SINGLE AUTOTERMS SELECTION FOR BLIND SOURCE SEPARATION IN TIME-FREQUENCY PLANE

Aleš Holobar [1], *Cédric Févotte* [2], *Christian Doncarli* [2], *Damjan Zazula*[1]
[1] Faculty of electrical engineering and computer science, 2000 Maribor, Slovenia
[2] IRCCyN, B.P 92101, 44321 Nantes cedex 03, France
e-mail: ales.holobar@uni-mb.si, cedric.fevotte@irccyn.ec-nantes.fr

ABSTRACT

We address in this paper a method to perform blind separation of over-determined instantaneous linear mixtures of non-stationary sources. Our contribution is focused on the identification of *single source autoterms* which correspond to diagonal Spatial Time-Frequency Distribution of the sources vector with only one non-zero diagonal entry. These latter matrices are processed into an iterative joint-diagonalization scheme which provides an estimation of the mixing matrix.

1 INTRODUCTION

The problem of blind source separation (BSS) in its simplest form (identify n mutually independent unknown sources from m linear and instantaneous mixtures - the observations - without any *a priori* knowledge on the sources nor on the mixture) has been widely studied and many solutions exist, but generally under additional hypotheses. Actually a favorable case consists in assuming the sources to be independently and identically distributed (i.i.d) and to have at least as many observations as sources. Numerous algorithms have been designed from these hypotheses (see [1] for a review of them).

Other works have been carried out for cases where one of the 'i' of the i.i.d hypothesis fails. The SOBI algorithm [2] takes advantage of the eventual temporal self-correlation of each source, that is to say when each source sequence is not independently distributed. Other works deal with the non-stationary case. The definition of "non-stationarity" through papers is not very unique. Non-stationarity can be settled on a statistic point of view: each source is no more identically distributed in time and its statistical properties are time/lag dependent; in this case independence has the standard statistical meaning (see [3] for source separation in this context). Besides, non-stationarity can be settled from a rather more deterministic edge: the frequency content of each realization of the sources is time-varying; in that case independence means that the sources energy location in time-frequency plane shall not completely overlap. From this point of view, A. Belouchrani and M. Amin have proposed a blind source separation method based on the joint-diagonalization of a set of *spatial t-f distribution matrices* corresponding to a set of source autoterms points [4]. But no procedure was given to blindly select the needed autoterms, and this is the aim of this paper.

Section 2 briefly exposes the SOBI algorithm and its extension to the joint-diagonalization of spatial t-f distribution matrices instead of correlation matrices. In section 3 we propose a method for the selection of single source autoterms; section 4 describes an iterative joint-diagonalization process which eliminates perturbing autoterms. Section 5 provides results on synthetic signals, conclusions and perspectives are discussed in section 6.

2 STATE OF THE ART

We consider the following model:

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

where:

- $\mathbf{x}(t) = [x_1(t), \ldots, x_m(t)]^T$ is a m-vector of observations,
- $\mathbf{s}(t) = [s_1(t), \ldots, s_n(t)]^T$ is a n-vector containing the zero-mean unknown sources; they are mutually independent for every time/lag,
- $\mathbf{A}$ is a $m \times n$ unknown full rank mixing matrix, with $m > n$ (over-determined case),
- $\mathbf{n}(t)$ is a i.i.d noise vector independent from the sources defined by

$$\mathrm{E}\left\{\mathbf{n}(t+\tau)\,\mathbf{n}(t)\right\} = \delta(\tau)\,\sigma^2\mathbf{I}_m$$

where $\mathbf{I}_m$ is the $m \times m$ identity matrix, $\delta(\tau)$ the Dirac impulse and σ^2 the unknown variance of the noise, identical on each observation.

The aim of blind source separation is to recover the sources from the only observations. Without any a priori knowledge on the sources or on the mixture, one can only recover the sources up to permutations, sign change, constant factor (scale factor and phase shift in the complex case), since these modifications can be balanced by the mixing matrix to provide exactly the same observations. This is the well known blind source separation *indeterminacy* property.

