($< N_s^{(a)}$) and $N_t^{(b)} = 1.66/\text{cycle}$ ($< N_t^{(a)}$) for T3b), also implying the preprocessing of the constrained source extraction T3a) is of benefit to the convergence speed of the unconstrained source extraction T3b). Moreover, these values of N_s, N_s/cycle, and N_t/cycle are basically much less than 10 indicating the fast convergence of the FKMA (as stated in Fact 2). These results are also consistent with R8) and R9).

One can also observe from Table I that the MSC-FKMA performs worst with $N_s = 7.35$ (i.e., the lowest computational expense) and that the NCMS-TSEA performs best at the highest computational expense with $N_t = 14.93$ and $N_s = 25.65$ (about 3.5 times the value associated with the MSC-FKMA), the MSC-TSEA performs second with $N_t = 8.65$ and $N_s = 11.70$ (about 1.6 times the value associated with the MSC-FKMA), the NCMS-FKMA performs third with $N_s = 14.93$ (about double of the value associated with the MSC-FKMA).

The computational expenses associated with the MSC and SVD are basically not significant compared with that associated with the FKMA. Overall speaking, Table I provides a comparison of performance and the computational complexities of the four BSS algorithms (using the kurtosis maximization criterion) based on the simulation results presented in Section V-B, although these comparisons also depend on the values of P, K, and L used. These results also support the efficacy of the proposed NCMS-TSEA, NCMS-FKMA, and MSC-TSEA with moderately higher computational load than the MSC-FKMA as mentioned in R8) and R10).

VI. Conclusion

Chi and Chen's FKMA only involves spatial processing for extraction of a single non-Gaussian (i.i.d. or colored) source. It performs well with super-exponential convergence rate, but its performance depends on the parameter $0 < \gamma(s_i[n]) \leq 1$ (see (14)) as stated in Observation 1. By taking the effects of $\gamma(s_i[n])$ on the source extraction performance into account, we have presented a novel blind source extraction algorithm, TSEA shown in Fig. 1, which operates cyclically using the FKMA for both of the temporal processing and spatial processing. The proposed TSEA outperforms the FKMA for $\gamma(s_i[n]) < 1$ in addition to sharing convergence speed and computational efficiency of the latter at each cycle. Note that the proposed TSEA is only applicable for sources that can be modeled as non-Gaussian linear processes.

As any other source extraction algorithms, it is straightforward to apply the proposed TSEA through the MSC procedure for the extraction of all the sources leading to the proposed MSC-TSEA. Because of performance degradation resultant from the error propagation in the MSC procedure as stated in R1), we further presented two noncancellation BSS algorithms, namely, NCMS-FKMA and NCMS-TSEA, that can extract a distinct source at each stage. The proposed NCMS-FKMA and NCMS-TSEA perform better than the existing MSC-FKMA and the proposed MSC-TSEA, respectively, with moderately higher computational complexities, as stated in R8) and R9). Note that the number K of sources was assumed known *a priori* in the proposed BSS algorithms. The determination of K based on normalized kurtosis is left as a future research.

We also provided some simulation results to show the efficacy of the proposed MSC-TSEA, NCMS-FKMA, and NCMS-TSEA together with a performance comparison with some existing BSS algorithms, and a computational load comparison with the MSC-FKMA (Table I). Two final concluding remarks from the simulation results are as follows.

R11) The proposed TSEA outperforms the FKMA no matter whether or not the cancellation procedure is involved.

R12) The proposed BSS algorithms using either TSEA or FKMA without involving the cancellation procedure performs better than involving the cancellation procedure, and meanwhile all the sources are not required to be temporally colored with different power spectra as required by most of SOS based algorithms. ■

