Blind Source Separation Based on Time–Frequency Signal Representations

Adel Belouchrani and Moeness G. Amin, *Senior Member, IEEE*

***Abstract*—Blind source separation consists of recovering a set of signals of which only instantaneous linear mixtures are observed. Thus far, this problem has been solved using statistical information available on the source signals. This paper introduces a new blind source separation approach exploiting the difference in the time–frequency $(t\text{-}f)$ signatures of the sources to be separated. The approach is based on the diagonalization of a combined set of "spatial $t\text{-}f$ distributions." In contrast to existing techniques, the proposed approach allows the separation of Gaussian sources with identical spectral shape but with different $t\text{-}f$ localization properties. The effects of spreading the noise power while localizing the source energy in the $t\text{-}f$ domain amounts to increasing the robustness of the proposed approach with respect to noise and, hence, improved performance. Asymptotic performance analysis and numerical simulations are provided.**

I. Introduction

Blind SOURCE separation is an emerging field of fundamental research with a broad range of applications. It is motivated by practical problems that involve several source signals and several sensors. Each sensor receives a linear mixture of the source signals. The problem of the blind source separation consists, then, of recovering the original waveforms of the sources without any knowledge of the mixture structure. This mixture is often a convolutive mixture. However, in this paper, our main concern is the blind identification of an instantaneous linear mixture, which corresponds to a linear memoryless channel. This choice is motivated not only by the fact that such a model is mathematically tractable but also by the applicability to various areas, including semiconductor manufacturing process [1], factor analysis [2], narrowband signal processing [3], and image reconstruction [4].

Thus far, the problem of the blind source separation has been solved using statistical information available on the source signals. The first solution to the source separation problem was proposed almost a decade ago [5] and was based on the cancellation of higher order moments assuming non-Gaussian and i.i.d source signals. Since then, other criteria based on minimizations of cost functions, such as the sum of square fourth-order cumulants [6]–[8], contrast functions [7], [9], or likelihood function [10], have been used by several researchers. Note that in the case of non-i.i.d source signals and even Gaussian sources, solutions based on second-order statistics are possible [11], [12].

Matsuaka *et al.* have shown that the problem of the separation of nonstationary signals can be solved using second-order decorrelation only [14]. They implicitly use the nonstationarity of the signal via a neural net approach. In this paper, we propose to take advantage explicitly of the nonstationarity property of the signals to be separated. This is done by resorting to the powerful tool of time–frequency (t-f) signal representations. The underlying problem can then be posed as a signal synthesis [15] from the t-f plane with the incorporation of the spatial information provided by the multisensor array. With the proposed approach, no masking is required, and the cross terms no longer represent ambiguity in the synthesis of multicomponent signals.

This paper introduces a new blind identification technique based on a joint diagonalization of a combined set of spatial t-f distributions (STFD's) that are a generalization of the t-f distribution to a vector signal. It is shown that under the linear data model the proposed STFD has the similar structure than the data spatial correlation matrix that we commonly use in array signal processing. The benefits of STFD over the spatial correlation matrix in a nonstationary signal environment is the direct exploitation of the information brought by the nonstationarity of the signals. Hence, the new approach exploits the difference between the t-f signatures of the sources. This method presents a number of attractive features. In contrast to blind source separation approaches using second-order and/or high order statistics, the proposed approach allows the separation of Gaussian sources with identical spectral shape but with different t-f localization properties. Moreover, the effects of spreading the noise power while localizing the source energy in the t-f domain amounts to increasing the robustness of the proposed approach with respect to noise.

The paper is organized as follows. In Section II, the problem of blind source separation is stated along with the relevant hypothesis. Spatial t-f distributions are introduced in Section III. Section IV presents the proposed t-f blind identification technique based on the diagonalization of a combined set of spatial t-f distributions (TFD's). In Section V, a closed-form expression of the asymptotic performance of the proposed method is derived. Numerical simulations illustrating the usefulness of the proposed technique are given in Section VI.

Manuscript received September 6, 1996; revised April 10, 1998. This work was supported by Rome Laboratories, Rome, NY, under Contract F30602-96-C-0077. The associate editor coordinating the review of this paper and approving it for publication was Prof Luis F. Chaparro.

A. Belouchrani is with the Electrical Engineering Department, National Polytechnic School of Algiers, Algiers, Algeria.

M. G. Amin is with the Department of Electrical and Computer Engineering, Villanova University, Villanova, PA 19085 USA (e-mail: moeness@ece.vill.edu).

Publisher Item Identifier S 1053-587X(98)07799-X.

II. Problem Formulation

A. Data Model

In most practical situations, we have to process multidimensional observations of the form

$$\mathbf{x}(t) = \mathbf{y}(t) + \mathbf{n}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \tag{1}$$

where $\mathbf{x}(t)$ is a noisy instantaneous linear mixture of source signals, and $\mathbf{n}(t)$ is the additive noise. This model is commonly used in the field of narrowband array processing. In this context, the vector $\mathbf{s}(t) = [s_1(t), \ldots, s_n(t)]^T$ consists of the signals emitted by n narrowband sources, whereas the vector $\mathbf{y}(t) = [y_1(t), \ldots, y_m(t)]^T$ contains the array output. Both vectors are sampled at time t. Matrix $\mathbf{A}$ is the transfer function between the sources and the array sensors. In the following, this matrix is referred to as the "array matrix" or the "mixing matrix."

B. Assumptions

The source signal vector $\mathbf{s}(t)$ is assumed to be a nonstationary multivariate process with

$$\text{H1): } \mathbf{R}_s = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbf{s}(t+\tau)\mathbf{s}^*(t) = \operatorname{diag}[r_{11}(\tau), \ldots, r_{nn}(\tau)] \tag{2}$$

where superscript * denotes the conjugate transpose of a vector, diag$[\cdot]$ is the diagonal matrix formed with the elements of its vector valued argument, and $r_{ii}(\tau) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1,T} \mathbf{s}_i(t+\tau)\mathbf{s}_i^*(t)$ denotes the autocorrelation of $\mathbf{s}_i(t)$. Assumption H1) means that the component $\mathbf{s}_i(t)$, $1 \le i \le n$ are mutually uncorrelated as their cross-correlations are equal to zero.

The additive noise $\mathbf{n}(t)$ is modeled as a stationary, temporally white, zero-mean complex random process independent of the source signals. For simplicity, we also require $\mathbf{n}(t)$ to be spatially white, i.e.,

$$(H2)\text{: } E(\mathbf{n}(t+\tau)\mathbf{n}^*(t)) = \sigma^2\delta(\tau)\mathbf{I} \tag{3}$$

where $\delta(\tau)$ is the Kronecker delta, and $\mathbf{I}$ denotes the identity matrix.

The $m \times n$ complex matrix $\mathbf{A}$ is assumed to have full column rank but is otherwise unknown. In contrast with traditional parametric methods, no specific array geometry or sensor characteristics are assumed, i.e., the array manifold is unknown.

The aim of blind source separation is to identify the mixture matrix and/or to recover the source signals from the array output $\mathbf{x}(t)$ *without any a priori knowledge of the array manifold*.

C. Problem Indeterminacies

This problem of blind source separation has two inherent ambiguities. First, it is not possible to know the original labeling of the sources; hence, any permutation of the estimated sources is also a satisfactory solution. The second ambiguity is that it is inherently impossible to uniquely identify the source signals. This is because the exchange of a fixed scalar factor between a source signal and the corresponding column of the mixture matrix $\mathbf{A}$ does not affect the observations as is shown by

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) = \sum_{i=1}^{n} \frac{\mathbf{a}_i}{\alpha_i}\alpha_i s_i(t) + \mathbf{n}(t) \tag{4}$$

where α_i is an arbitrary complex factor, and $\mathbf{a}_i$ denotes the ith column of $\mathbf{A}$.

We take advantage of the second indeterminacy by treating the source signals as if they have unit power so that the dynamic range of the sources is accounted for by the magnitude of the corresponding columns of $\mathbf{A}$. This normalization *convention* turns out to be convenient in the sequel; it does not affect the performance results presented below. Since the sources are assumed to be uncorrelated, we have

$$\mathbf{R}_s = \mathbf{I} \text{ so that } \mathbf{R}_y \stackrel{\text{def}}{=} \lim_{T\to\infty} \frac{1}{T} \sum_{t=1,T} \mathbf{y}(t)\mathbf{y}^*(t) = \mathbf{A}\mathbf{A}^H \tag{5}$$

where the superscript ^{H} denotes the complex conjugate transpose of a matrix. This normalization still leaves undetermined the ordering and the phases of the columns of $\mathbf{A}$. Hence, the blind source separation must be understood as the identification of the mixing matrix and/or the recovering of the source signals up to a fixed permutation and some complex factors.