2.1 Stationary Ergodic Sources

We briefly recall here the SOBI method and the joint-diagonalization concept, see [2] for details. We note:

$$\mathbf{R}_{\mathbf{ss}}(\tau) = \mathrm{E}\left\{\mathbf{s}(t+\tau)\,\mathbf{s}^*(t)\right\} = \mathrm{diag}[\rho_1(\tau)\ldots\rho_n(\tau)] \tag{2}$$

with $\rho_i(\tau) = \mathrm{E}\left\{s_i(t+\tau)\,s_i{}^*(t)\right\}$. The indeterminacy property allows us to set, *without any loss of generality*:

$$\mathbf{R}_{\mathbf{ss}}(0) = \mathbf{I}_n \tag{3}$$

Equation (1) leads to:

$$\mathbf{R}_{\mathbf{xx}}(\tau) = \mathrm{E}\left\{\mathbf{x}(t+\tau)\,\mathbf{x}^*(t)\right\} = \mathbf{A}\mathbf{R}_{\mathbf{ss}}(\tau)\mathbf{A}^H + \delta(\tau)\,\sigma^2\mathbf{I}_m \tag{4}$$

In practice, the autocorrelation matrices $\mathbf{R}_{\mathbf{xx}}(\tau)$ of the observations are estimated from the classical ergodic formula.

2.1.1 Step 1 - Whitening

We look for a $n \times m$ matrix $\mathbf{W}$ such that $\mathbf{W}\mathbf{A}\mathbf{A}^H\mathbf{W}^H = \mathbf{I}_n$ ($\mathbf{W}$ is a *whitening matrix*). In discrete-time, from (3) and (4) we have:

$$\mathbf{R}_{\mathbf{xx}}[0] = \mathbf{A}\mathbf{A}^H + \sigma^2\mathbf{I}_m \tag{5}$$

We note $[\lambda_1 \dots \lambda_m]$ the eigenvalues of $\mathbf{R_{xx}}[0]$, sorted in decreasing order. We note $[\mathbf{h}_1 \dots \mathbf{h}_m]$ the corresponding eigenvectors. $\mathbf{A}$ being a $m \times n$ full rank matrix, $\mathbf{A}\mathbf{A}^H$ has m-n zero eigenvalues. Then, from equation (5), we have $\sigma^2 = \lambda_{n+1} = \dots = \lambda_m$. Thus σ^2 can be estimated by the mean value $\hat{\sigma}^2$ of the m-n smallest eigenvalues of $\mathbf{R_{xx}}[0]$. Moreover an estimation of $\mathbf{W}$ is:

$$\hat{\mathbf{W}} = [(\lambda_1 - \hat{\sigma}^2)^{-1/2}\ \mathbf{h}_1, \dots, (\lambda_n - \hat{\sigma}^2)^{-1/2}\ \mathbf{h}_n] \qquad (6)$$

2.1.2 Step 2 - Joint-Diagonalization

We note $\mathbf{U} = \mathbf{WA}$, from the definition of $\mathbf{W}$, $\mathbf{U}$ is unitary. If we now compute

$$\underline{\mathbf{R}}_{\mathbf{xx}}(\tau) = \mathbf{W}\mathbf{R}_{\mathbf{xx}}(\tau)\mathbf{W}^H \qquad (7)$$

we have from equation (4), for $\tau \neq 0$:

$$\begin{aligned} \underline{\mathbf{R}}_{\mathbf{xx}}(\tau) &= (\mathbf{WA})\ \mathbf{R_{ss}}(\tau)\ (\mathbf{WA})^H \qquad (8) \\ &= \mathbf{U}\ \mathrm{diag}[\rho_1(\tau) \dots \rho_n(\tau)]\ \mathbf{U}^H \qquad (9) \end{aligned}$$