Appendix I
Proof of Theorem 1

For ease of later use, let the mixing matrix $\mathbf{A}$ and the source signal vector $\mathbf{s}[n]$ be partitioned as

$$\mathbf{A} = [\boldsymbol{\mathcal{A}}_1 \,|\, \boldsymbol{\mathcal{A}}_2] \tag{I.1}$$

$$\mathbf{s}^{\mathrm{T}}[n] = [s_1[n], s_2[n], \ldots, s_{\ell-1}[n], \mathbf{s}_\ell^{\mathrm{T}}[n]] \tag{I.2}$$

where

$$\boldsymbol{\mathcal{A}}_1 = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{\ell-1}] \tag{I.3}$$

$$\boldsymbol{\mathcal{A}}_2 = [\mathbf{a}_\ell, \mathbf{a}_{\ell+1}, \ldots, \mathbf{a}_K] \tag{I.4}$$

$$\mathbf{s}_\ell[n] = [s_\ell[n], s_{\ell+1}[n], \ldots, s_K[n]]^{\mathrm{T}}. \tag{I.5}$$

An inequality regarding matrix rank which is needed in the proof below is as follows [31]:

$$\begin{aligned}\mathrm{rank}(\mathbf{P}) + \mathrm{rank}(\mathbf{Q}) - m &\leq \mathrm{rank}(\mathbf{PQ}) \\ &\leq \min\{\mathrm{rank}(\mathbf{P}), \mathrm{rank}(\mathbf{Q})\}\end{aligned} \tag{I.6}$$

where $\mathbf{P}$ is a $p \times m$ matrix and $\mathbf{Q}$ is an $m \times q$ matrix. Moreover, by the assumption A1), one can easily prove the following fact (also needed in the proof below).

Fact 3: The matrices $\boldsymbol{\mathcal{A}}_1$ defined by (I.3) and $\boldsymbol{\mathcal{A}}_2$ defined by (I.4) are of full column rank with $\mathrm{rank}(\boldsymbol{\mathcal{A}}_1) = \ell - 1$, and $\mathrm{rank}(\boldsymbol{\mathcal{A}}_2) = K - \ell + 1$, respectively. Moreover, $\mathcal{R}(\boldsymbol{\mathcal{A}}_1) \cap \mathcal{R}(\boldsymbol{\mathcal{A}}_2) = \mathbf{0}_P$.

Assume that, without loss of generality, all the source estimates obtained at the end of the $(\ell-1)$th stage are $\{e_i[n] = \alpha_i s_i[n], i = 1, 2, \ldots, \ell-1\}$ under $\mathcal{A}1$), $\mathcal{A}2$), and the noise-free assumption. Therefore, the $P \times (\ell - 1)$ matrix $\mathbf{C}_\ell$ (defined by (26)) which consists of the associated $\ell - 1$ channel estimates obtained by (15) is given by

$$\mathbf{C}_\ell = \left[\frac{\mathbf{a}_1}{\alpha_1}, \frac{\mathbf{a}_2}{\alpha_2}, \ldots, \frac{\mathbf{a}_{\ell-1}}{\alpha_{\ell-1}}\right] \quad \text{(by (1) and (15))}. \tag{I.7}$$

Let

$$\mathbf{A}_\ell = \mathbf{C}_\ell^\perp \boldsymbol{\mathcal{A}}_2. \tag{I.8}$$

By (33) and (I.8), we have

$$\begin{aligned}\mathbf{x}_\ell[n] &= \mathbf{C}_\ell^\perp \mathbf{x}[n] = \mathbf{C}_\ell^\perp \mathbf{A}\mathbf{s}[n] \text{ (by (1))} \\ &= [\mathbf{C}_\ell^\perp \boldsymbol{\mathcal{A}}_1 \,|\, \mathbf{C}_\ell^\perp \boldsymbol{\mathcal{A}}_2]\mathbf{s}[n] \quad \text{(by (I.1))} \\ &= \mathbf{A}_\ell \mathbf{s}_\ell[n] \quad \text{(by (I.2) and (I.8))}\end{aligned} \tag{I.9}$$

where in the derivation of (I.9), we have used the fact that $\mathbf{C}_\ell^\perp \boldsymbol{\mathcal{A}}_1 = \mathbf{0}_{P\times(\ell-1)}$ because of $\mathcal{R}(\mathbf{C}_\ell) = \mathcal{R}(\boldsymbol{\mathcal{A}}_1)$.