III. Spatial Time–Frequency Distributions

The discrete-time form of the Cohen's class of TFD's, for signal $x(t)$, is given by [16]

$$D_{xx}(t,f) = \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \phi(m,l)x(t+m+l) \times x^*(t+m-l)e^{-j4\pi fl} \tag{6}$$

where t and f represent the time index and the frequency index, respectively. The kernel $\phi(m,l)$ characterizes the distribution and is a function of both the time and lag variables. The cross-TFD of two signals $x_1(t)$ and $x_2(t)$ is defined by

$$D_{x_1x_2}(t,f) = \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \phi(m,l)x_1(t+m+l) \times x_2^*(t+m-l)e^{-j4\pi fl}. \tag{7}$$

Expressions (6) and (7) are now used to define the following data *spatial t-f distribution* (STFD) *matrix*,

$$\mathbf{D}_{\mathbf{xx}}(t,f) = \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \phi(m,l)\mathbf{x}(t+m+l) \times \mathbf{x}^*(t+m-l)e^{-j4\pi fl} \tag{8}$$

where $[\mathbf{D}_{\mathbf{xx}}(t,f)]_{ij} = D_{x_ix_j}(t,f)$, for $i,j = 1,\ldots,n$.

A more general definition of the STFD matrix can be given as

$$\mathbf{D}_{\mathbf{xx}}(t,f) = \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \mathbf{\Phi}(m,l) \odot \mathbf{x}(t+m+l) \times \mathbf{x}^*(t+m-l)e^{-j4\pi fl} \tag{9}$$

where $\odot$ designates the Hadamard product, and $[\mathbf{\Phi}(m,l)]_{ij} = \phi_{ij}(m,l)$ is the t-f kernel associated with the pair of the sensor data $x_i(t)$ and $x_j(t)$.

Under the linear data model of (1) and assuming noise-free environment, the STFD matrix takes the simple structure in

$$\mathbf{D_{xx}}(t,f) = \mathbf{A}\mathbf{D_{ss}}(t,f)\mathbf{A}^H \tag{10}$$

where $\mathbf{D_{ss}}(t,f)$ is the signal TFD matrix whose entries are the auto- and cross-TFD's of the sources. We note that $\mathbf{D_{xx}}(t,f)$ is of dimension $m \times m$, whereas $\mathbf{D_{ss}}(t,f)$ is of $n \times n$ dimension. For narrowband array signal processing applications, matrix $\mathbf{A}$ holds the spatial information and maps the auto- and cross-TFD's of the sources into auto- and cross-TFD's of the data.

Expression (10) is similar to that which commonly used in blind source separation [12] and direction-of-arrival (DOA) estimation problems, relating the signal correlation matrix to the data spatial correlation matrix. The two subspaces spanned by the principle eigenvectors of $\mathbf{D_{xx}}(t,f)$ and the columns of A are, therefore, identical. Since the off-diagonal elements are crossterms of $\mathbf{D_{ss}}(t,f)$, then this matrix is diagonal for each t-f point that corresponds to a true power concentration, i.e., signal auto-term. In the sequel, we consider the t-f points that satisfy this property. In practice, to simplify the selection of autoterms, we apply a smoothing kernel $\phi(m,l)$ that significantly decreases the contribution of the crossterms in the t-f plane. This kernel can be a member of the reduced-interference distribution (RID) introduced in [17], or it can be signal dependent, which matches the underlying signal characteristics [18]–[20].

IV. A Two-Step Blind Identification Approach

In this section, we present a new blind identification approach based on two step processing; the first step consists of whitening the data in order to transform the mixing matrix $\mathbf{A}$ into a unitary matrix. The second step consits then of retreiving this unitary matrix by joint diagonalizing a set of data-STFD matrices.

A. First Step

The first processing step consists of whitening the signal part $\mathbf{y}(t)$ of the observation. This is achieved by applying a *whitening matrix* $\mathbf{W}$ to $\mathbf{y}(t)$, i.e., an $n \times m$ matrix satisfying

$$\begin{aligned}\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbf{W}\mathbf{y}(t)\mathbf{y}^*(t)\mathbf{W}^H &= \mathbf{W}\mathbf{R}_y\mathbf{W}^H \\ &= \mathbf{W}\mathbf{A}\mathbf{A}^H\mathbf{W}^H = I \end{aligned} \tag{11}$$

where $\mathbf{R}_y = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T}\mathbf{y}(t)\mathbf{y}^*(t)$ is the autocorrelation matrix of the noiseless array output. Equation (11) shows that if $\mathbf{W}$ is a whitening matrix, then $\mathbf{WA}$ is a $n \times n$ unitary matrix. It follows that for any whitening matrix $\mathbf{W}$, there exists a $n \times n$ unitary matrix $\mathbf{U}$ such that $\mathbf{WA} = \mathbf{U}$. As a consequence, matrix $\mathbf{A}$ can be factored as

$$\mathbf{A} = \mathbf{W}^{\#}\mathbf{U} \tag{12}$$

where the superscript $^{\#}$ denotes the Moore–Penrose pseudoinverse. This whitening procedure reduces the determination of the $m \times n$ mixture matrix $\mathbf{A}$ to that of a unitary $n \times n$ matrix $\mathbf{U}$. The whitened data vector $\mathbf{z}(t) = \mathbf{W}\mathbf{x}(t)$ still obeys a linear model

$$\mathbf{z}(t) \overset{\text{def}}{=} \mathbf{W}\mathbf{x}(t) = \mathbf{W}(\mathbf{A}\mathbf{s}(t) + \mathbf{n}(t)) = \mathbf{U}\mathbf{s}(t) + \mathbf{W}\mathbf{n}(t). \tag{13}$$

The signal part of the whitened process now is a "unitary mixture" of the source signals. Note that all information contained in the autocorrelation matrix is "exhausted" after the whitening procedure in the sense that changing $\mathbf{U}$ in (13) to any other unitary matrix leaves the autocorrelation matrix of $\mathbf{z}(t)$ unchanged. Note also that besides whitening the signal part of the observations, multiplication by a whitening matrix $\mathbf{W}$ reduces the array output to an n-dimensional vector.

Under the assumption of the linear model (1), the data autocorrelation matrix has the following structure:

$$\mathbf{R} \overset{\text{def}}{=} \lim_{T\to\infty} \frac{1}{T} \sum_{t=1,T} \mathbf{x}(t)\mathbf{x}^*(t) = \mathbf{R}_y + \sigma^2\mathbf{I}. \tag{14}$$

Combining (14) and (5), we have

$$\mathbf{A}\mathbf{A}^H = \mathbf{R} - \sigma^2\mathbf{I} \tag{15}$$

Hence, from (11) and (15), the whitening matrix $\mathbf{W}$ can be determined from the array output autocorrelation $\mathbf{R}$. As shown by (12), finding a whitening matrix still leaves undetermined some unitary factor in $\mathbf{A}$. This "missing factor" $\mathbf{U}$ can be determined from higher order statistics, as investigated in [6], [21], [7], or from second order statistics as proposed in [12]. As explained below, by exploiting the t-f dependence structure (10), the missing rotation may be also retrieved from spatial time frequency distributions at properly chosen t-f points.

B. Second Step

By pre and postmultiplying the STFD matrices $\mathbf{D_{xx}}(t,f)$ by $\mathbf{W}$, we obtain the *whitened STFD-matrices* as

$$\mathbf{D_{zz}}(t,f) = \mathbf{W}\mathbf{D_{xx}}(t,f)\mathbf{W}^H \tag{16}$$

which is, in essence, the STFD of the whitened data vector $\mathbf{z}$. From the definition of $\mathbf{W}$ and (10), we may express $\mathbf{D_{zz}}(t,f)$ as

$$\mathbf{D_{zz}}(t,f) = \mathbf{U}\mathbf{D_{ss}}(t,f)\mathbf{U}^H. \tag{17}$$

Since the matrix $\mathbf{U}$ is unitary and $\mathbf{D_{ss}}(t,f)$ is diagonal, (17) shows that any whitened data STFD-matrix is diagonal in the basis of the columns of the matrix $\mathbf{U}$ (the eigenvalues of $\mathbf{D_{zz}}(t,f)$ being the diagonal entries of $\mathbf{D_{ss}}(t,f)$). As a consequence, the missing unitary matrix $\mathbf{U}$ may be obtained as a unitary diagonalizing matrix of a whitened STFD matrix for some t-f point corresponding to a signal autoterm. More formally, we have the folowing theorem.

