

An MSE-Based Method to Avoid Permutation/Gain Indeterminacy in Frequency-Domain Blind Source Separation

Adriana Dapena · Daniel Iglesia ·
Carlos J. Escudero

Received: 23 April 2008 / Revised: 31 March 2009 / Published online: 19 February 2010
© Springer Science+Business Media, LLC 2010

Abstract This paper presents a novel technique for separating convolutive mixtures of statistically independent non-Gaussian signals without resorting to an *a priori* knowledge of the sources or the mixing system. This problem is solved in the frequency domain by transforming the convolutive mixture into several instantaneous mixtures which are independently separated using blind source separation (BSS) algorithms. First, the instantaneous mixture at one frequency is solved using the joint approximate diagonalization of eigenmatrices (JADE) technique, and the other mixtures are then separated using the mean squared error (MSE) criterion. As a special case of this method, we consider the separation of non-Gaussian temporally white signals transmitted in blocks with zero padding between them.

Keywords Convolutive mixtures · Blind source separation · JADE · Frequency-domain systems · Unsupervised algorithms

1 Introduction

Blind source separation (BSS) consists of recovering a set of unknown signals (sources) from mixtures recorded by an array of sensors (observations). The term

This work has been supported by Ministerio de Industria, Turismo y Comercio (Grant CSD2008-00010), Ministerio de Educación y Ciencia (Grant TEC2007-68020-C04-01) and Xunta de Galicia (PGIDT06TIC10501PR).

A. Dapena (✉) · D. Iglesia · C.J. Escudero
Departamento de Electrónica y Sistemas, Facultad de Informática, Universidade da Coruña, Campus de Elviña 5, 15071, Corunna, Spain
e-mail: adriana@udc.es

D. Iglesia
e-mail: dani@udc.es

C.J. Escudero
e-mail: escudero@udc.es

blind (or unsupervised) refers to the fact that the sources and the mixing system are completely unknown [2]. Since the seminal work of [6, 9], rising research interest in BSS has been motivated by the huge variety of application areas where this problem appears [10]: audio processing, image compression, digital communication, etc.

The mixtures are often considered as instantaneous and linear, so that each observation is a sum of differently weighted sources. Assuming this model, the sources can be recovered by using a multiple-input multiple-output (MIMO) system whose coefficients are adapted by using a given statistical criterion. However, in many real-world applications, such as acoustics [11, 13], motor monitoring [8], and digital communication [7], each source contributes to the sum with multiple delays corresponding to multiple paths. The observations are then modeled as convolutive mixtures of the sources. Finding a solution to this problem is more complex than in the case of instantaneous mixtures because the separating system must recover all the sources and eliminate the memory of the mixing system.

It is interesting to note that a convolutive mixture can be transformed into several instantaneous mixtures by using a time-to-frequency transform. This interpretation allows the separation problem to be decomposed into smaller problems at each frequency. However, if the instantaneous mixture at each frequency is independently separated, the frequency-domain sources can be recovered with a different order and gain. When this undesirable situation occurs, the time-domain sources cannot be recovered using a frequency-to-time transform.

The classical scheme for solving permutation/gain indeterminacy consists of including additional stages in the separating system. For a review of this approach, the reader is referred to [14]. Capdevielle et al. [3] have proposed a method to avoid the permutation indeterminacy by recovering the continuity of the frequency spectra. Due to the independence of the sources, the cross-correlation between the frequency-domain outputs corresponding to different sources is zero and, therefore, the frequency-domain outputs corresponding to the same source can be determined by maximizing the cross-correlation between them. This idea has also been used in [8, 15]. The disadvantage of this method is its high computational cost, since the Fourier transform is calculated with a window shift of one point. In addition, it can only be used for temporally colored sources. Computationally less expensive methods for solving the permutation indeterminacy have been proposed in [7, 12] for temporally white sources. The idea is also to cluster the outputs taking into account the statistical dependence between frequency-domain outputs corresponding to the same source, but the discrete Fourier transform (DFT) is applied to nonoverlapped windows: Mejuto et al. [12] have proposed maximizing fourth-order cross-cumulants while Dapena et al. [7] utilize cross-correlation.

To the best of our knowledge, gain indeterminacy for the case of temporally white sources has been addressed only in Dapena et al. [7]. The idea consists of estimating the scale introduced at each frequency by considering the cross-correlation between frequency-domain outputs and the convergence point of the BSS algorithm used to separate the instantaneous mixture at each frequency.

The main contribution of this work is to present a separating system that does not suffer from permutation/gain indeterminacy. We will show that if the sources in different frequencies are uncorrelated and independent, permutation/gain indeterminacy can be avoided by selecting the separating coefficients in order to minimize the

mean squared error (MSE) between the frequency-domain outputs in adjacent frequencies. Using this strategy, all the frequency-domain sources are extracted with the same order and gain, and the sources can be recovered using the frequency-to-time transform.

This paper is structured as follows. In Sect. 2, we describe the signal model. In Sect. 3 we present the main idea to avoid the permutation/gain indeterminacy and we propose the novel system. Section 4 particularizes the system to the case of temporally white signals transmitted in burst. Section 5 presents some simulation results, and Sect. 6 contains the conclusions.

Notation. Throughout this paper, upper case and lower case bold characters denote matrices and vectors, respectively. $(\cdot)^T$, $(\cdot)^H$ and $(\cdot)^{-1}$ denote the transpose, the conjugate transpose, and the pseudo-inverse of the matrix argument, respectively. \mathbf{a}_k represents the k th column of matrix \mathbf{A} and $\mathbf{a}_{(1:M)}$ denotes a vector containing the elements from 1 to M of vector \mathbf{a} . $E[\cdot]$ is the expectation operator.