The optimum $e[n]$, denoted $e_\ell[n]$, by maximizing $J(e[n] = \boldsymbol{v}^{\mathrm{T}}\mathbf{x}_\ell[n])$ without any constraint on $\boldsymbol{v}$ is known to be

$$e_\ell[n] = \boldsymbol{v}_\ell^{\mathrm{T}} \mathbf{x}_\ell[n] = \alpha_k s_k[n] \quad \text{(by (I.9) and Fact 1)} \tag{I.10}$$

if the unknown $P \times (K - \ell + 1)$ mixing matrix $\mathbf{A}_\ell$ associated with $\mathbf{x}_\ell[n]$ (see (I.9)) is of full column rank (i.e., $\mathrm{rank}(\mathbf{A}_\ell) = K - \ell + 1$), where α_k is an unknown nonzero constant and $s_k[n] \notin S_\ell$ is one of the source signals in $\mathbf{s}_\ell[n]$. Therefore, what remains to prove is $\mathrm{rank}(\mathbf{A}_\ell) = K - \ell + 1$.

Substituting (28) into (I.8), one can obtain

$$\mathbf{A}_\ell = \mathbf{C}_\ell^{\perp}\boldsymbol{\mathcal{A}}_2 = \mathbf{U}_\ell\mathbf{U}_\ell^{\mathrm{H}}\boldsymbol{\mathcal{A}}_2 = \mathbf{U}_\ell\mathbf{D} \tag{I.11}$$

where

$$\begin{aligned}\mathbf{D} &= \mathbf{U}_\ell^{\mathrm{H}}\boldsymbol{\mathcal{A}}_2 \\ &= [\mathbf{U}_\ell^{\mathrm{H}}\mathbf{a}_\ell, \mathbf{U}_\ell^{\mathrm{H}}\mathbf{a}_{\ell+1}, \ldots, \mathbf{U}_\ell^{\mathrm{H}}\mathbf{a}_K]. \quad \text{(by (I.4))}\end{aligned} \tag{I.12}$$

Again, by the fact of $\mathbf{C}_\ell^{\perp}\boldsymbol{\mathcal{A}}_1 = \mathbf{0}_{P\times(\ell-1)}$ and by (28), (I.7) and Fact 3, one can easily show that

$$\mathbf{U}_\ell^{\mathrm{H}} \cdot \mathbf{a}_i = \mathbf{0}_{P-\ell+1}, \quad i = 1, 2, \ldots, \ell-1, \tag{I.13}$$

i.e., $\mathcal{N}(\mathbf{U}_\ell^{\mathrm{H}}) = \mathcal{R}(\boldsymbol{\mathcal{A}}_1)$ because the nullity of $\mathbf{U}_\ell^{\mathrm{H}}$ is equal to $\ell-1$ (by (29)) and $\mathrm{rank}(\boldsymbol{\mathcal{A}}_1) = \ell-1$ (by Fact 3).

Next, let us prove that the $(P-\ell+1)\times(K-\ell+1)$ matrix $\mathbf{D}$ given by (I.12) is of full column rank, i.e., $\mathrm{rank}(\mathbf{D}) = K-\ell+1$. The matrix $\mathbf{D}$ (see (I.12)) is of full column rank if

$$\begin{aligned}&\mathbf{U}_\ell^{\mathrm{H}}(c_\ell\mathbf{a}_\ell + c_{\ell+1}\mathbf{a}_{\ell+1} + \cdots + c_K\mathbf{a}_K) \\ &\qquad = \mathbf{0}_{P-\ell+1} \quad \text{(by (I.12) and (29))}\end{aligned} \tag{I.14}$$

leads to a unique solution of $c_\ell = c_{\ell+1} = \cdots = c_K = 0$. Because of $\mathcal{N}(\mathbf{U}_\ell^{\mathrm{H}}) = \mathcal{R}(\boldsymbol{\mathcal{A}}_1)$ (by (I.13)), it can be easily inferred that

$$\begin{aligned}&c_\ell\mathbf{a}_\ell + c_{\ell+1}\mathbf{a}_{\ell+1} + \cdots + c_K\mathbf{a}_K \\ &\qquad = c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_{\ell-1}\mathbf{a}_{\ell-1}\end{aligned} \tag{I.15}$$