Theorem 1 (First Identifiability Condition): Let (t_a, f_a) be a t-f point corresponding to a signal autoterm and $\mathbf{V}$ be a unitary matrix such that

$$\mathbf{V}^H \mathbf{D}_{\mathbf{zz}}(t_a, f_a)\mathbf{V} = \text{diag}[d_1, \ldots, d_n]. \tag{18}$$

$$\text{For all } 1 \leq i \neq j \leq n \quad D_{s_i s_i}(t_a, f_a) \neq D_{s_j s_j}(t_a, f_a). \tag{19}$$

Then

- $\mathbf{V}$ is equal to $\mathbf{U}$ up to a phase and a permutation of its columns;
- there exists a permutation γ on $\{1, \ldots n\}$ such that

$$\left[D_{s_1 s_1}(t_a, f_a), \ldots, D_{s_n s_n}(t_a, f_a)\right] = [d_{\gamma(1)}, \ldots, d_{\gamma(n)}].$$

This is a direct consequence of the spectral theorem for normal matrices (see [22, Th. 2.5.4]). We recall that an $n \times n$ matrix $\mathbf{M}$ is said to be *normal* if $\mathbf{M}\mathbf{M}^H = \mathbf{M}^H\mathbf{M}$. According to Theorem 1, for the (t_a, f_a) point, if the diagonal elements of $\mathbf{D}_{\mathbf{ss}}(t_a, f_a)$ are all distinct, the missing unitary matrix $\mathbf{U}$ may be "uniquely" (i.e. up to permutation and phase shifts) retrieved by computing the eigendecomposition of $\mathbf{D}_{\mathbf{zz}}(t_a, f_a)$. However, when the t-f signatures of the different signals are not highly overlapping or frequently intersecting, which is likely to be the case, the selected (t_a, f_a) point often corresponds to a single signal auto-term, rendering matrix $\mathbf{D}_{\mathbf{ss}}(t_a, f_a)$ deficient. That is, only one diagonal element of $\mathbf{D}_{\mathbf{ss}}(t_a, f_a)$ is different from zero. It follows that the determination of the matrix $\mathbf{U}$ from the eigendecomposition of a single whitened data STFD-matrix is no longer "unique" in the sense defined above.

The situation is more favorable when considering *simultaneous diagonalization* of a combined set $\{\mathbf{D}_{\mathbf{zz}}(t_i, f_i) \mid i = 1, \ldots, p\}$ of p STFD matrices. This amounts to incorporating several t-f points in the source separation problem.

Theorem 2 (Second Identifiability Condition): Let (t_1, t_1), $(t_2, t_2), \ldots, (t_K, t_K)$ be K t-f points corresponding to signal autoterms, and let $\mathbf{V}$ be a unitary matrix such that

$$\begin{aligned}&\text{For all } 1 \leq k \leq K\\ &\quad \mathbf{V}^H \mathbf{D}_{\mathbf{zz}}(t_k, f_k)\mathbf{V} = \text{diag}[d_1(k), \ldots, d_n(k)] \qquad (20)\\ &\text{For all } 1 \leq i \neq j \leq n \quad \text{there exists } k, \quad 1 \leq k \leq K\\ &\quad D_{s_i s_i}(t_k, f_k) \neq D_{s_j s_j}(t_k, f_k). \qquad (21)\end{aligned}$$

Then

- $\mathbf{V}$ is equal to $\mathbf{U}$ up to a phase and a permutation of its columns;
- there exists a permutation γ on $\{1, \ldots n\}$ such that

$$\begin{aligned}&\left[D_{s_1 s_1}(t_k, f_k), \ldots, D_{s_n s_n}(t_k, f_k)\right]\\ &\quad = [d_{\gamma(1)}(k), \ldots, d_{\gamma(n)}(k)] \quad 1 \leq k \leq K.\end{aligned}$$

This is a consequence of the uniqueness of the joint diagonalization: (see [12, Th. 3]). Again, the existence of a unitary matrix $\mathbf{V}$ that simultaneously diagonalizes the set of STFD matrices $[\mathbf{D}_{\mathbf{zz}}(t_1, f_1), \ldots, \mathbf{D}_{\mathbf{zz}}(t_K, f_K)]$ is guaranteed by (17) for any choice of t-f points corresponding to signal autoterms. Condition (21), although it is is weaker than condition (19), it is not always satisfied. In particular, in the trivial case where the sources show identical t-f signatures, the mixing matrix $\mathbf{A}$ cannot be identified by resorting to Theorem 2. Conversely, when the source signals have different t-f signatures, it is always possible to find a set of t-f points corresponding to signal autoterms such that condition (21) is met.

C. Joint Diagonalization

The *joint diagonalization* [12], [13] can be explained by first noting that the problem of the diagonalization of a single $n \times n$ normal matrix $\mathbf{M}$ is equivalent to the minimization of the criterion [23]

$$C(\mathbf{M}, \mathbf{V}) \stackrel{\text{def}}{=} -\sum_i |\mathbf{v}_i^* \mathbf{M} \mathbf{v}_i|^2 \tag{22}$$

over the set of unitary matrices $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$. Hence, the joint diagonalization of a set $\{\mathbf{M}_k \mid k = 1 \cdots K\}$ of K arbitrary $n \times n$ matrices is defined as the minimization of the following JD criterion:

$$C(\mathbf{V}) \stackrel{\text{def}}{=} -\sum_k C(\mathbf{M}_k, \mathbf{V}) = -\sum_{ki} |\mathbf{v}_i^* \mathbf{M}_k \mathbf{v}_i|^2 \tag{23}$$

under the same unitary constraint.

It is important to note that the above definition of joint diagonalization does *not* require the matrix set under consideration to be exactly and simultaneously diagonalized by a single unitary matrix. This is because we do not require the off-diagonal elements of all the matrices to be cancelled by a unitary transform; a joint diagonalizer is simply a minimizer of the JD criterion. If the matrices in $\{\mathbf{M}_k \mid k = 1 \cdots K\}$ are not exactly joint diagonalizable, the JD criterion cannot be zeroed, and the matrices can only be approximately joint diagonalized. Hence, an (approximate) joint diagonalizer defines a kind of an "average eigenstructure." Note that a numerically efficient algorithm for solving (23) exists in [12] and is based on a generalization of the Jacobi technique [23].

D. Summary

Based on the previous sections, we can introduce a time–frequency separation (TFS) algorithm. The TFS is defined by the following implementation.

1) Estimate the autocorrelation matrix $\hat{\mathbf{R}}$ from T data samples. Denote by $\lambda_1, \ldots, \lambda_n$ the n largest eigenvalues and $\mathbf{h}_1, \ldots, \mathbf{h}_n$ the corresponding eigenvectors of $\hat{\mathbf{R}}$.
2) Under the white noise assumption, an estimate $\hat{\sigma}^2$ of the noise variance is the average of the $m - n$ smallest eigenvalues of $\hat{\mathbf{R}}$. The whitened signals are $\mathbf{z}(t) = [z_1(t), \ldots, z_n(t)]^T$ computed by $z_i(t) = (\lambda_i - \hat{\sigma}^2)^{-\frac{1}{2}} \; \mathbf{h}_i^* \mathbf{x}(t)$ for $1 \leq i \leq n$. This is equivalent to forming a whitening matrix by $\hat{\mathbf{W}} = [(\lambda_1 - \hat{\sigma}^2)^{-\frac{1}{2}} \mathbf{h}_1, \ldots, (\lambda_n - \hat{\sigma}^2)^{-\frac{1}{2}} \mathbf{h}_n]^H$.
3) Form K matrices by computing the STFD of $\mathbf{z}(t)$ for a fixed set of (t_i, f_i) points, $i = 1, \ldots, K$, corresponding to signal autoterms.
4) A unitary matrix $\hat{\mathbf{U}}$ is then obtained as joint diagonalizer of the set $\{\mathbf{D}_{\mathbf{zz}}(t_i, f_i) \mid i = 1, \ldots, K\}$.
5) The source signals are estimated as $\hat{\mathbf{s}}(t) = \hat{\mathbf{U}}^H \hat{\mathbf{W}} \mathbf{x}(t)$, and/or the mixing matrix $\mathbf{A}$ is estimated as $\hat{\mathbf{A}} = \hat{\mathbf{W}}^\# \hat{\mathbf{U}}$.

V. Asymptotic Performance Analysis

In this section, an asymptotic performance analysis of the proposed method is carried out. To ease the derivations, we make the following two assumptions.

H1′) Each source signal $s_i(t)$ is a deterministic sequence.

H2′) The additive noise $\mathbf{n}(t)$ is complex circular Gaussian process.

To eliminate the phase and permutation indeterminacies, we shall assume that they are fixed in such a way that the matrix estimate $\hat{\mathbf{A}}$ is closer to the true mixture matrix $\mathbf{A}$ than to some other matrix equal to $\mathbf{A}$ up to a phase and a permutation of its columns. In addition, the STFD are computed at a set of t-f points $\{(t_i, f_i), i = 1, \ldots, K\}$ such that the identifiability condition of Theorem 2 is satisfied.