2 Signal Model: Instantaneous and Convolutional Mixtures

Let $x_i(n)$, $i = 1, 2, \dots, M$ be the M measured signals, termed observations. The observations are instantaneous linear mixtures of N ($N \leq M$) unknown random processes $s_i(n)$, $i = 1, 2, \dots, N$, termed sources, whose exact probability density function is unknown. We only assume that they are complex-valued, zero-mean, stationary, non-Gaussian distributed, and statistically independent. These hypotheses are satisfied in a large number of applications.

In the basic BSS problem, the observations consist of a sum of differently weighted sources. In matrix notation, the vector of observations $\mathbf{x}(n) = [x_1(n), \dots, x_M(n)]^T$ can be written as

$$\mathbf{x}(n) = \mathbf{A}\mathbf{s}(n), \quad (1)$$

where $\mathbf{s}(n) = [s_1(n), \dots, s_N(n)]^T$ is the vector of sources and \mathbf{A} is an $M \times N$ matrix. The sources can be recovered from the observations by using a MIMO system whose output is a linear combination of the observations,

$$\mathbf{y}(n) = \mathbf{W}^H \mathbf{x}(n), \quad (2)$$

where \mathbf{W} is the $M \times N$ separating matrix. The aim in BSS is to select the matrix \mathbf{W} to recover all the sources from the observations, without using training sequences or information about the mixing system. Note that, combining both (1) and (2) together, the outputs can be written as a linear combination of the sources

$$\mathbf{y}(n) = \mathbf{G}\mathbf{s}(n), \quad (3)$$

where $\mathbf{G} = \mathbf{W}^H \mathbf{A}$ is an $N \times N$ matrix representing the overall mixing/separating system. From this (3), it can be concluded that the sources are recovered when the matrix \mathbf{G} has the form

$$\mathbf{G} = \mathbf{\Delta}\mathbf{P}, \quad (4)$$

where \mathbf{A} is an $N \times N$ diagonal invertible matrix and \mathbf{P} is an $N \times N$ permutation matrix. This means that BSS algorithms can permute and scale the sources.

In many real-world applications, the sources are received at the antennas by multiple paths [14]. In this case, the observations are expressed as convolutive mixtures of the sources

$$\mathbf{x}(n) = \sum_{l=-\infty}^{\infty} \mathbf{A}(l)\mathbf{s}(n-l), \quad (5)$$

where $\mathbf{A}(l)$ is an unknown $M \times N$ matrix representing the mixing system whose elements, $a_{ij}(l)$, model the propagation conditions from the j th source to the i th observation. One way of solving the convolutive problem consists in transforming the convolutive mixture into several instantaneous mixtures by using the Fourier transform given by

$$x_i[\omega_k, n] = \sum_{t=0}^{L-1} x_i(n(L-1)+t)e^{-j\omega_k t}, \quad n = 0, 1, \dots, \quad (6)$$

where $\omega_k = 2\pi k/L$, $k = 0, 1, \dots, L-1$ denotes the frequency and L denotes the number of points in the DFT. Using the properties of the DFT, the frequency-domain observation vector at each frequency is an instantaneous mixture of the frequency-domain sources given by

$$\mathbf{x}[\omega_k, n] = \mathbf{A}[\omega_k]\mathbf{s}[\omega_k, n], \quad (7)$$

where $\mathbf{A}[\omega_k]$ represents the mixing coefficients in the frequency domain. As a result, we can recover the frequency-domain sources by using a MIMO system with output

$$\mathbf{y}[\omega_k, n] = \mathbf{W}^H[\omega_k]\mathbf{x}[\omega_k, n], \quad (8)$$

where $\mathbf{W}[\omega_k]$ is the $M \times N$ coefficients matrix which is typically obtained using BSS algorithms proposed in the case of instantaneous mixtures.

Notice that BSS algorithms for instantaneous mixtures suffer from the permutation/gain indeterminacy given in (4). This means that if the separating system at each frequency is computed independently, the frequency-domain sources can be recovered in a different order (permutation indeterminacy) and with different gain (gain indeterminacy). When this occurs, the sources cannot be recovered using the inverse DFT (IDFT).

3 The FD-BSS Cascade System

We assume the following conditions.

- C1. The convolutive mixture can be transformed into several instantaneous mixtures using the DFT given in (6).

C2. The cross-correlation between the sources in different frequencies (ω_k and ω_r) satisfies

$$E[s_i[\omega_k, n]s_i^*[\omega_r, n]] \neq 0, \tag{9}$$

$$E[s_i[\omega_k, n]s_j^*[\omega_r, n]] = 0 \quad \text{if } i \neq j, \tag{10}$$

where the cross-correlation is estimated by sample averaging over the R samples

$$E[s_i[\omega_k, n]s_j^*[\omega_r, n]] = \frac{1}{R} \sum_{n=0}^{R-1} s_i[\omega_k, n]s_j^*[\omega_r, n]. \tag{11}$$

C3. The mixing matrix $\mathbf{A}[\omega]$ in all the frequencies is invertible.

Figure 1 shows the proposed separating system. Each particular observation, $x_j(t)$, $j = 1, \dots, M$, is split into nonoverlapped segments of L points and the L-DFT is computed over them. The next step consists in computing the separating coefficients at *one* specific frequency, denoted by ω_r , by using a given unsupervised algorithm whose output, $\mathbf{y}[\omega_r, n]$, will be used as reference in the adjacent frequencies $r - 1$ and $r + 1$. The separating matrices in these frequencies, $\mathbf{W}[\omega_{r-1}]$ and $\mathbf{W}[\omega_{r+1}]$, are computed by using the MSE criterion defined by

$$\begin{aligned} \mathbf{W}[\omega_k] &= \arg \min_{\mathbf{W}[\omega_k]} J(\mathbf{W}[\omega_k]) \\ &= \arg \min_{\mathbf{W}[\omega_k]} E[|\mathbf{y}[\omega_k, n] - \mathbf{\Delta}_k \mathbf{y}[\omega_r, n]|^2], \quad k \in \{r - 1, r + 1\}, \end{aligned} \tag{12}$$

where $\mathbf{\Delta}_k$ is an $N \times N$ diagonal matrix that must be determined by taking into account the statistics of the sources. Notice that condition C2 guarantees that the frequency-domain sources corresponding to the same time-domain sources are correlated and,