Thus, we see that for any $\tau \neq 0$, $\mathbf{U}$ is a diagonalizing matrix of $\underline{\mathbf{R}}_{\mathbf{xx}}(\tau)$. Then, $\mathbf{U}$ could be estimated from the eigenvectors of $\underline{\mathbf{R}}_{\mathbf{xx}}(\tau)$, for any $\tau \neq 0$. In practice, in order to balance the errors made on the estimation of $\underline{\mathbf{R}}_{\mathbf{xx}}(\tau)$ (due to the finite length signals and to the use of an estimation of $\mathbf{W}$) and then to get a more robust estimation of $\mathbf{U}$, the idea of SOBI is to *joint-diagonalize* a set of K matrices $\{\underline{\mathbf{R}}_{\mathbf{xx}}(\tau), \tau = \tau_1 \dots \tau_K\}$. This provides us with an estimation of $\hat{\mathbf{U}} = \hat{\mathbf{W}}\hat{\mathbf{A}}$ which enables us to compute $\hat{\mathbf{A}}$ and then the sources (see [5] for a joint-diagonalization procedure). We have the following uniqueness property of joint-diagonalization [2]:

Property *Let $\mathcal{M}$=$\{\mathbf{M}_1, \dots, \mathbf{M}_K\}$ be a set of K matrices where, for $1 \leq k \leq K$, matrix $\mathbf{M}_k$ is in the form $\mathbf{M}_k = \mathbf{U}\mathbf{D}_k\mathbf{U}^H$ with $\mathbf{U}$ a unitary matrix and $\mathbf{D}_k = diag[d_1(k), \dots, d_n(k)]$. Then any unitary joint-diagonalizer of $\mathcal{M}$ is equal to $\mathbf{U}$ (up to column permutations and phase shifts) if and only if*

$$\forall \quad 1 \leq i \neq j \leq n, \quad \exists\ k,\ 1 \leq k \leq K \quad / \quad d_i(k) \neq d_j(k)$$

This theorem is very important and the selection of $\{\tau_1 \dots \tau_K\}$ strongly depends on it. It will further guarantee the exactness of our solution provided by the joint-diagonalization of a set of spatial t-f distribution matrices corresponding to the so called "single autoterms" *from every source.*

2.2 Non-Stationary Sources

Now we assume that the sources are non-stationary, which means that their statistical properties are time and lag dependent. The covariance of vectors $\mathbf{s}(t)$ and $\mathbf{x}(t)$ are then:

$$\begin{aligned} \mathbf{R_{ss}}(t,\tau) &= \mathrm{E}\{\mathbf{s}(t+\tau)\,\mathbf{s}^*(t)\} = \mathrm{diag}[\rho_1(t,\tau), \dots, \rho_n(t,\tau)] \\ \mathbf{R_{xx}}(t,\tau) &= \mathbf{A}\ \mathrm{diag}[\rho_1(t,\tau), \dots, \rho_n(t,\tau)]\ \mathbf{A}^H + \delta(\tau)\,\sigma^2\mathbf{I}_m \end{aligned} \qquad (10)$$

Provided we can find like previously a $m \times n$ matrix $\mathbf{W}$ such as $\mathbf{WA}(\mathbf{WA})^H = \mathbf{I}_n$, $\mathbf{U} = \mathbf{WA}$ still diagonalizes $\mathbf{W}\mathbf{R_{xx}}(t,\tau)\mathbf{W}^H$ for every $(t, \tau \neq 0)$. Equivalently, in the time-frequency plane, equation (10) becomes:

$$\mathcal{SPWV}_{\mathbf{xx}}(t,f) = \mathbf{A}\ \mathcal{SPWV}_{\mathbf{ss}}(t,f)\ \mathbf{A}^H + \sigma^2\,\mathbf{I}_m \qquad (11)$$

where $\mathcal{SPWV}_{\mathbf{xx}}(t,f)$ is the Wigner-Ville Spectra defined by $\mathcal{SPWV}_{\mathbf{xx}}(t,f) = \int_{-\infty}^{+\infty} \mathbf{R_{xx}}(t,\tau)e^{-j2\pi f\tau}\,d\tau$. Considering a low noise environment we assume that the noise term in (11) can be neglected such that $\mathcal{SPWV}_{\mathbf{xx}}(t,f) \approx \mathbf{A}\ \mathcal{SPWV}_{\mathbf{ss}}(t,f)\ \mathbf{A}^H$. Thus, since $\mathcal{SPWV}_{\mathbf{ss}}(t,f)$ is still diagonal, $\mathbf{U}$ approximately diagonalizes $\mathbf{W}\mathcal{SPWV}_{\mathbf{xx}}(t,f)\mathbf{W}^H$ for any (t,f).