where $c_i, i = 1, 2, \ldots, K$ are scalars. Note, from (I.15), that the vector on the left-hand side of (I.15) belongs to $\mathcal{R}(\boldsymbol{\mathcal{A}}_2)$ while that on the right-hand side of (I.15) belongs to $\mathcal{R}(\boldsymbol{\mathcal{A}}_1)$. Because of $\mathcal{R}(\boldsymbol{\mathcal{A}}_1)\cap\mathcal{R}(\boldsymbol{\mathcal{A}}_2) = \mathbf{0}_P$ (by Fact 3), one can infer, from (I.15), that

$$\begin{aligned}&c_\ell\mathbf{a}_\ell + c_{\ell+1}\mathbf{a}_{\ell+1} + \cdots + c_K\mathbf{a}_K \\ &\qquad = c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_{\ell-1}\mathbf{a}_{\ell-1} = \mathbf{0}_P\end{aligned} \tag{I.16}$$

which implies that $c_\ell = c_{\ell+1} = \cdots = c_K = 0$ because of $\mathrm{rank}(\boldsymbol{\mathcal{A}}_2) = K-\ell+1$ (by Fact 3). Thus, we have proven that $\mathrm{rank}(\mathbf{D}) = K-\ell+1$.

By (I.11) and (I.6) (with $p = P \geq m = P-\ell+1 \geq q = K-\ell+1$), it can be easily inferred that the $P\times(K-\ell+1)$ matrix $\mathbf{A}_\ell$ must be of full column rank with $\mathrm{rank}(\mathbf{A}_\ell = \mathbf{U}_\ell\mathbf{D}) = K-\ell+1$ because of $\mathrm{rank}(\mathbf{U}_\ell) = P-\ell+1$ (by (29)) and $\mathrm{rank}(\mathbf{D}) = K-\ell+1 \leq \mathrm{rank}(\mathbf{U}_\ell)$. Thus, we have completed the proof. ∎

ACKNOWLEDGMENT

The authors would like to sincerely thank Dr. W.-K. Ma, Dr. C.-C. Feng, Dr. C.-Y. Chen, and Dr. C.-H. Chen for their advice and suggestions during the preparation of the manuscript.

REFERENCES

[1] P. Comon, "Independent component analysis, a new concept?," *Signal Process.*, vol. 36, pp. 287–314, Apr. 1994.

[2] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. New York: Wiley-Interscience, 2001.

[3] A. Cichocki and S. Amari, *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*. New York: Wiley, 2002.

[4] L. Tong, R.-W. Liu, V. C. Soon, and Y.-F. Huang, "Indeterminacy and identifiability of blind identification," *IEEE Trans. Circuits Syst.*, vol. 38, pp. 499–509, May 1991.

[5] 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, Feb. 1997.

[6] N. Delfosse and P. Loubaton, "Adaptive blind separation of independent sources: A deflation approach," *Signal Process.*, vol. 45, pp. 59–83, Feb. 1995.

[7] L. Tong, V. C. Soon, Y.-F. Huang, and R.-W. Liu, "AMUSE: A new blind identification algorithm," in *Proc. IEEE Int. Symp. Circuits Systems*, New Orleans, LA, May 1–3, 1990, pp. 1784–1787.

[8] C. Chang, Z. Ding, S. F. Yau, and F. H. Y. Chan, "A matrix-pencil approach to blind separation of non-white sources in white noise," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Seattle, WA, May 12–15, 1998, pp. 2485–2488.

[9] C. L. Nikias and A. P. Petropulu, *Higher-Order Spectral Analysis: A Nonlinear Signal Processing Framework*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[10] J.-F. Cardoso, "Source separation using higher order moments," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Glasgow, U.K., May 23–26, 1989, pp. 2109–2112.

[11] A. Hyvärinen and E. Oja, "A fixed-point algorithm for independent component analysis," *Neur. Comput.*, vol. 9, pp. 1483–1492, 1997.

[12] Z. Ding, "A new algorithm for automatic beamforming," in *Proc. 25th Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, Nov. 4–6, 1991, pp. 689–693.