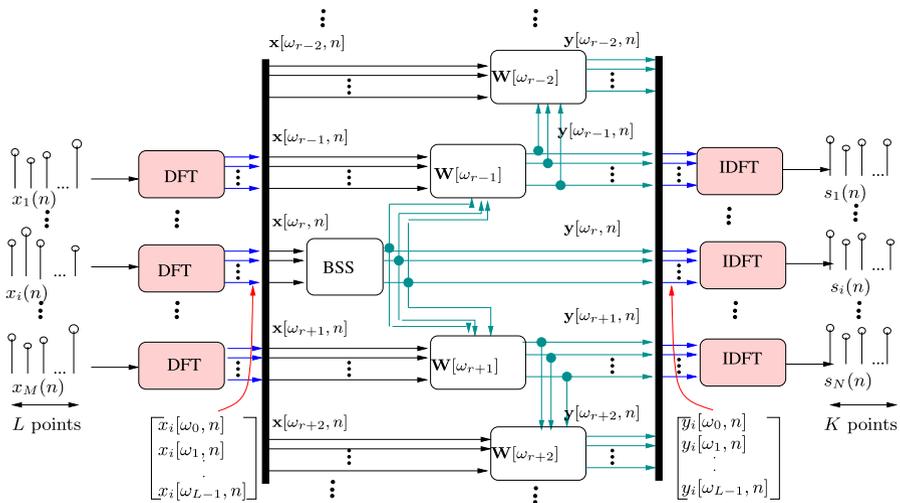


Fig. 1 Scheme of the FD-BSS cascade system

therefore, the criterion (12) can be used to measure the dependence between $\mathbf{y}[\omega_k, n]$ and $\mathbf{y}[\omega_r, n]$.

It is well known that the solution to the MSE criterion (12) is given by

$$\mathbf{W}_o[\omega_k] = \Delta_k^* \mathbf{R}_x^{-1}[\omega_k, \omega_k] \mathbf{R}_{xy}[\omega_k, \omega_r], \tag{13}$$

where

$$\begin{aligned} \mathbf{R}_x[\omega_k, \omega_k] &= E[\mathbf{x}[\omega_k, n] \mathbf{x}^H[\omega_k, n]], \\ \mathbf{R}_{xy}[\omega_k, \omega_r] &= E[\mathbf{x}[\omega_k, n] \mathbf{y}^H[\omega_r, n]]. \end{aligned} \tag{14}$$

If $\mathbf{R}_x[\omega_k, \omega_k]$ is not full rank, a pseudo-inverse is computed instead of the inverse in (13).

The outputs of this frequency $\mathbf{y}[\omega_k, n] = \mathbf{W}_o^H[\omega_k] \mathbf{x}[\omega_k, n]$ are now used as a reference at the adjacent frequency. The last stage of our separating system consists of computing the IDFT of each output in all the frequencies, $y_i[\omega_k, n], k = 0, \dots, L - 1$, to obtain the estimation of the sources in the time domain.

3.1 The Matrix Δ_k

We will determine the form of matrix Δ_k in order to ensure the same permutation and gain in all the frequencies. Using $\mathbf{x}[\omega_k, n] = \mathbf{A}[\omega_k] \mathbf{s}[\omega_k, n]$ and $\mathbf{y}[\omega_r, n] = \mathbf{G}[\omega_r] \mathbf{s}[\omega_r, n]$, the matrix $\mathbf{R}_{xy}[\omega_k, \omega_r]$ can be written as

$$\mathbf{R}_{xy}[\omega_k, \omega_r] = \mathbf{A}[\omega_k] \mathbf{R}_s[\omega_k, \omega_r] \mathbf{G}^H[\omega_r], \tag{15}$$

where $\mathbf{R}_s[\omega_k, \omega_r] = E[\mathbf{s}[\omega_k, n] \mathbf{s}^H[\omega_r, n]]$. By a similar reasoning, the other term in equation (14) can be written as follows:

$$\mathbf{R}_x[\omega_k, \omega_k] = \mathbf{A}[\omega_k] \mathbf{R}_s[\omega_k, \omega_k] \mathbf{A}^H[\omega_k], \tag{16}$$

where $\mathbf{R}_s[\omega_k, \omega_k] = E[\mathbf{s}[\omega_k] \mathbf{s}^H[\omega_k]]$. Substituting (15) and (16) in (13), the separating matrix takes the form

$$\begin{aligned} \mathbf{W}_o[\omega_k] &= \Delta_k^* (\mathbf{A}[\omega_k] \mathbf{R}_s[\omega_k, \omega_k] \mathbf{A}^H[\omega_k])^{-1} (\mathbf{A}[\omega_k] \mathbf{R}_s[\omega_k, \omega_r] \mathbf{G}^H[\omega_r]) \\ &= \Delta_k^* (\mathbf{A}^{-H}[\omega_k] \mathbf{R}_s^{-1}[\omega_k, \omega_k] \mathbf{A}^{-1}[\omega_k]) (\mathbf{A}[\omega_k] \mathbf{R}_s[\omega_k, \omega_r] \mathbf{G}^H[\omega_r]) \\ &= \Delta_k^* \mathbf{A}^{-H}[\omega_k] \mathbf{R}_s^{-1}[\omega_k, \omega_k] \mathbf{R}_s[\omega_k, \omega_r] \mathbf{G}^H[\omega_r], \end{aligned} \tag{17}$$