A. Performance Index

Rather than estimating the variance of the coefficients of the mixing matrix, it is more relevant with respect to the source separation issue to compute an index that quantifies the performance in terms of interference rejection, as follows. Assume that at each time instant t, an estimate of the vector of source signals is computed by applying the pseudoinverse of the estimated mixture matrix to the received signal $\mathbf{x}(t)$, i.e.,

$$\hat{\mathbf{s}}(t) = \hat{\mathbf{A}}^{\#}\mathbf{x}(t) = \hat{\mathbf{A}}^{\#}\mathbf{A}\mathbf{s}(t) + \hat{\mathbf{A}}^{\#}\mathbf{n}(t) \tag{24}$$

where $\hat{\mathbf{A}}^{\#} = \hat{\mathbf{U}}^H\hat{\mathbf{W}}$. We stress that in general, this procedure is not optimal for recovering the source signals based on an estimate $\hat{\mathbf{A}}$. For large enough sample size T, matrix $\hat{\mathbf{A}}$ should be close to the true mixing matrix $\mathbf{A}$ so that $\hat{\mathbf{A}}^{\#}\mathbf{A}$ is close to the identity matrix. The performance index used in the sequel is the interference-to-signal ratio (ISR), which is defined as

$$\mathcal{I}_{pq} = E|(\hat{\mathbf{A}}^{\#}\mathbf{A})_{pq}|^2. \tag{25}$$

This actually defines an ISR because by our normalization convention (5), we have $\mathcal{I}_{pp} \simeq 1$ for large enough T. Thus, $\mathcal{I}_{pq}$ measures the ratio of the power of the interference of the qth source to the power of the pth source signal estimated as in (25). As a measure of the global quality of the separation, we also define a global rejection level

$$\mathcal{I}_{\text{perf}} \stackrel{\text{def}}{=} \sum_{q \neq p} \mathcal{I}_{pq}. \tag{26}$$

B. Outline of the Performance Analysis

The purpose of this section is to give a closed-form expression of the mean rejection level (25). Giving the details of a rigorous proof goes far beyond the scope of this paper. We present only an outline of the derivation below with additional mathematical details in the Appendix.

The matrix estimate $\hat{A}$ is a "function" of the data autocorrelation matrix $\mathbf{R}$, and the STFD matrices $(\mathbf{D_{xx}}(t_1, f_1), \mathbf{D_{xx}}(t_2, f_2), \ldots, \mathbf{D_{xx}}(t_K, f_K))$ of the observed signal $\mathbf{x}(t)$. The computation proceeds in two steps: First, we compute the leading term in the Taylor series expansion of $|(\hat{A}^{\#}A)_{pq}|^2$ (see Lemma 1). Then, by computing the expectation of the aforementioned leading term, we obtain the desired result.

Lemma 1: The Taylor series expansion of $|(\hat{\mathbf{A}}^{\#}\mathbf{A})_{pq}|^2$ is given for $p \neq q$ by

$$\begin{aligned}|(\hat{\mathbf{A}}^{\#}\mathbf{A})_{pq}|^2 = {} & |\alpha_{pq}C_{pq}|^2 + \sum_{k=1}^{K} \alpha_{pq}\alpha_{pq}(k)[C_{qp}C_{pq}(k) \\ & + C_{pq}C_{qp}(k)] + \sum_{k=1,l=1}^{K,K} \alpha_{pq}(k)\alpha_{pq}(l) \\ & \times C_{pq}(k)C_{qp}(l) \\ & + O\left(\|\delta\mathbf{R}\|^3 + \sum_{k=1}^{K} \|\delta\mathbf{D_{xx}}(t_k, f_k)\|^3\right)\end{aligned}$$

with

$$\begin{aligned}\mathbf{d}_r &= \left[D_{s_rs_r}(t_1, f_1), \ldots, D_{s_rs_r}(t_K, f_K)\right]^T \\ \alpha_{pq} &= 1 + \frac{|\mathbf{d}_p|^2 - |\mathbf{d}_q|^2}{|\mathbf{d}_p - \mathbf{d}_q|^2} \\ \alpha_{pq}(k) &= \frac{D^*_{s_ps_p}(t_k, f_k) - D^*_{s_qs_q}(t_k, f_k)}{|\mathbf{d}_p - \mathbf{d}_q|^2} \\ \mathbf{C} &= -\frac{1}{2}\mathbf{A}^{\#}\delta\mathbf{R}\mathbf{A}^{\#^H} + \frac{\mathrm{Tr}(\Pi\delta\mathbf{R})}{2(m-n)}(\mathbf{A}^H\mathbf{A})^{-1} \\ \mathbf{C}(k) &= \frac{1}{2}\mathbf{A}^{\#}\delta\mathbf{D_{xx}}(t_k, f_k)\mathbf{A}^{\#^H} \\ \delta\mathbf{R} &= \hat{\mathbf{R}} - \mathbf{R} \\ \delta\mathbf{D_{xx}}(t_k, f_k) &= \mathbf{D_{xx}}(t_k, f_k) - \mathbf{A}\,\mathrm{diag}\big[D_{s_1s_1}(t_k, f_k), \ldots \\ &\quad D_{s_ns_n}(t_k, f_k)\big]\mathbf{A}^H \\ &= \mathbf{A}\big\{\mathbf{D_{ss}}(t_k, f_k) - \mathrm{diag}\big[D_{s_1s_1}(t_k, f_k), \ldots \\ &\quad D_{s_ns_n}(t_k, f_k)\big]\big\}\mathbf{A}\mathbf{D_{sn}}(t_k, f_k) \\ &\quad + \mathbf{D_{ns}}(t_k, f_k)\mathbf{A}^H + \mathbf{D_{nn}}(t_k, f_k)\end{aligned}$$

where $\mathrm{Tr}(\cdot)$ is the trace of the matrix, Π denotes the orthogonal projector on the noise subspace (i.e., the subspace orthogonal to the range of matrix $\mathbf{A}$), and $C_{pq}(\cdot)$ is the pqth element of the matrix $\mathbf{C}(\cdot)$.

The proof of the above lemma follows closely the same steps used in [12]. For more details, see Appendix A.

The expectation of $|(\hat{\mathbf{A}}^{\#}\mathbf{A})_{pq}|^2$ can be computed from the expectations of $|C_{pq}|^2$, $C_{qp}C_{pq}(k)$, and $C_{qp}(l)C_{pq}(k)$. Conditions H1′) and H2′) reduce this computation to simple algebra, yielding

$$\begin{aligned}E(|C_{pq}|^2) = {} & \frac{1}{4}\left|r_{T\,pq}\right|^2 + \frac{1}{4T}\bigg[\sigma^2\left(r_{T\,pp}J_{qq} + r_{T\,qq}J_{pp}\right) \\ & + \sigma^4\left(J_{pp}J_{qq} + \frac{|J_{pq}|^2}{m-n}\right)\bigg] \\ E(C_{qp}C_{pq}(k)) = {} & -\frac{1}{4}r_{T\,qp}D_{s_ps_q}(t_k, f_k) - \frac{\sigma^2}{4}\bigg[\left|r_{T\,qp}\right|^2 J_{pq} \\ & + \frac{1}{T}\left(D_{s_ps_p}(t_k, f_k)J_{qq}\right. \\ & \left. + D_{s_qs_q}(t_k, f_k)J_{pp}\right)\bigg] - \frac{\sigma^4}{4T}J_{qq}J_{pp}\end{aligned}$$

$$\begin{aligned}E(C_{pq}(k)C_{qp}(l)) =& \frac{1}{4}D_{s_ps_q}(t_k,f_k)D_{s_qs_p}(t_l,f_l)\\&+\frac{1}{4}\big[\sigma^2\big(D_{s_ps_q}(t_k,f_k)J_{qp}\\&+D_{s_qs_p}(t_l,f_l)Jpq\\&+F_{s_ps_p}(k,l)J_{qq}+F_{s_qs_q}(l,k)J_{pp}\big)\\&+\sigma^4\big(|J_{pq}|^2+J_{pp}J_{qq}\phi_{kl}\big)\big]\end{aligned}$$

where we have set

$$\begin{aligned}r_{Tpq} &= \frac{1}{T}\sum_{t=1}^{T}s_p(t)s_q^*(t)\\ J_{pq} &= (\mathbf{A}^H\mathbf{A})^{-1}_{pq}\\ F_{s_ps_p}(k,l) &= \sum_{v'=-\infty}^{+\infty}\sum_{v=-\infty}^{+\infty}\left[\sum_{m=-\infty}^{+\infty}\phi(m,v)\right.\\&\quad\times\phi(m-v-v'+t_k-t_l,v')s_p(t_k+m+v)\\&\quad\left.\times\ s_p^*(t_k+m-v-2v')\right]e^{-j4\pi f_kv}e^{-j4\pi f_lv'}\\ \phi_{kl} &= \sum_{v=-\infty}^{+\infty}\left[\sum_{m=-\infty}^{+\infty}\phi(m,v)\phi^*(m+(t_k-t_l),v)\right]\\&\quad\times e^{-j4\pi(f_k-f_l)v}.\end{aligned}$$

It can be readily shown that [24]

$$\lim_{T\to\infty}TE\|\delta\mathbf{R}\|^3=0$$
$$\lim_{T\to\infty}TE\|\delta\mathbf{D_{xx}}(t_k,f_k)\|^3=0.$$