However, in practice, we can't estimate $\mathbf{R_{xx}}(t,\tau)$ from the classic ergodic formula like previously. Ideally, the knowledge of a set of different realizations of the observation signals is required to make a spatial average instead of of a temporal average. Though, if the signals are "locally" stationary and ergodic, one can make an estimation of $\mathbf{R_{xx}}(t,\tau)$ from only one realization of $\mathbf{x}(t)$ with a smooth moving window $\phi(t,\tau)$:

$$\mathbf{R_{xx}}(t,\tau) \approx \mathbf{R}^{\phi}_{\mathbf{xx}}(t,\tau) \stackrel{def}{=} \int_{-\infty}^{+\infty} \phi(t-u,\tau)\ \mathbf{x}(u+\tau)\mathbf{x}^*(u)\ du \qquad (12)$$

Equivalently, $\mathcal{SPWV}_{\mathbf{xx}}(t,f)$ is approximated by:

$$\mathcal{SPWV}_{\mathbf{xx}}(t,f) \approx \mathbf{D}^{\phi}_{\mathbf{xx}}(t,f) \stackrel{def}{=} \int_{-\infty}^{+\infty} \mathbf{R}^{\phi}_{\mathbf{xx}}(t,\tau)\ e^{-j2\pi f\tau}\ d\tau \qquad (13)$$

$\mathbf{D}^{\phi}_{\mathbf{xx}}(t,f)$ is an element of the Cohen class of Spatial Time-Frequency Distributions (STFD) of $\mathbf{x}(t)$ for the particular kernel $\phi(t,\tau)$.

2.2.1 Step 1 - Whitening

Since the sources are not stationary, we can't anymore take advantage of the BSS indeterminacy on scale by setting $\mathbf{R_{ss}}(t,0) = \mathbf{I}_n$ and compute $\mathbf{W}$ like previously. However, we can normalize the time average of $\mathbf{R_{ss}}(t,0)$. In the discrete-time case it means we set:

$$\lim_{N\to+\infty} \frac{1}{2N} \sum_{t=-N}^{N} \mathbf{R_{ss}}[t,0] = \mathbf{I}_n \qquad (14)$$

From (10) we have $\mathbf{R_{xx}}[t,0] = \mathbf{A}\mathbf{R_{ss}}[t,0]\mathbf{A}^H + \sigma^2\mathbf{I}_m$. Thus:

$$\lim_{N\to+\infty} \frac{1}{2N} \sum_{t=-N}^{N} \mathbf{R_{xx}}[t,0] = \mathbf{A}\mathbf{A}^H + \sigma^2\mathbf{I}_m \qquad (15)$$

So in practice, for finite signals of length N, we compute the eigen-elements of $\frac{1}{N}\sum_{t=1}^{N} \mathbf{R}^{\phi}_{\mathbf{xx}}[t,0] \approx \mathbf{A}\mathbf{A}^H + \sigma^2\mathbf{I}_m$ and estimate $\mathbf{W}$ like in section 2.1.1.

2.2.2 Step 2 - Joint-Diagonalization

We propose as in [4] to joint-diagonalize a set of whitened matrices $\underline{\mathbf{D}}^{\phi}_{\mathbf{xx}}(t,f) = \mathbf{W}\mathbf{D}^{\phi}_{\mathbf{xx}}(t,f)\mathbf{W}^H$ corresponding to (t,f) points for which $\mathbf{D}^{\phi}_{\mathbf{ss}}(t,f)$ is diagonal (instead of $\mathcal{SPWV}_{\mathbf{ss}}(t,f)$). Now the question is, in front of real situations, when is $\mathbf{D}^{\phi}_{\mathbf{ss}}(t,f)$ diagonal ?