where $\mathbf{A}^{-H}[\omega_k] = (\mathbf{A}^H[\omega_k])^{-1}$. Using this equation we can express the gain matrix at the k -th frequency as follows:

$$\begin{aligned} \mathbf{G}[\omega_k] &= \mathbf{W}_o^H[\omega_k] \mathbf{A}[\omega_k] \\ &= \Delta_k (\mathbf{A}^{-H}[\omega_k] \mathbf{R}_s^{-1}[\omega_k, \omega_k] \mathbf{R}_s[\omega_k, \omega_r] \mathbf{G}^H[\omega_r])^H \mathbf{A}[\omega_k] \\ &= \Delta_k (\mathbf{G}[\omega_r] \mathbf{R}_s^H[\omega_k, \omega_r] \mathbf{R}_s^{-H}[\omega_k, \omega_k] \mathbf{A}^{-1}[\omega_k]) \mathbf{A}[\omega_k] \\ &= \Delta_k \mathbf{G}[\omega_r] \mathbf{R}_s^H[\omega_k, \omega_r] \mathbf{R}_s^{-H}[\omega_k, \omega_k] = \Delta_k \mathbf{G}[\omega_r] \mathbf{P}[\omega_k, \omega_r], \end{aligned} \tag{18}$$

where $\mathbf{R}_s^{-H}[\omega_k, \omega_k] = (\mathbf{R}_s[\omega_k, \omega_k]^H)^{-1}$ and $\mathbf{P}[\omega_k, \omega_r] = \mathbf{R}_s^H[\omega_k, \omega_r]\mathbf{R}_s^{-H}[\omega_k, \omega_k]$. Note that the permutation/gain indeterminacy is solved when the matrix $\mathbf{P}[\omega_k, \omega_r]$ is removed in (18). In particular, when the condition C2 is verified, $\mathbf{R}_s[\omega_k, \omega_r]$ and $\mathbf{R}_s[\omega_k, \omega_k]$ are diagonal matrices with entries $E[s_i[\omega_k, n]s_i^*[\omega_r, n]]$ and $E[|s_i[\omega_k, n]|^2]$, respectively. This implies that $\mathbf{P}[\omega_k, \omega_r]$ is a diagonal matrix

$$\mathbf{P}[\omega_k, \omega_r] = \begin{bmatrix} \frac{E[s_1^*[\omega_k, n]s_1[\omega_r, n]]}{E[|s_1[\omega_k, n]|^2]} & 0 & \dots & 0 \\ 0 & \frac{E[s_2^*[\omega_k, n]s_2[\omega_r, n]]}{E[|s_2[\omega_k, n]|^2]} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{E[s_N^*[\omega_k, n]s_N[\omega_r, n]]}{E[|s_N[\omega_k, n]|^2]} \end{bmatrix}. \tag{19}$$

As a result, the matrix Δ_k must be computed by inverting the matrix $\mathbf{P}[\omega_k, \omega_r]$ given in (19).

3.2 The Reference Frequency

A new question arises when the convolutive problem is transformed into instantaneous mixtures: not all instantaneous separation algorithms are designed for separating mixtures of complex-valued sources. In addition, the convergence of some algorithms, like Complex FastICA [1], depends on the selection of some nonlinear functions and step-size parameters.

In order to separate the sources at the reference frequency, we have selected the blind identification algorithm by joint approximate diagonalization of eigenmatrices (JADE) proposed in [4], which works with complex-valued signals and whose convergence does not depend on nonlinearities. The JADE algorithm can be described by the following steps (a MATLAB implementation of JADE is available on www.tsi.enst.fr/~cardoso/guidesepsou.html):

- Step 1. Compute the whitening matrix $\mathbf{U}[\omega_r]$ from the sample covariance $\mathbf{R}_x[\omega_r, \omega_r]$ and obtain the whitened process $\mathbf{z}[\omega_r] = \mathbf{U}[\omega_r]\mathbf{x}[\omega_r]$.
- Step 2. Compute the fourth-order cumulants matrices $\mathbf{C}^{k,l}$, $k, l = 1, \dots, M$ of $\mathbf{z}[\omega_r]$. Each matrix has dimension $M \times M$ and is defined by

$$\begin{aligned} \mathbf{C}^{k,l} &= c_4(\mathbf{z}[\omega_r], \mathbf{z}^H[\omega_r], z_k[\omega_r], z_l^*[\omega_r]) \\ &= \begin{bmatrix} c_4(z_1, z_1^*, z_k, z_l^*) & c_4(z_1, z_2^*, z_k, z_l^*) & \dots & c_4(z_1, z_{N_r}^*, z_k, z_l^*) \\ c_4(z_2, z_1^*, z_k, z_l^*) & c_4(z_2, z_2^*, z_k, z_l^*) & \dots & c_4(z_2, z_{N_r}^*, z_k, z_l^*) \\ \vdots & \vdots & \ddots & \vdots \\ c_4(z_M, z_1^*, z_k, z_l^*) & c_4(z_M, z_2^*, z_k, z_l^*) & \dots & c_4(z_M, z_M^*, z_k, z_l^*) \end{bmatrix}, \end{aligned} \tag{20}$$

where $z_i = z_i[\omega_r]$ and $c_4(\cdot)$ represents the fourth-order cross-cumulant,

$$\begin{aligned} c_4(z_1, z_2, z_3, z_4) &= E[z_1z_2z_3z_4] - E[z_1z_2]E[z_3z_4] - E[z_1z_3]E[z_2z_4] \\ &\quad - E[z_1z_4]E[z_2z_3]. \end{aligned} \tag{21}$$