Using the above results, the ISR is asymptotically given by

$$\mathcal{I}_{pq}=\mathcal{I}_{pq}^0+\sigma^2\mathcal{I}_{pq}^1+\sigma^4\mathcal{I}_{pq}^2 \tag{27}$$

where the coefficients of the expansion are

$$\begin{aligned}\mathcal{I}_{pq}^0 =& \frac{1}{4}\left[\alpha_{pq}^2|r_{Tpq}|^2-\sum_{k=1}^{K}\alpha_{pq}\alpha_{pq}(k)\right.\\&\times\big[r_{Tpq}D_{s_qs_p}(t_k,f_k)+r_{Tqp}D_{s_ps_q}(t_k,f_k)\big]\\&\left.+\sum_{k=1,l=1}^{K,K}\alpha_{pq}(k)\alpha_{pq}(l)D_{s_ps_q}(t_k,f_k)D_{s_qs_p}(t_l,f_l)\right]\end{aligned}$$

$$\begin{aligned}\mathcal{I}_{pq}^1 =& \frac{1}{4}\left[\frac{\alpha_{pq}^2}{T}\big(r_{Tpp}J_{qq}+r_{Tqq}J_{pp}\big)\right.\\&-\sum_{k=1}^{K}\alpha_{pq}\alpha_{pq}(k)\Big[r_{Tpq}J_{qp}+r_{Tqp}J_{pq}\\&+\frac{2}{T}\big(D_{s_ps_p}(t_k,f_k)J_{qq}+D_{s_qs_q}(t_k,f_k)J_{pp}\big)\Big]\\&+\sum_{k=1,l=1}^{K,K}\alpha_{pq}(k)\alpha_{pq}(l)(D_{s_ps_q}(t_k,f_k)J_{qp}\\&\left.+D_{s_qs_p}(t_l,f_l)J_{pq}+F_{s_ps_p}(k,l)J_{qq}+F_{s_qs_q}(k,l)J_{pp}\right]\end{aligned}$$

$$\begin{aligned}\mathcal{I}_{pq}^2 =& \frac{1}{4}\left[\frac{1}{T}\left[\alpha_{pq}^2\left(J_{qq}J_{pp}+\frac{|J_{pq}|^2}{m-n}\right)\right.\right.\\&\left.-2\sum_{k=1}^{K}\alpha_{pq}\alpha_{pq}(k)J_{qq}J_{pp}\right]\\&\left.+\sum_{k=1,l=1}^{K,K}\alpha_{pq}(k)\alpha_{pq}(l)\big(|J_{pq}|^2+\phi_{kl}J_{pp}J_{qq}\big)\right].\end{aligned}$$

C. Discussion

For high signal-to-noise ratio, the expansion (27) of the ISR is dominated by the first term $\mathcal{I}_{pq}^0$. Below, some comments on this term are given.

- If the sources p and q have identical t-f signatures over the chosen t-f points (i.e., $\mathbf{d}_p=\mathbf{d}_q$), the corresponding ISR $\mathcal{I}_{pq}\to\infty$. This confirms the statement of Theorem 2.
- As the correlation function r_{Tpq} of the sources p and q and the crossterms $D_{s_ps_q}(t_k,f_k)$ vanish, the corresponding ISR given by $\mathcal{I}_{pq}$ also vanishes, yielding a perfect separation.
- $\mathcal{I}_{pq}^0$ is independent of the mixing matrix. In the array processing context, it means that performance in terms of interference rejection is unaffected by the array geometry and, in particular, by the number of sensors. The performance depends only on the sample size and the t-f signatures of the sources.
- In (27), the sum over k is a sum over time and frequency. Hence, the joint diagonalization can be seen as a kind of averaging. Indeed, the choice of a large number of t-f points increases the performance.

Note that from the above analysis, the choice of the t-f kernel has a direct impact on the performance of the proposed method. Optimal smoothing kernels could, at least theoretically, be obtained by extending the previous derivations. This point is left to further study.

VI. Performance Evaluation

In this section, the performance of the t-f separation (TFS) method, as investigated via computer simulations, is reported. Evaluation of the domain of validity of the asymptotic performance expansion developed in the previous section are also presented.

A. Numerical Experiments

Example 1: This example deals with real source signals.

Two speech signals sampled at 8000 Hz are mixed by the mixing matrix

$$\mathbf{A}=\begin{bmatrix}1.0 & 0.6 & 0.4\\ 0.5 & 1.0 & 0.8\end{bmatrix}^T.$$

The plots of the two individual speech signals are shown in Fig. 1 and their TFD's are displayed in Fig. 2. Speech "1" and "2" of a male speaker are the words "Cars" and "Cats," respectively. The TFD's of the observed speech signals at three sensors are shown in Fig. 3. Fig. 4 shows the TFD's of the speech signals estimated by TFS. It is clear that TFS works

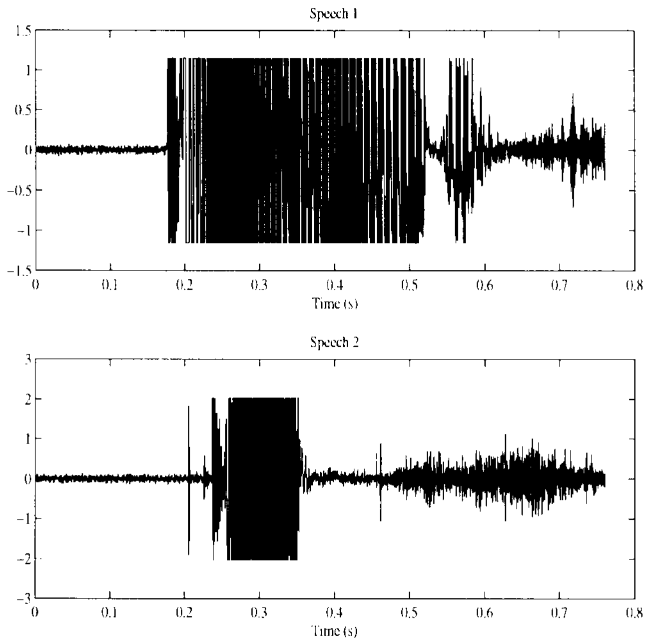

Fig. 1. Plots of individual speeches.

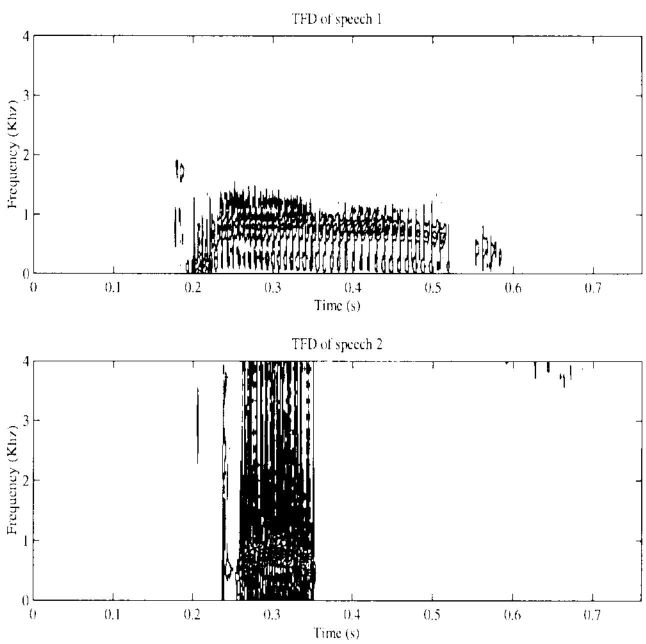

Fig. 2. TFD's of the individual speeches.

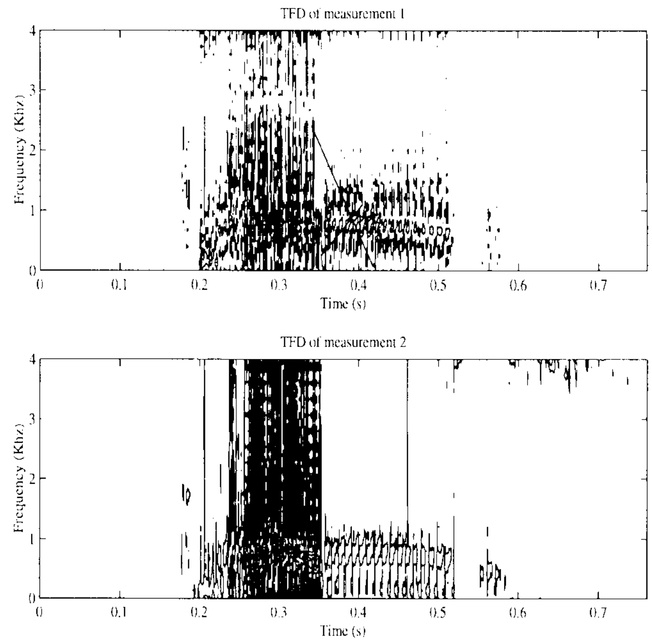

Fig. 3. TFD's of the observed signals.