3 ITERATIVE SELECTION OF SINGLE AUTOTERMS

The matrix $\mathbf{D}^{\phi}_{\mathbf{ss}}(t,f)$ is diagonal for (t,f) points such that the cross-STFD terms are very low (ideally zero) and for which at least one diagonal entry is non-zero. In the sequel such (t,f) points will be referred to as autoterms positions, while the (t,f) points with non-zero cross-STFD terms, and very low (ideally zero) diagonal entries in $\mathbf{D}^{\phi}_{\mathbf{ss}}(t,f)$ will be referred to as crossterms positions. The success of the BSS algorithm in determining the unitary matrix $\mathbf{U}$ will in practice strongly depend on the correct selection of autoterms. Therefore it is crucial to have a selection criterion that is able to distinguish between autoterms and crossterms points based only on the whitened STFD matrices of observations

$\underline{\mathbf{D}}^{\phi}_{\mathbf{xx}}(t,f)$. Since $\mathbf{U}$ is a unitary matrix, the following relation is valid for observed STFD matrices corresponding to the crossterms positions:

$$\begin{aligned} \text{trace}(\underline{\mathbf{D}}^{\phi}_{\mathbf{xx}}(t,f)) &= \text{trace}(\mathbf{U}\mathbf{D}^{\phi}_{\mathbf{ss}}(t,f)\mathbf{U}^H) \\ &= \text{trace}(\mathbf{D}^{\phi}_{\mathbf{ss}}(t,f)) \approx 0. \end{aligned} \quad (16)$$

If two sources share a same particular frequency at the very same time, that frequency will very likely appear with non-zero value in the cross-TFD of those sources [6]. Hence, the corresponding source STFD matrix will not be diagonal anymore. That leads us to look only for *single autoterms* positions, which means (t,f) points such that $\mathbf{D}^{\phi}_{\mathbf{ss}}(t,f)$ is diagonal *with only one non-zero diagonal entry*. In such a case the matrix $\mathbf{D}^{\phi}_{\mathbf{ss}}(t,f)$ has only one non-zero eigenvalue. The following relation holds:

$$\text{eig}(\underline{\mathbf{D}}^{\phi}_{\mathbf{xx}}(t,f)) = \text{eig}(\mathbf{U}\mathbf{D}^{\phi}_{\mathbf{ss}}(t,f)\mathbf{U}^H) = \text{eig}(\mathbf{D}^{\phi}_{\mathbf{ss}}(t,f)) \quad (17)$$

where $\text{eig}(\mathbf{M})$ denotes the eigenvalues of the matrix $\mathbf{M}$. In reality the source STFD matrices $\mathbf{D}^{\phi}_{\mathbf{ss}}(t,f)$ are not exactly diagonal (since $\mathbf{D}^{\phi}_{\mathbf{ss}}(t,f)$ is not exactly the Wigner-Ville Spectra of $\mathbf{s}(t)$) . Hence, the "should-be zero" eigenvalues of single autoterms STFD matrices are not exactly equal to zero. Therefore,we propose the following criterion:

$$\textbf{if } \left(\frac{\max(|\ \text{eig}(\underline{\mathbf{D}}^{\phi}_{\mathbf{xx}}(t,f))\ |)}{\sum |\ \text{eig}(\underline{\mathbf{D}}^{\phi}_{\mathbf{xx}}(t,f))\ |} \right) > \varepsilon \quad (18)$$
$$\textbf{then } (t,f) \text{ is a single autoterm position}$$

The threshold value ε is bounded to the interval $[0,1]$. The closer to 1, the better candidates for joint diagonalization will be chosen. However, because of the noise and the estimation errors, setting ε close to 1 can result in ignoring the autoterms of particular source signals. Although the autoterms of one (but only one) particular source signals can be ignored without significant decrease in the BSS performance, one should try his best to find the autoterms of as many different source signals as possible. Ignoring the autoterms of two or more sources introduces indeterminacy in the process of joint-diagonalization, because the necessary condition of the joint-diagonalization uniqueness property is not fulfilled anymore. Therefore, we propose the following procedure:

Algorithm 1 - Selection of single autoterms

1. Form the matrix of traces
$$\mathbf{TrM}(t,f) = \text{trace}(\underline{\mathbf{D}}^{\phi}_{\mathbf{xx}}(t,f)) \quad (19)$$
2. Form the matrix of the eigenvalue criterion
$$\mathbf{EigM}(t,f) = \frac{\max(|\ \text{eig}(\underline{\mathbf{D}}^{\phi}_{\mathbf{xx}}(t,f))\ |)}{\sum |\ \text{eig}(\underline{\mathbf{D}}^{\phi}_{\mathbf{xx}}(t,f))\ |} \quad (20)$$
3. Form the joint criteria matrix
$$\mathbf{CrM}(t,f) = BiOp[\mathbf{TrM}(t,f) > \varepsilon] \cdot \mathbf{EigM}(t,f) \quad (21)$$
where $BiOp[true] = 1$, $BiOp[false] = 0$ and ε is a certain trace threshold value.
4. For different decreasing levels l_i (see below) select (t,f) points belonging to the neighborhood of the extrema of $\mathbf{CrM}(t,f)$ greater than l_i.

The level l_i should decrease exponentially in dependence of the number of selected (t,f) points (in small steps at the beginning and in larger steps at the end) ensuring the majority of selected (t,f) points to belong to single autoterms, e.g:

$$l_i = e^{-\alpha \frac{k_i}{K}} \quad (22)$$

where k_i is the total number of gathered (t,f) points at iteration i, α is a slope coefficient and K is the final number of (t,f) points we want to gather. The superpositions of autoterms (at least two sources present at the same particular (t,f) point) and crossterms that pass the selection criteria (due to low level l in the final iterations of step 4 in algorithm 1) are eliminated by the iterative joint-diagonalization process, explained in next section.

4 ITERATIVE JOINT DIAGONALIZATON

Selected (t,f) points determine a set of K observation STFD matrices $\underline{\mathbf{D}}^{\phi}_{\mathbf{xx}}(t_k,f_k)$ that enter the procedure of joint diagonalization. However, because of the noise, of the autoterms selection errors and of the approximation of the Wigner-Ville Spectra, a single unitary matrix $\mathbf{U}$ cannot be an exact joint-diagonalizer of all the observations STFD matrices. Sources STFD matrices $\mathbf{D}^{\phi}_{\mathbf{ss}}(t_k,f_k)$, that can be calculated by pre- and post-multiplying the corresponding observation STFD matrices $\underline{\mathbf{D}}^{\phi}_{\mathbf{xx}}(t_k,f_k)$ with $\mathbf{U}$, are only quasi-diagonal with their anti-diagonal elements as close to zero as possible.

Assuming that algorithm 1 ensures the selected (t,f) points to be mainly single autoterms positions, we can make the following assertion: joint-diagonalization will more successfully diagonalize the single autoterms matrices than matrices corresponding to superposition of autoterms or even to crossterms. Hence, a simple intuitive idea that significantly improves the performance of the BSS algorithm, is to eliminate the "least diagonalizable" $\underline{\mathbf{D}}^{\phi}_{\mathbf{xx}}(t_k,f_k)$ matrices from the process of joint-diagonalization. We propose the following criterion:

$$\forall k = 1...K \qquad \textbf{if } \left(\frac{(N-1)\sum_{i=1}^{N} |\ \mathbf{D}^{'\phi}_{s_i s_i}(t_k,f_k)\ |}{\sum_{1\le i\ne j\le N} |\ \mathbf{D}^{'\phi}_{s_i s_j}(t_k,f_k)\ |} \right) < \gamma$$
$$\textbf{then remove } \underline{\mathbf{D}}^{\phi}_{\mathbf{xx}}(t_k,f_k) \quad (23)$$

where $\mathbf{D}^{'\phi}_{\mathbf{ss}}(t_k,f_k) = \mathbf{U}^H \underline{\mathbf{D}}^{\phi}_{\mathbf{xx}}(t_k,f_k)\mathbf{U}$, $\mathbf{D}^{'\phi}_{s_i s_j}(t_k,f_k)$ is the (i,j)-th element of $\mathbf{D}^{'\phi}_{\mathbf{ss}}(t_k,f_k)$, and $\mathbf{U}$ is the joint-diagonalizer of the set of $\underline{\mathbf{D}}^{\phi}_{\mathbf{xx}}(t_k,f_k)$ matrices. The criterion (23) can be successively applied (eliminating all non-diagonal $\underline{\mathbf{D}}^{'\phi}_{\mathbf{ss}}(t_k,f_k)$ matrices and recalculating the $\mathbf{U}$ matrix at each iteration) until there are no matrices left for elimination.