- Step 3. Compute the $M^2 \times M^2$ matrix $\mathbf{B} = [\hat{\mathbf{c}}^{1,1}, \hat{\mathbf{c}}^{1,2}, \dots, \hat{\mathbf{c}}^{1,M}, \dots, \hat{\mathbf{c}}^{M,M}]$ where $\hat{\mathbf{c}}^{i,j} = [(\mathbf{c}_1^{i,j})^T, (\mathbf{c}_2^{i,j})^T, \dots, (\mathbf{c}_M^{i,j})^T]^T$ is an $M^2 \times 1$ vector formed by the columns of $\mathbf{C}^{i,j}$ ($\mathbf{c}_k^{i,j}$ denotes the k th column of $\mathbf{C}^{i,j}$).
- Step 4. Perform the eigendecomposition of \mathbf{B} and select the N more significant eigenpairs $\{\lambda_i, \mathbf{m}_i\}, i = 1, \dots, N$. From these eigenpairs, compute the $M \times M$ matrices $\mathbf{M}_i = \lambda_i[\mathbf{m}_{i(1:M)} \mathbf{m}_{i(M+1:2M)} \dots], i = 1, \dots, N$.
- Step 5. Diagonalize jointly the set of matrices $\mathbf{M}_i, i = 1, \dots, N$ to obtain the matrix $\mathbf{W}[\omega_r]$. This step is implemented by extending the single-matrix Jacobi technique to several matrices as described in [5].
- Step 6. Estimate the mixing matrix $\mathbf{A}[\omega_r]$ as $\hat{\mathbf{A}}[\omega_r] = \mathbf{U}^H[\omega_r]\mathbf{W}[\omega_r]$ and recover the frequency-domain sources as $\mathbf{y}[\omega_r, n] = \hat{\mathbf{A}}^{-1}[\omega_r]\mathbf{x}[\omega_r, n]$.

Although in most cases JADE provides an adequate separation, we have observed that the degree of separation depends on the eigenvalue spread of the matrix diagonalized in Step 5. We propose the following procedure to select the reference frequency:

- Set as reference frequency $r = 0$.
- Step 1: Compute JADE for the frequency ω_r and evaluate the eigenvalue spread of the matrix diagonalized in Step 5. We define the eigenvalue spread as follows:

$$e_i = \sum_{i=1} |\lambda_i - \lambda_{i+1}|, \tag{22}$$

where $\lambda_i, i = 1, 2 \dots$ denotes the eigenvalues of this matrix, with $\lambda_i > \lambda_{i+1}$.

- Step 2: If the eigenvalue spread is less than a threshold β , use this frequency as reference. If this is not the case, set $r = r + 1$ and go to Step 1.

If the eigenvalue spread is greater than β for all the frequencies, we set r to the frequency with the lowest eigenvalue spread. Note that this procedure increases the computational load because JADE is used in several frequencies. As shown in Sect. 5, however, this scheme significantly improves the performance of the FD-BSS cascade system.

4 A Particular Case: Temporally White Sources

In this section, we particularize the results presented in Sect. 3.1 to the case of temporally white stationary sources. In order to satisfy condition C1, we assume that the sources are transmitted in bursts with guard time between frames.

First, we are going to introduce a more convenient notation to represent the terms in (7). Recall that the i th frequency-domain observation has been obtained by applying the L-DFT to nonoverlapped windows, as is shown in (6). In a compact form, this can be written as $x_i[\omega_k, n] = \hat{\mathbf{f}}_k^T \mathbf{x}_i(n)$ where $\hat{\mathbf{f}}_k = [1, e^{-j2\pi k/L}, \dots, e^{-j2\pi k(L-2)/L}, e^{-j2\pi k(L-1)/L}]^T$ is the Fourier vector and $\mathbf{x}_i = [x_i(n), x_i(n+1), \dots, x_i(n+L-1)]^T$ is a window containing L samples of the i th observation.

In addition, assuming that the mixing system can be modeled as a finite impulse response (FIR) filter of order P , the entries in the mixing matrices $\mathbf{A}[\omega_k], k =$

0, . . . , L - 1, have the form

$$a_{ij}[\omega_k] = \hat{\mathbf{f}}_{k(0:P-1)}^T \mathbf{a}_{ij}, \tag{23}$$

where $\mathbf{a}_{ij} = [a_{ij}(0), a_{ij}(1), \dots, a_{ij}(P - 1)]^T$ contains the P nonzero values representing the propagation coefficients and $\hat{\mathbf{f}}_{k(0:P-1)} = [1, e^{-j2\pi k/L}, \dots, e^{-j2\pi k(P-2)/L}, e^{-j2\pi k(P-1)/L}]^T$.

Considering that the sources are transmitted in frames of K samples, $\mathbf{s}_j(n) = [s_j(n), s_j(n + 1), \dots, s_j(n + K - 1)]^T$, the j th frequency-domain source can be written as

$$s_j[\omega_k, n] = \hat{\mathbf{f}}_{k(0:K-1)}^T \mathbf{s}_j(n). \tag{24}$$

For convenience, we will denote the truncated Fourier vector as $\mathbf{f}_k = \hat{\mathbf{f}}_{k(0:K-1)} = [1, e^{-j2\pi k/L}, \dots, e^{-j2\pi k(K-2)/L}, e^{-j2\pi k(K-1)/L}]^T$. When the sources are temporally white stationary signals, the cross-correlation between the frequency-domain sources in two frequencies, $s_i[\omega_r, n]$ and $s_j[\omega_k, n]$, is given by

$$E[s_i^*[\omega_k, n]s_j[\omega_r, n]] = E[s_i^*(n)s_j(n)]\mathbf{f}_k^H \mathbf{f}_r = E[s_i^*(n)s_j(n)]\mathbf{f}_k^H \mathbf{f}_r. \tag{25}$$

Since the sources are statistically independent and assuming $E[|s_i(n)|^2] = 1$, we conclude

$$\begin{aligned} E[s_i^*[\omega_k, n]s_j[\omega_r, n]] &= 0, \quad i \neq j, \quad \forall k, r, \\ E[s_i^*[\omega_k, t]s_i[\omega_r, n]] &= E[|s_i(n)|^2]\mathbf{f}_k^H \mathbf{f}_r = \mathbf{f}_k^H \mathbf{f}_r, \\ E[|s_i[\omega_k, n]|^2] &= \mathbf{f}_k^H \mathbf{f}_k = K. \end{aligned} \tag{26}$$

Note that, in order to ensure condition C2, the block size K and the number of frequencies L must be selected such that $\mathbf{f}_k^H \mathbf{f}_r \neq 0$.