[13] Y. Inouye and K. Tanebe, "Super-exponential algorithms for multichannel blind deconvolution," *IEEE Trans. Signal Process.*, vol. 48, no. 3, pp. 881–888, Mar. 2000.

[14] Z. Ding and T. Nguyen, "Stationary points of a kurtosis maximization algorithm for blind signal separation and antenna beamforming," *IEEE Trans. Signal Process.*, vol. 48, no. 6, pp. 1587–1596, Jun. 2000.

[15] C.-Y. Chi and C.-Y. Chen, "Blind beamforming and maximum ratio combining by kurtosis maximization for source separation in multipath," in *Proc. IEEE Workshop Signal Processing Advances Wireless Communications*, Taoyuan, Taiwan, R.O.C., Mar. 20–23, 2001, pp. 243–246.

[16] C.-Y. Chi, C.-Y. Chen, C.-H. Chen, and C.-C. Feng, "Batch processing algorithms for blind equalization using higher-order statistics," *IEEE Signal Process. Mag.*, vol. 20, no. 1, pp. 25–49, Jan. 2003.

[17] O. Shalvi and E. Weinstein, "Super-exponential methods for blind deconvolution," *IEEE Trans. Inf. Theory*, vol. 39, pp. 504–519, Mar. 1993.

[18] K. Yang, T. Ohira, Y. Zhang, and C.-Y. Chi, "Super-exponential blind adaptive beamforming," *IEEE Trans. Signal Process.*, vol. 52, no. 6, pp. 1549–1563, Jun. 2004.

[19] J.-F. Cardoso, "Infomax and maximum likelihood for blind source separation," *IEEE Signal Process. Lett.*, vol. 4, no. 4, pp. 112–114, Apr. 1997.

[20] K. Matsuoka and M. Kawamoto, "A neural net for blind separation of nonstationary signal sources," in *Proc. IEEE Int. Conf. Neural Networks*, Orlando, FL, Jun. 27–Jul. 2, 1994, vol. 1, pp. 221–232.

[21] S. Choi and A. Cichocki, "Blind separation of nonstationary and temporally correlated sources from noisy mixtures," in *Proc. IEEE Signal Processing Society Workshop Neural Networks Signal Processing*, Sydney, NSW, Australia, Dec. 11–13, 2000, vol. 1, pp. 405–414.

[22] ——, "Blind separation of nonstationary sources in noisy mixtures," *IEEE Electron. Lett.*, vol. 36, pp. 848–849, Apr. 2000.

[23] D.-T. Pham and J.-F. Cardoso, "Blind separation of instantaneous mixtures of nonstationary sources," *IEEE Trans. Signal Process.*, vol. 49, no. 9, pp. 1837–1848, Sep. 2001.

[24] J. Zhang, L. Wei, and Y. Wang, "Computational decomposition of molecular signatures based on blind source separation of non-negative dependent sources with NMF," in *Proc. IEEE Workshop Neural Networks Signal Processing*, Sep. 17–19, 2003, pp. 409–418.

[25] J. K. Tugnait, "Identification and deconvolution of multichannel linear non-Gaussian processes using higher order statistics and inverse filter criteria," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 658–672, Mar. 1997.

[26] C.-Y. Chi and C.-H. Chen, "Cumulant based inverse filter criteria for MIMO blind deconvolution: Properties, algorithms, and application to DS/CDMA systems in multipath," *IEEE Trans. Signal Process.*, vol. 49, no. 7, pp. 1282–1299, Jul. 2001.

[27] C.-Y. Chi, C.-H. Chen, and C.-Y. Chen, "Blind MAI and ISI suppression for DS/CDMA systems using HOS-based inverse filter criteria," *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1368–1381, Jun. 2002.

[28] C.-Y. Chi, C.-Y. Chen, C.-H. Chen, C.-C. Feng, and C.-H. Peng, "Blind identification of SIMO systems and simultaneous estimation of multiple time delays from HOS-based inverse filter criteria," *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 2749–2761, Oct. 2004.

[29] J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications," *Proc. IEEE*, vol. 79, pp. 278–305, Mar. 1991.

[30] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: The Johns Hopkins Univ. Press, 1996.