The parameter γ is usually set to 1. In other words, iterative joint-diagonalization eliminates the observation STFD matrices $\underline{\mathbf{D}}^{\phi}_{\mathbf{xx}}(t_k,f_k)$ that correspond to $\mathbf{D}^{'\phi}_{\mathbf{ss}}(t_k,f_k)$ matrices with the average absolute diagonal value lower than the average absolute non-diagonal value.

5 RESULTS WITH SYNTHETIC SIGNALS

In this section, the performance of our BBS algorithm, as investigated via computer simulations, is reported. The following three signals were chosen as source signals:

$$s_1(t) = e^{-j2\pi(0.00024414t^2+0.05t)} \quad (24)$$
$$s_2(t) = e^{-j2\pi(4.13\sin(0.0154\pi t)+0.25t)} \quad (25)$$
$$s_3(t) = e^{-j2\pi(0.0000017872t^3-0.0014t^2+0.4027t)} \quad (26)$$

Four observation signals were gathered setting the mixing matrix $\mathbf{A}$ to:

$$\mathbf{A} = \begin{bmatrix} 1.0 & 0.6 & 0.4 & 0.4 \\ 0.5 & 1.0 & 0.8 & 0.6 \\ 0.3 & 0.5 & 1.0 & 0.3 \end{bmatrix}^T \quad (27)$$

A complex white gaussian noise was added to the signals. The Bessel's kernel (a kernel based on the Bessel function of

the first kind) was used for the computation of the STFD's, since it provides a good reduction of the interferences within a single source signal that come from interactions of particular frequency in different times, or interactions of different frequencies in particular time [6]. The trace $\mathbf{TrM}(t,f)$ of the whitened STFD matrices $\underline{\mathbf{D}}^{\phi}_{\mathbf{xx}}(t,f)$ is depicted in Fig.1. Notice the concentrations of power in the proximity of superpositions of sources.

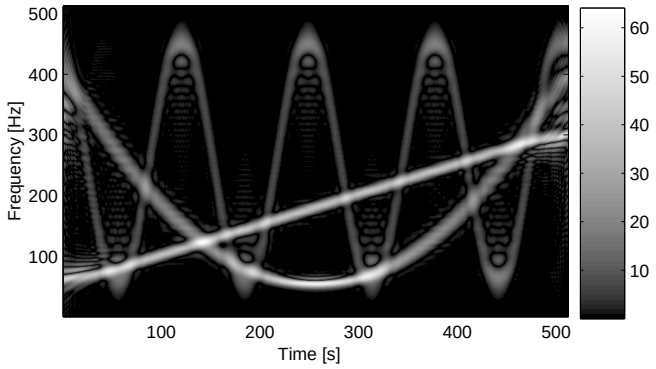

Figure 1: Trace of the whitened STFD matrices $\underline{\mathbf{D}}^{\phi}_{\mathbf{xx}}(t,f)$

5.1 Single vs. superposition of autoterms

A set of 20 (t,f) points $(A1)$ that correspond to autoterms of a single source signal $s_i; i = 1,2,3$, and a set of 20 (t,f) points $(A2)$ that correspond to a superposition of autoterms of two source signals were chosen manually. Different combinations of elements from $A1$ and $A2$ were used for the determination of whitened observation STFD matrices to be joint-diagonalized, always picking 20 (t,f) points (m from $A1$ and n from $A2$, with $m+n=20$). To assure the independence of a particular (t,f) point, 1000 different combinations of (t,f) points were tested for each fixed (m,n) pair. The means of the correlation coefficients between the original and estimated sources were computed over 100 runs of the simulation with SNR set to 10 dB. As shown in Fig. 2, (t,f) points that correspond to a superposition of autoterms of two or more sources severely alter the unitary matrix $\mathbf{U}$ and, therefore, decrease the accuracy of the source separation.