Substituting (26) in (19), we conclude that the matrix $\mathbf{P}[\omega_k]$ has the form

$$\mathbf{P}[\omega_k] = \frac{E[s_i^*[\omega_k, n]s_i[\omega_r, n]]}{E[|s_i[\omega_k, n]|^2]} \mathbf{I} = \frac{\mathbf{f}_k^H \mathbf{f}_r}{K} \mathbf{I}, \tag{27}$$

where \mathbf{I} is the $N \times N$ identity matrix. As a result, the matrix in (13) has the form $\Delta_k = \alpha_k \mathbf{I}$ with

$$\begin{aligned} \alpha_k &= K / (\mathbf{f}_k^H \mathbf{f}_r) = \frac{K}{\sum_{m=0}^{K-1} e^{-j2\pi(r-k)m/L}} \\ &= \frac{K(1 - e^{-j2\pi(r-k)/L})}{1 - e^{-j2\pi(r-k)K/L}}. \end{aligned} \tag{28}$$

In particular, when the reference signals are the frequency-domain outputs obtained in the previous frequency, $r = k - 1$, the parameter α_k is selected as

$$\alpha_k = \frac{K(1 - e^{j2\pi/L})}{1 - e^{j2\pi K/L}}. \tag{29}$$

In the other case, where $r = k + 1$, the parameter α_k is given by

$$\alpha_k = \frac{K(1 - e^{-j2\pi/L})}{1 - e^{-j2\pi K/L}}. \tag{30}$$

Note that the parameters K and L must be selected in order to guarantee that K/L is not integer.

5 Simulation Results

In this section, we will illustrate the behavior of the proposed separating system. We have used complex-valued temporally white sources (4-QAM) divided in blocks of $K = 11$ symbols. The mixing system is modeled as FIR filters of $P = 6$ taps with randomly generated coefficients. The L-DFT has been applied to nonoverlapped windows of $L = K + P - 1 = 16$ points of the observations.

Fig. 2 Final gains obtained using JADE to separate the instantaneous mixtures in all the frequencies

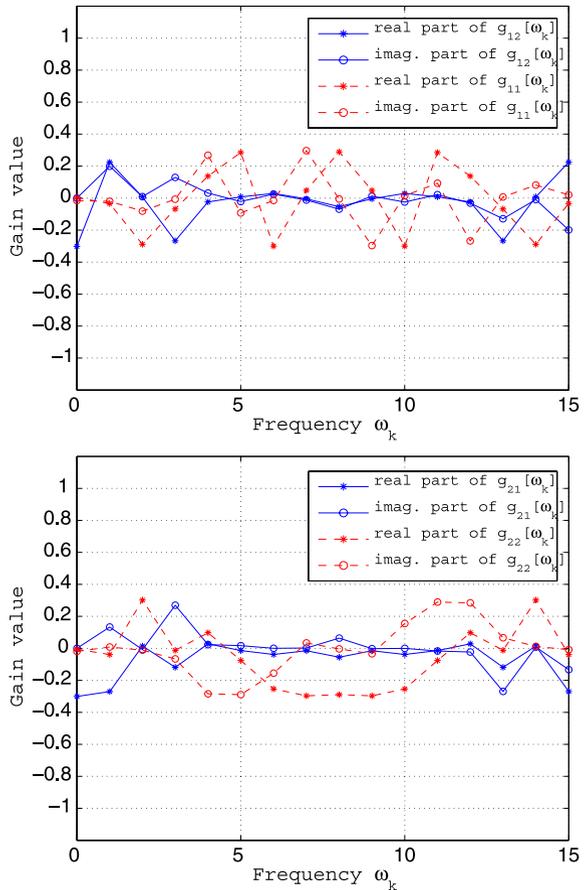
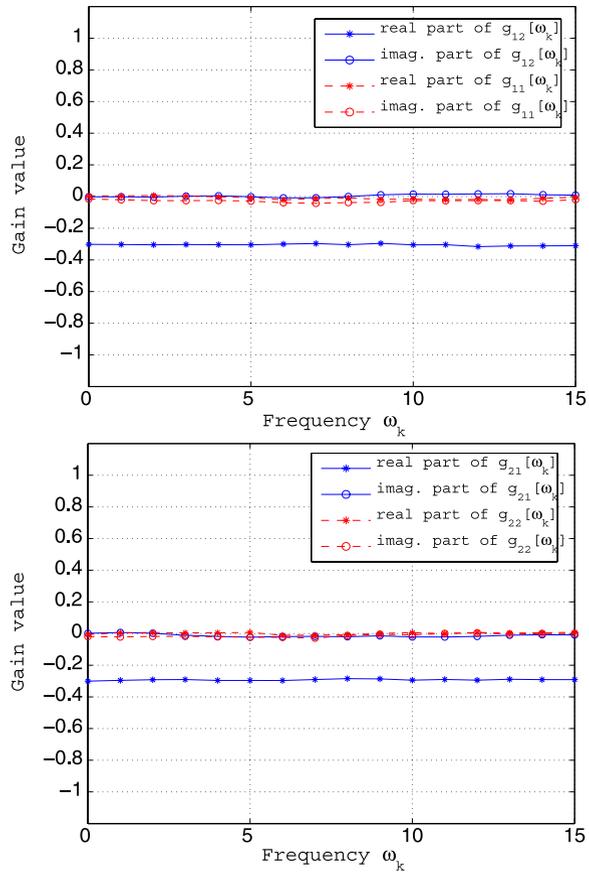


Fig. 3 Final gains obtained using the FD-BSS Cascade system with $r = 0$



In order to show the importance of solving the permutation/gain indeterminacy, we have performed an experiment using JADE to separate the frequency-domain mixtures in the 16 frequencies. The blocks contain 5,000 symbols of the frequency-domain sources. Figure 2 shows the real and the imaginary parts of the coefficients into the 2×2 gain matrices $\mathbf{G}[\omega_k]$, $k = 0, \dots, 15$. It is apparent that some frequency-domain sources have been recovered in a different order and with different complex-value gains. As a result, the sources have not been recuperated using the IDFT, as can be seen in Fig. 4(a).