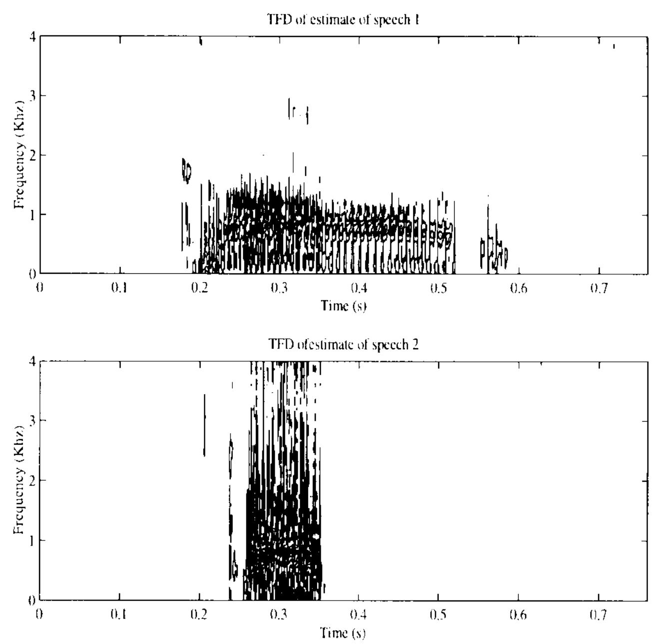

Fig. 4. TFD's of the estimate signals.

well in this case. The purpose of this example is to test the algorithm when speech signals are used. It may not reflect a real speech environment.

To assess the robustness of the TFS algorithm with respect to noise, we corrupt the observed speech signals by an additive white Gaussian noise, and we compare in Fig. 5 the performance of the second-order blind identification (SOBI) algorithm proposed in [12] and the TFS algorithm over the [0–20 dB] range of signal-to-noise ratio (SNR). The mean rejection levels are evaluated here over 100 Monte-Carlo runs with $T = 6084$ samples. It is evident from Fig. 5 that in this case, the TFS algorithm outperforms SOBI algorithm. The increase of this robustness of the TFS algorithm with respect to noise may be explain by the effect of spreading the noise power and of localizing the source energy in the t-f domain.

Example 2: In this example, we consider a uniform linear array of three sensors having half wavelength spacing and receiving signals from two sources in the presence of white Gaussian noise. The sources arrive from different directions $\phi_1 = 0$ and $\phi_2 = 20°$ (the particular structure of the array manifold is, of course, not exploited here). The source signals are generated by filtering a complex circular white Gaussian

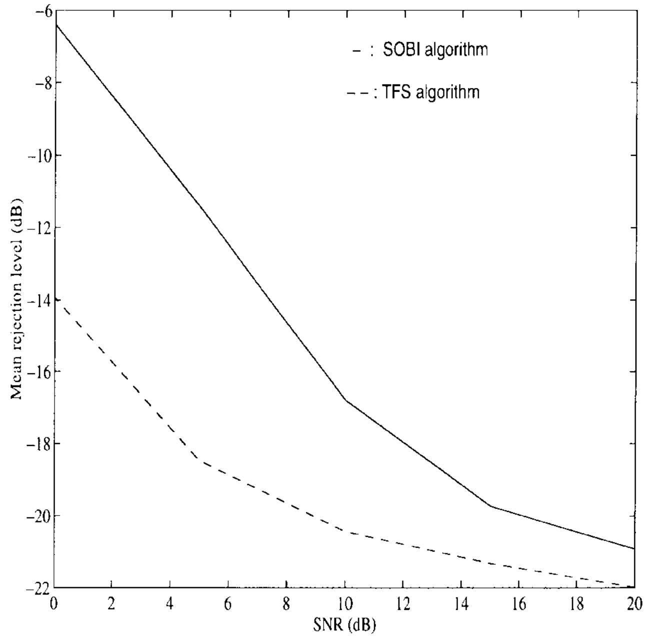

Fig. 5. Performance of SOBI and TFS algorithms versus SNR.

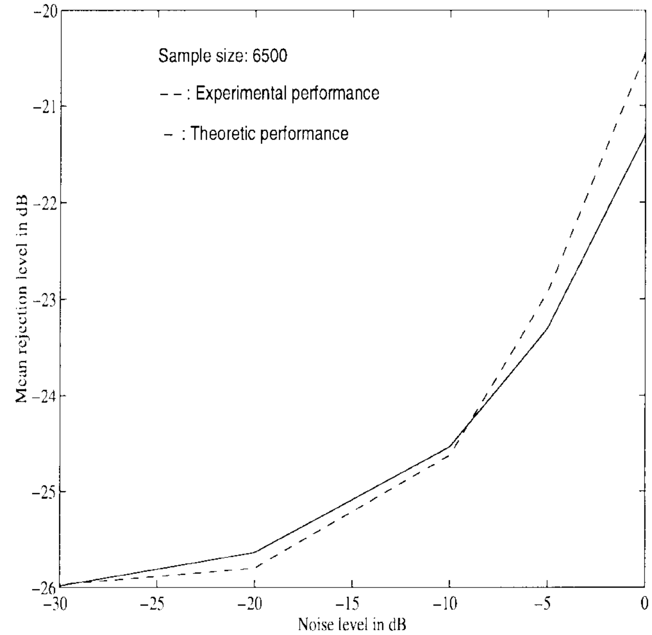

Fig. 6. Performance validation versus σ^2.

TABLE I
PERFORMANCE OF SOBI AND TFS ALGORITHMS VERSUS δf

Spectral shift (δf)	*Mean Rejection level in dB*	
	SOBI	TFS
0.000	-8.86	-12.22
0.002	-10.01	-12.21
0.010	-10.18	-12.34
0.050	-11.09	-12.53
0.200	-12.92	-12.54

processes by an AR model of order one with coefficient $a_1 = \rho \exp(j2\pi f_1(t))$ and $a_2 = \rho \exp(j2\pi f_2(t))$, where we have

$$f_1(t) = \begin{cases} 0.0625, & \text{for } t = 1:400 \\ 0.1250, & \text{for } t = 401:450 \\ 0.3750, & \text{for } t = 451:850 \end{cases}$$

$$f_2(t) = \begin{cases} 0.3750, & \text{for } t = 1:400 \\ 0.1250 + \delta f, & \text{for } t = 401:450 \\ 0.0625, & \text{for } t = 451:850 \end{cases}$$

$$\rho = 0.85.$$

The signal-to-noise ratio (SNR) is set at 5 dB. The kernel used for the computation of the TFD's is the Choi–Williams kernel [16], which provides a good reduction of the crossterms. For the TFS algorithm, eight TFD matrices are considered. The corresponding t-f points are those of the highest power in the t-f domain. The mean rejection level is evaluated over 500 Monte-Carlo runs.

Table I shows the mean rejection level in dB versus the "spectral shift" δf both for the SOBI [12] and TFS algorithms. Note that for $\delta f = 0$, the two *Gaussian* source signals have *identical spectral shape*. In this case, while SOBI fails[1] in separating the two sources, TFS succeeds. Note also that in contrast to SOBI, TFS presents constant performance with respect to δf.

[1] We admit that a source separation algorithm fails when the mean rejection level is greater than -10 dB.

B. Experimental Validation of the Performance Analysis

This section deals with the evaluation of the domain of validity of the first-order performance approximation (27).

The same settings than in Example 2 are used, with the exception of the source signals, which are deterministic sinusoids at frequencies $f_1 = 0.4375$ and $f_2 = 0.0625$. The TFD's are computed using windowed Wigner distribution. The chosen window width is $M = 2L+1$, with $L = 32$. The identification is performed using $\frac{T}{M}$ STFD matrices spaced in time by M samples (T being the sample size). The overall rejection level is evaluated over 500 independent runs.

In Fig. 6, the rejection level $\mathcal{I}_{\text{perf}}$ is plotted in decibels as a function of the noise power σ^2 (also expressed in decibels). In Fig. 7, the rejection level $\mathcal{I}_{\text{perf}}$ is plotted in decibels as against sample size. Both Figs. 6 and 7 show that the approximation is better at high SNR and for large sample size. This means that the asymptotic conditions are reached faster in this range of parameters.