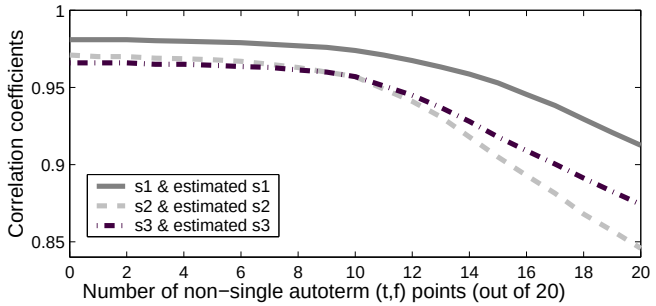

Figure 2: Correlation coefficients between the original and the estimated sources in dependence of the number of (t,f) points that correspond to a superposition of chosen autoterms

5.2 Contribution of the proposed selection

In [4] the suggested (t,f) points for joint-diagonalization were those of the highest power in the t-f domain. We compared this selection criterion with ours. Both algorithms were tested over the [8-40 dB] range of SNR. The means of the correlation coefficients between the original and estimated signals were evaluated over 100 runs of the algorithms. It is evident from Fig. 3 that our approach improves greatly the one from [4]. This can be explained by noticing that (t,f) points of the highest power in the t-f domain correspond to superposition of autoterms (see Fig. 1).

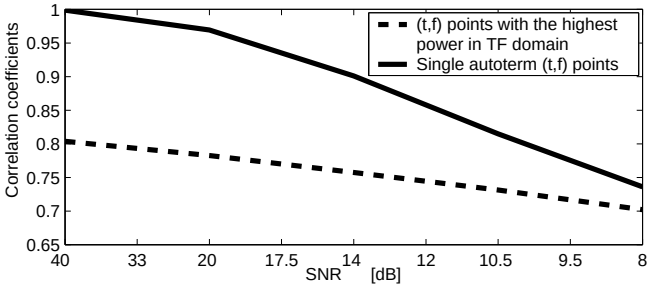

Figure 3: Correlation coefficients between the original and the estimated sources in dependence of SNR

6 CONCLUSIONS AND PERSPECTIVES

The BSS algorithm presented in [4] has proved to be very robust in low noise environments. In this paper we have concentrated on the selection of (t,f) points. We showed that the selection of the right matrices that enter the joint diagonalization procedure is crucial for the performance of BSS. Two general criteria for selecting (t,f) points were proposed. Criterion (16) eliminates the crossterms, while criterion (18) emphasizes the single autoterms. Both criteria were combined in Algorithm 1. An iterative joint-diagonalization procedure was proposed and with its help the BSS performance is no longer so strongly dependent on the selection of the single autoterm (t,f) points.

The definition of the decreasing level l in Algorithm 1 is not very strict and should be the object of further investigation. However, it guarantees the success of the iterative joint-diagonalization. Hence, as demonstrated in this paper, when combined, both algorithms greatly improve the performance and robustness of the BSS algorithm.

No word was spoken about choosing the most appropriate t-f distribution. Generally, the distribution that minimizes the interferences in the t-f plane is preferable, since it simplifies the selection of (t,f) points that represent true signal power. In our tests the Bessel reduced interference t-f distribution was used. However, other reduced interference t-f distributions, and signal dependent t-f distributions, that match the underlying signal characteristics are also possible.

References

[1] J. F. Cardoso. Blind signal separation: statistical principles. *Proceedings of the IEEE. Special issue on blind identification and estimation*, 9(10):2009–2025, October 1998.

[2] A. Belouchrani, K. A. Meraim, J. F. Cardoso, and É. Moulines. A blind source separation technique based on second order statistics. *IEEE Transactions on Signal Processing*, 45(2):434–44, February 1997.

[3] D. T. Pham and J. F. Cardoso. Blind separation of instantaneous mixtures of non stationary sources. *IEEE Transactions on Signal Processing*, 49(9):1837–1848, 2001.

[4] A. Belouchrani, and A. G. Moeness. Blind source separation based on time-frequency signal representation. *IEEE Transactions on Signal Processing*, 46(11), 1998.

[5] J. F. Cardoso and A. Souloumiac. Jacobi angles for simultaneous diagonalization. *SIAM J. Mat. Anal. Appl.*, 17(1):161–164, January 1996.

[6] L. Cohen. *Time-Frequency Analysis*. Prentice Hall PTR, Englewoo Cliffs, New Jersey, 1995.