For this experiment, Fig. 3 shows the gains obtained in all the frequencies using the FD-BSS cascade system proposed in Sect. 3 taking $r = 0$ as a reference. It is apparent that the frequency-domain sources have been recovered in the same order and with the same gain. This allows an adequate estimation of the time-domain sources to be obtained, as can be seen in Fig. 4(b).

In the second experiment, we have compared the error probability obtained using the following strategies:

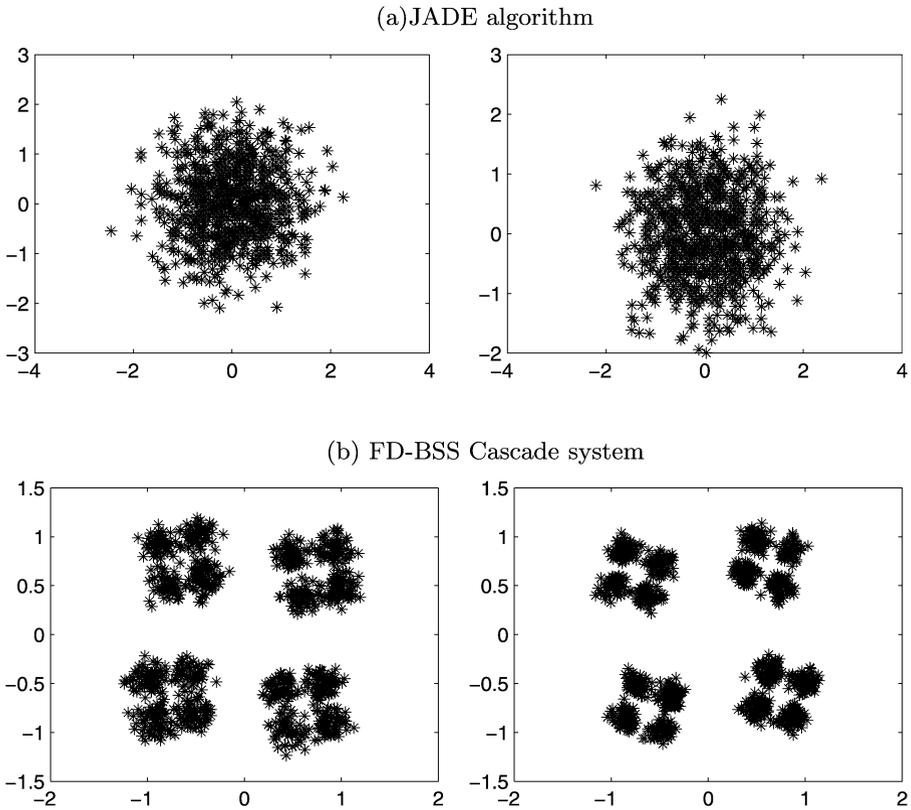


Fig. 4 Recovered sources obtained in the simulations. Note that the signal constellations are not recovered when the JADE algorithm is used to separate the mixtures in all frequencies

- The mixtures in all the frequencies are separated by using JADE. In this case, the permutation/gain ambiguity is not eliminated.
- The permutation indeterminacy is solved using the method proposed in [12], and the gain indeterminacy is solved using the method proposed in [7].
- Both the permutation and the gain indeterminacies are solved using the approach proposed in [7].
- FD-BSS cascade system with $r = 0$.
- FD-BSS cascade system selecting the reference frequency using the procedure described in Sect. 3.2 with $\beta = 0.02$. In order to clarify the selection of parameter β , we have performed 1,000 independent simulations varying the reference frequency from 0 to 15. We have evaluated the eigenvalue spread (22) in Step 5 of JADE and the final error probability. The block size has been 5,000 symbols. The results are plotted in Fig. 5. Note that a poor performance has been obtained in some cases where the eigenvalue spread is close to 0.01. For this reason, we have selected $\beta = 0.02$.

Table 1 shows the error probability obtained for 1,000 independent simulations and considering blocks of 3,000, 5,000, and 7,000 symbols. It is apparent that the

Fig. 5 Error probability versus eigenvalue spread

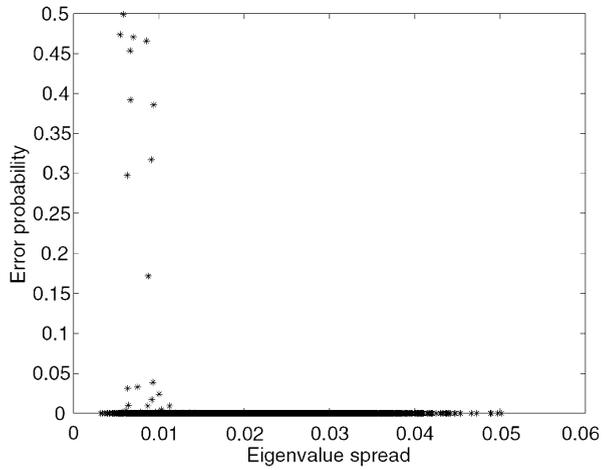


Table 1 Comparison between frequency-domain BSS approaches in terms of error probability for various block sizes