VII. CONCLUSION

In this paper, a new blind separation approach using t-f distributions (STFD's) is introduced. It is devised to primarily separate sources with temporal nonstationary signal characteristics. The new approach is based on the joint diagonalization of a combined set of spatial t-f distribution matrices. The later are made up of the auto- and cross-TFD's of the data snapshots across the multisensor array, and they are expressed in terms of the TFD matrices of the sources. The TFD matrices of the data and sources appear, respectively, in place of the spatial and signal correlation matrices commonly used under stationary

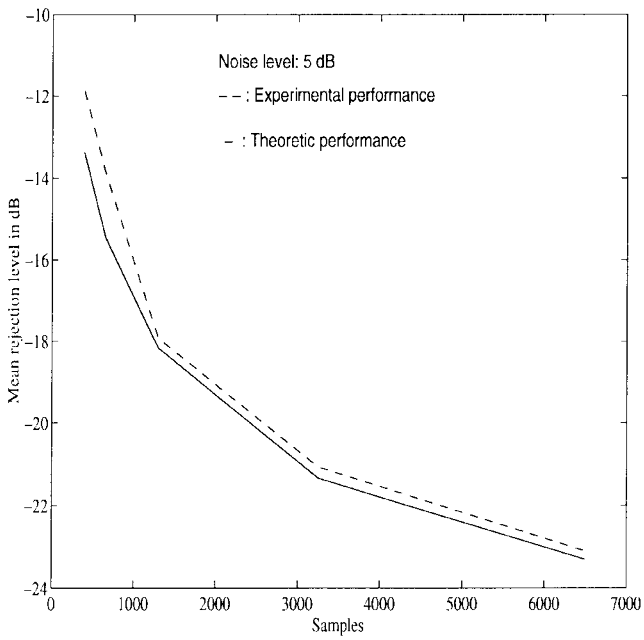

Fig. 7. Performance validation versus samples size (T).

environment. The diagonal structure of the TFD matrix of the sources is essential for the proposed approach and is enforced by incorporating only the T-F points corresponding to the signal autoterms. The off-diagonal elements are crossterms that become negligible by using a reduced-interference distribution kernel. We have focused on the TFD's of Cohen's class, however. We can use any other bilinear t-f distributions signal representations, such as the affine and hyperbolic classes.

The proposed approach shows a number of attractive features. In contrast to blind source separation approaches using second-order and/or high-order statistics, it allows the separation of Gaussian sources with identical spectral shapes but with different t-f localization properties. The effect of spreading the noise power while localizing the source energy in the t-f domain amounts to increasing the robustness of the proposed approach with respect to noise. The paper has included numerical experiments of simple nonstationary signals as well as real source signal data. These experiments have demonstrated the effectiveness of the proposed technique in separating a wide class of signals. The asymptotic performance analysis of the proposed technique has been provided.

Appendix A
Proof for Lemma 1

In this section, a sketch of the proof for Lemma 1 is presented. We follow the same steps as in [12]. The square modulus $|\hat{I}_{pq}|^2$ is expressed as

$$|\hat{I}_{pq}|^2 = |(\hat{\mathbf{U}}^H\hat{\mathbf{W}}\mathbf{A})_{pq}|^2. \tag{28}$$

We decompose the matrix $\hat{\mathbf{W}}\mathbf{A}$ under its polar form

$$\hat{\mathbf{V}}\hat{\mathbf{H}} = \hat{\mathbf{W}}\mathbf{A} \tag{29}$$

where $\hat{\mathbf{V}}$ is a unitary matrix, and $\hat{\mathbf{H}}$ is a non-negative semidefined hermitian matrix; matrix $\hat{\mathbf{H}}$ verifies $\hat{\mathbf{H}}^2 = \mathbf{A}^H\hat{\mathbf{W}}^H\hat{\mathbf{W}}\mathbf{A}$ (see [22, th. 7.3.2, p. 412]). According to the convention outlined in Section C, matrix $\hat{\mathbf{H}}$ is expected to be close to the identity matrix; let $\delta\mathbf{H} = \hat{\mathbf{H}} - \mathbf{I}$ denote the estimation error of the hermitian part of $\hat{\mathbf{W}}\mathbf{A}$. Using standard perturbation calculus (see, for example, [25]), it can be shown that

$$\begin{aligned}\delta\mathbf{H} \simeq & -\frac{1}{2}\mathbf{A}^{\#}\delta\mathbf{R}\mathbf{A}^{\#^H} + \frac{1}{2(m-n)}\mathrm{Tr}(\Pi\delta\mathbf{R})(\mathbf{A}^H\mathbf{A})^{-1} \\ & + o(\delta\mathbf{R}).\end{aligned} \tag{30}$$

From the polar decomposition (29), the whitened STFD matrices can be similarly approximated at the first order, for all $k \neq 0$, as

$$\begin{aligned}&\mathbf{D}_{\mathbf{zz}}(t_k, f_k) \\ &\quad = \hat{\mathbf{W}}(\mathbf{A}\,\mathrm{diag}\big[D_{s_1s_1}(t_k, f_k), \ldots, D_{s_ns_n}(t_k, f_k)\big]\mathbf{A}^H \\ &\qquad + \mathbf{D}_{\mathbf{xx}}(t_k, f_k) - \mathbf{A}\,\mathrm{diag}\big[D_{s_1s_1}(t_k, f_k), \ldots, \\ &\qquad D_{s_ns_n}(t_k, f_k)\big]\mathbf{A}^H)\hat{\mathbf{W}}^H \\ &\quad = \hat{\mathbf{V}}\big(\hat{\mathbf{H}}\,\mathrm{diag}\big[D_{s_1s_1}(t_k, f_k), \ldots, D_{s_ns_n}(t_k, f_k)\big]\hat{\mathbf{H}} \\ &\qquad + \hat{\mathbf{V}}^H\hat{\mathbf{W}}\delta\mathbf{D}_{\mathbf{xx}}(t_k, f_k)\hat{\mathbf{W}}^H\hat{\mathbf{V}}\big)\hat{\mathbf{V}}^H.\end{aligned} \tag{31}$$

The joint diagonalization criterion aims at searching the unitary matrix that minimizes the off-diagonal elements of a set of matrices: here, the whitened STFD matrix $\mathbf{D}_{\mathbf{zz}}(t_k, f_k)$. It can be shown (see a discussion in [26]) that if the set of matrices entering in the JD are multiplied by a *common* unitary matrix, then the result of the JD will simply be multiplied by this common matrix. Formally, let $\mathbf{N}_1, \ldots, \mathbf{N}_p$ be arbitrary matrices, and let $\mathbf{U}$ be an arbitrary unitary matrix; then, $JD\{\mathbf{U}\mathbf{N}_1\mathbf{U}^H, \ldots, \mathbf{U}\mathbf{N}_p\mathbf{U}^H\} = \mathbf{U}JD\{\mathbf{N}_1, \ldots, \mathbf{N}_p\}$. Applying this result in our situation, it comes from (31) that the unitary matrix $\hat{\mathbf{U}}$ resulting from the JD of the set of whitened STFD matrices $\mathbf{D}_{\mathbf{zz}}(t_1, f_1), \ldots, \mathbf{D}_{\mathbf{zz}}(t_K, f_K)$ can be decomposed as

$$\hat{\mathbf{U}} = \hat{\mathbf{V}}\hat{\mathbf{U}}_0$$

where the matrix $\hat{\mathbf{U}}_0$ minimizes the JD criterion for the matrices

$$\begin{aligned}\mathbf{M}_k \stackrel{\text{def}}{=} & \hat{\mathbf{H}}\,\mathrm{diag}\big[D_{s_1s_1}(t_k, f_k), \ldots, D_{s_ns_n}(t_k, f_k)\big]\hat{\mathbf{H}} \\ & + \hat{\mathbf{V}}^H\hat{\mathbf{W}}\delta\mathbf{D}_{\mathbf{xx}}(t_k, f_k)\hat{\mathbf{W}}^H\hat{\mathbf{V}} \\ = & \,\mathrm{diag}\big[D_{s_1s_1}(t_k, f_k), \ldots, D_{s_ns_n}(t_k, f_k)\big] \\ & + \mathrm{diag}\big[D_{s_1s_1}(t_k, f_k), \ldots, D_{s_ns_n}(t_k, f_k)\big]\delta\mathbf{H} \\ & + \delta\mathbf{H}\,\mathrm{diag}\big[D_{s_1s_1}(t_k, f_k), \ldots, D_{s_ns_n}(t_k, f_k)\big] \\ & + \hat{\mathbf{V}}^H\hat{\mathbf{W}}\delta\mathbf{D}_{\mathbf{xx}}(t_k, f_k)\hat{\mathbf{W}}^H\hat{\mathbf{V}} + o(\delta\mathbf{D}_{\mathbf{xx}}(t_k, f_k)) \\ = & \,\mathrm{diag}\big[D_{s_1s_1}(t_k, f_k), \ldots, D_{s_ns_n}(t_k, f_k)\big] \\ & + \mathrm{diag}\big[D_{s_1s_1}(t_k, f_k), \ldots, D_{s_ns_n}(t_k, f_k)\big]\delta\mathbf{H} \\ & + \delta\mathbf{H}\,\mathrm{diag}\big[D_{s_1s_1}(t_k, f_k), \ldots, D_{s_ns_n}(t_k, f_k)\big] \\ & + \mathbf{A}^{\#}\delta\mathbf{D}_{\mathbf{xx}}(t_k, f_k)\mathbf{A}^{\#^H} + o(\delta\mathbf{D}_{\mathbf{xx}}(t_k, f_k)) \\ = & \,\mathrm{diag}\big[D_{s_1s_1}(t_k, f_k), \ldots, D_{s_ns_n}(t_k, f_k)\big] \\ & + \xi_k + o(\delta\mathbf{D}_{\mathbf{xx}}(t_k, f_k)), \quad 1 \leq k \leq K\end{aligned}$$

where

$$\begin{aligned}\xi_k \stackrel{\text{def}}{=} & \,\mathrm{diag}\big[D_{s_1s_1}(t_k, f_k), \ldots, D_{s_ns_n}(t_k, f_k)\big]\delta\mathbf{H} \\ & + \delta\mathbf{H}\,\mathrm{diag}\big[D_{s_1s_1}(t_k, f_k), \ldots, D_{s_ns_n}(t_k, f_k)\big] \\ & + \mathbf{A}^{\#}\delta\mathbf{D}_{\mathbf{xx}}(t_k, f_k)\mathbf{A}^{\#^H}.\end{aligned}$$