[31] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge Univ. Press, 1985.

[32] C.-Y. Chi and C.-H. Chen, "Two-dimensional frequency-domain blind system identification using higher order statistics with application to texture synthesis," *IEEE Trans. Signal Process.*, vol. 49, no. 4, pp. 864–877, Apr. 2001.

Chong-Yung Chi (S'83–M'83–SM'89) was born in Taiwan, R.O.C., on August 7, 1952. He received the B.S. degree from the Tatung Institute of Technology, Taipei, Taiwan, R.O.C., in 1975, the M.S. degree from the National Taiwan University, Taipei, Taiwan, R.O.C., in 1977, and the Ph.D. degree from the University of Southern California, Los Angeles, in 1983, all in electrical engineering.

From July 1983 to September 1988, he was with the Jet Propulsion Laboratory, Pasadena, CA, where he worked on the design of various spaceborne radar remote sensing systems, including radar scatterometers, SARs, altimeters, and rain mapping radars. From October 1988 to July 1989, he was a visiting Specialist at the Department of Electrical Engineering, National Taiwan University. He has also been with National Tsing Hua University, Hsinchu, Taiwan, R.O.C., as a Professor with the Department of Electrical Engineering since August 1989 and the Institute of Communications Engineering (ICE) since August 1999 (as well as the Chairman of ICE from August 2002 to July 2005). He was a visiting Researcher at the Advanced Telecommunications Research (ATR) Institute International, Kyoto, Japan, in May and June 2001. He has published more than 120 technical papers in radar remote sensing, system identification and estimation theory, deconvolution and channel equalization, digital filter design, spectral estimation, higher order statistics (HOS)-based signal processing, and wireless communications. His current research interests include signal processing for wireless communications, statistical signal processing and digital signal processing and their applications.

Dr. Chi is an active member of Society of Exploration Geophysicists, a member of European Association for Signal Processing, and an active member of the Chinese Institute of Electrical Engineering. He was a Technical Committee Member of both the 1997 and 1999 IEEE Signal Processing Workshop on HOS and the 2001 IEEE Workshop on Statistical Signal Processing (SSP). He was also a member of International Advisory Committee of TENCON 2001 and a Member of International Program Committee of the Fourth International Symposium on Independent Component Analysis and Blind Source Separation (ICA 2003). He was a Co-Organizer and a general Co-Chairman of 2001 IEEE SP Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2001), the International Liaison of SPAWC 2003, and the Technical Committee Member and International Liaison (Asia and Australia) of SPAWC 2004. He was a Program Committee Member of the International Conference on Signal Processing (ICSP 2003) and a Technical Committee Member of the Third IEEE International Symposium on Signal Processing and Information Technology (ISSPIT 2003). He was also the Chair of Information Theory Chapter of IEEE Taipei Section (July 2001 to June 2003). He was Technical Program Committee (TPC) Member of the Emerging Networks, Technologies & Standards Symposium, WirelessCom 2005, Track TPC (TTPC) Member of the 2005 International Conference on Communications, Circuits and Systems, and TPC Member of First IEEE International Workshop on Computational Advances in Multisensor Adaptive Processing (IEEE CAMSAP 2005). Currently, he is an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and an Editorial Board Member of JASP (*The EURASIP Journal on Applied Signal Processing*), a member of Editorial Board of *EURASIP Signal Processing Journal*, a member of IEEE Technical Committee on Signal Processing Theory and Methods, TPC Member of the 2006 International Conference on Communications (ICC) Wireless and Ad hoc Sensor Networks, and TPC Member of the 2006 International Wireless Communications and Computing Conference (IWCCC 2006).

Chun-Hsien Peng was born in Taiwan, R.O.C., on November 21, 1978. He received the B.S. degree from the Department of Electrical Engineering, National Taipei University of Technology, Taipei, Taiwan, R.O.C., in 2001. He is currently working towards the Ph.D. degree at the Institute of Communications Engineering, Department of Electrical Engineering, National Tsing Hua University, Hsinchu, Taiwan, R.O.C.

His research interests include statistical signal processing, digital signal processing, and wireless communications.