Approach	Block size		
	3,000	5,000	7,000
JADE	0.6783	0.6779	0.6753
JADE & Permutation [12] & Gain [7]	1.68×10^{-2}	1.40×10^{-3}	2.68×10^{-5}
JADE & Permutation [7] & Gain [7]	3.68×10^{-2}	5.40×10^{-3}	7.41×10^{-4}
FD-BSS Cascade $r = 0$	1.03×10^{-2}	2.30×10^{-3}	2.39×10^{-5}
FD-BSS Cascade $\beta = 0.02$	3.10×10^{-3}	8.45×10^{-7}	0

Table 2 Comparison between frequency-domain BSS approaches in terms of error probability and processing time

	JADE	JADE & Permutation [12] & Gain [7]	JADE & Permutation [7] & Gain [7]	FD-BSS Cascade ($r = 0$)	FD-BSS Cascade ($\beta = 0.02$)
Error probability	0.67	1.40×10^{-3}	5.40×10^{-3}	2.30×10^{-3}	8.45×10^{-7}
Time (sec.)	1.9440	2.3216	2.1953	0.1950	0.3230

FD-BSS cascade system provides a better performance than the other approaches. The improvement is considerably greater when the reference frequency is selected using the procedure described in Sect. 3.2 with $\beta = 0.02$.

We have also evaluated the time required to recover 5,000 frequency-domain symbols of each source, this being the intermediate value in our experiment. For the case of the FD-BSS cascade system with parameter $\beta = 0.02$, we have averaged the results of 1,000 independent simulations. Table 2 shows the error probability and the time measured using MATLAB code running in a Centrino Duo Processor at 1.66 GHz. The low computational load of the FD-BSS cascade system is apparent. Notice that using $\beta = 0.02$ the time remains low and the probability error is considerably reduced.

6 Conclusions

This paper proposes a simple technique to separate convolutive mixtures of statistically independent sources without resorting to an a priori knowledge of the sources or the mixing system. It has been shown that, in those applications where there is statistical dependence between the frequency-domain sources in different frequencies, it is possible to solve the permutation/gain indeterminacy by using a simple MSE-based unsupervised strategy. The basic idea is to separate the instantaneous mixtures using a cascade strategy beginning at a specific frequency (reference frequency), which is separated using a well-known unsupervised algorithm. The other instantaneous mixtures are then separated using the MSE criterion. Finally, the time-domain sources are recovered by applying the IDFT to the frequency-domain outputs.

The separation degree obtained at the reference frequency is an important issue, since any errors are propagated through the cascade system. In order to separate the sources at this frequency, we have used the blind identification algorithm JADE, which works with complex-valued signals and whose convergence does not depend on nonlinearities. We have determined that JADE provides an adequate separation when the eigenvalue spread of the matrix to be diagonalized is greater than a given threshold. This simple selection rule produces a very good performance in comparison with previous approaches. Simulation results show that the proposed system is a promising strategy for signal detection and wideband digital communication. For these applications, however, the proposed scheme needs to be extended to take into account other undesirable effects such as noise and synchronization errors.

References

1. E. Bingham, A. Hyvärinen, A fast fixed-point algorithm for independent component analysis of complex-valued signals. *Int. J. Neural Syst.* **10**(1), 1–8 (2000)
2. X. Cao, R. Wen-Liu, General approach to blind source separation. *IEEE Trans. Signal Process.* **44**(4), 562–571 (1996)
3. V. Capdevielle, C. Serviére, J.L. Lacoume, Blind separation of wide-band sources in the frequency domain, in *Proc. ICASSP 1995* (1995), pp. 2080–2083
4. J.F. Cardoso, A. Souloumiac, Blind beamforming for non-Gaussian signals. *IEE Proc. F* **140**(6), 362–370 (1993)
5. J.F. Cardoso, A. Souloumiac, Jacobi angles for simultaneous diagonalization. *SIAM J. Matrix Anal. Appl.* **17**(1), 161–164 (1996)
6. P. Comon, C. Jutten, J. Héroult, Blind separation of sources. Part II: Problems statement. *Signal Process.* **24**(1), 11–20 (1991)
7. A. Dapena, L. Castedo, A novel frequency domain approach for separating convolutive mixtures of temporally white signals. *Digit. Signal Process.* **13**(2), 304–316 (2003)
8. G. Gelle, M. Colas, C. Serviére, BSS for fault detection and machine monitoring time or frequency domain approach?, in *Proc. ICA 2000*, June (2000), pp. 555–560
9. J. Héroult, C. Jutten, B. Ans, Détection de grandeurs primitives dans un message composite par une architecture neuromimétique en apprentissage non supervisé, in *Actes 10ème Colloque GRETSI* (1985), pp. 1017–1022
10. A. Hyvärinen, E. Oja, Independent component analysis: algorithms and applications. *Neural Netw.* **13**(4–5), 411–430 (2000)
11. A. Mansour, N. Benckrouw, C. Gervaise, Blind separation of underwater acoustic signals, in *ICA'06* (2006), pp. 181–188
12. C. Mejuto, A. Dapena, L. Castedo, Frequency-domain infomax for blind separation of convolutive mixtures, in *ICA'00* (2000), pp. 315–320

13. N. Murata, S. Ikeda, A. Ziehe, An approach to blind source separation based on temporal structure of speech signals. *Neurocomputing* **41**(1–4), 1–24 (2001)
14. M. Pedersen, J. Larsen, U. Kjems, L. Parra, A survey of convolutive blind source separation methods, in *Springer Handbook on Speech Processing and Speech Communication* (Springer, Berlin, 2007)
15. C. Servière, Separation of speech signals under reverberant conditions, in *EUSIPCO 2004* (2004), pp. 1693–1696