Hence, (28) can be written as

$$|\hat{I}_{pq}|^2 = \left|\left(\hat{\mathbf{U}}_0^H \hat{\mathbf{H}}\right)_{pq}\right|^2.$$

As shown in [26], the unitary matrix $\hat{\mathbf{U}}_0$ is given at first order by

$$\hat{\mathbf{U}}_0 = \mathbf{I} + \delta\mathbf{U}_0$$
$$\delta\mathbf{U}_0 = \frac{1}{2}\sum_{r\neq s}\sum_{k=1}^{K}\left(\alpha_{rs}(k)\Pi_r\xi_k\Pi_s + \alpha_{rs}^*(k)\Pi_r{\xi_k}^H\Pi_s\right)$$
$$\left(\delta\mathbf{U}_0^H = -\delta\mathbf{U}_0\right) \quad (32)$$

where $\Pi_r = \mathbf{e}_r\mathbf{e}_r^*$ is the orthogonal projector on the rth vector column $\mathbf{e}_r$ of the identity matrix $\mathbf{I}_n$. The performance index becomes

$$\begin{aligned}|\hat{I}_{pq}|^2 &= |(\mathbf{I} - \delta\mathbf{U}_0)(\mathbf{I} + \delta\mathbf{H})|_{pq}^2 \\ &\simeq |\delta\mathbf{H} - \delta\mathbf{U}_0|_{pq}^2 \quad \text{for } p \neq q. \end{aligned} \quad (33)$$

Including expressions (30) and (32) in (33) leads to the Taylor expansion of Lemma 1.

REFERENCES

[1] C. M. Berrah, "Parameter yield estimation for a MOSFET integrated ciruit," in *Proc. 1990 IEEE ISCAS*, pp. 2260–2263.
[2] E. E. Cureton and R. B. D'Agostino, *FACTOR ANALYSIS an Applied Approach*. New York: Lawrence Erlbaum, 1983.
[3] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propagat.*, vol. AP-34, pp. 276–280, 1986.
[4] G. Demoment, "Image reconstruction and restoration: Overview of common estimation structures and problems," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 2024–2036, Oct. 1989.
[5] C. Jutten and J. Hérault, "Détection de grandeurs primitives dans un message composite par une architecture de calcul neuromimétrique en apprentissage non supervisé," in *Proc. Gretsi*, Nice, France, 1985.
[6] M. Gaeta and J.-L. Lacoume, "Source separation without a priori knowledge: The maximum likelihood solution," in *Proc. EUSIPCO*, 1990, pp. 621–624.
[7] P. Comon, "Independent component analysis, a new concept?" *Signal Process.*, vol. 36, pp. 287–314, 1994.
[8] J.-F. Cardoso and A. Souloumiac, "An efficient technique for blind separation of complex sources," in *Proc. IEEE SP Workshop Higher Order Statist.*, Lake Tahoe, CA, 1993.
[9] E. Moreau and O. Macchi, "New self-adaptive algorithms for source separation based on contrast functions," in *Proc. IEEE SP Workshop Higher Order Statist.*, Lake Tahoe, CA, 1993.
[10] A. Belouchrani and J.-F. Cardoso, "Maximum likelihood source separation for discrete sources," in *Proc. EUSIPCO*, 1994, pp. 768–771.
[11] L. Tong and R. Liu, "Blind estimation of correlated source signals," in *Proc. Asilomar Conf.*, Nov. 1990.
[12] A. Belouchrani, K. A. Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique using second order statistics," *IEEE Trans. Signal Processing*, vol. 45, pp. 434–444, Feb. 1997.
[13] M. Wax and J. Sheinvald, "A least-squares approach to joint diagonalization," *IEEE Signal Processing Lett.*, vol. 4, pp. 52–53, Feb. 1997.
[14] K. Matsuoka, M. Ohya, and M. Kawamoto, "A neural net for blind separation of nonstationary signals," *Neural Networks*, vol. 8, pp. 411–419, 1995.
[15] F. Hlawatsch and W. Krattenthaler, "Blinear signal synthesis," *IEEE Trans. Signal Processing*, vol. 40, pp. 352–363, Feb. 1992.
[16] L. Cohen, *Time-Frequency Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
[17] J. Jeong and W. Williams, "Kernel design for reduced interference distributions," *IEEE Trans. Signal Processing*, vol. 40, pp. 402–412, Feb. 1992.
[18] R. Baraniuk and D. Jones, "A signal dependent time-frequency representation: Optimum kernel design," *IEEE Trans. Signal Processing*, vol. 41, pp. 1589–1603, Apr. 1993.
[19] B. Ristic and B. Boashash, "Kernel design for time-frequency signal analysis using the Radon transform," *IEEE Trans. Signal Processing*, vol. 41, pp. 1996–2008, May 1993.
[20] M. Amin, G. Venkatesan, and J. Carroll, "A constrained weighted least square approach for time-frequency kernel design," *IEEE Trans. Signal Processing*, vol. 44, pp. 1111–1123, May 1996.
[21] J.-F. Cardoso and A. Souloumiac, "Blind beamforming for non Gaussian signals," *Proc. Inst. Elect. Eng.*, vol. 140, no. 6, pp. 362–370, 1993.
[22] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.
[23] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: Johns Hopkins Univ. Press, 1989.
[24] M. Rosenblatt, *Stationary Sequences and Random Fields*. Berlin, Germany: Birkhauser-Verlag, 1985.
[25] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "A second order blind source separation technique: Implementation and performance," Tech. Rep. 94D027, Telecom Paris, Signal Dept., 1994.
[26] J.-F. Cardoso, "Perturbation of joint diagonalizers," Tech. Rep. 94D023, Telecom Paris, Signal Dept., 1994.

Adel Belouchrani received the State Engineering degree in 1991 from the National Polythechic School of Algiers, Algiers, Algeria, the M.Sc. degree in signal processing from the Institut National Polytechnique de Grenoble (INPG), Grenoble, France, in 1992, and the Ph.D. degree in signal and image processing from Ecole Nationale Superieure des Telecommunications (ENST), Paris, France, in 1995.

He was a Visiting Scholar at the Electrical Engineering and Computer Science Department, University of California, Berkeley, from 1995 to 1996, working on fast adaptive blind equalization and carrier phase tracking. He was with the Department of Electrical and Computer Engineering, Villanova University, Villanova, PA, as a Research Associate from 1996 to 1997. He also served as a consultant to Comcast Inc., Philadelphia, PA, during this period. In February 1997, he was a visiting scientist at the Laboratory for Artificial Brain Systems, Riken, Japan. From August 1997 to October 1997, he was with Alcatel ETCA, Belgium, working on the Very High Speed Digital Subscriber Line (VDSL). He is currently with the Electrical Engineering Department, National Polytechnic School of Algiers, as an Assistant Professor. His research interests are in the areas of digital communications and statistical signal processing including (blind) array processing, performance analysis, blind source separation, blind equalization, systems identification, damped sinusoids estimation, adaptive algorithms, expectation-maximization techniques applied to communications, nonstationary signals, and spread spectrum communications.

Moeness G. Amin (SM'91) received the B.Sc. degree in 1976 from Cairo University, Cairo, Egypt, the M.Sc. degree in 1980 from University of Petroleum and Minerals, Dharan, Saudi Arabia, and the Ph.D. degree in 1984 from the University of Colorado, Boulder, all in electrical engineering.

In 1984, he joined University of Colorado, Denver, as a Visiting Assistant Professor. He has been on the Faculty of the Department of Electrical and Computer Engineering at Villanova University, Villanova, PA, since 1985, where is now a Professor. His current research interests are in the areas of time–frequency analysis, spread spectrum communications, smart antennas, and blind signal processing.

Dr. Amin is currently an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and a member of the Technical Committee of the IEEE Signal Processing Society on Signal Processing for Communications. He was the General Chair of the 1994 IEEE International Symposium on Time–Frequency and Time-Scale Analysis and is the General Chair of the 2000 IEEE Workshop on Statistical Signal and Array Processing, Poconos, PA. He is the recipient of the 1997 IEEE Philadelphia Section Award for "Outstanding Intellectual and Organizational Contributions to the IEEE Philadelphia Section in the Area of Signal Processing." He is also the recipient of the 1997 Villanova University Outstanding Faculty Research Award. Over the past four years, he chaired three All-Day Workshops on Smart Antennas, Recent Trends in Adaptive Filtering, and Advanced Signal Processing Applications in Wireless Communications. He is a member of Sigma Xi, Eta Kappa Nu, and Phi Kappa Phi.