NON-ORTHOGONAL ZERO-DIAGONALIZATION FOR SOURCE SEPARATION BASED ON TIME-FREQUENCY REPRESENTATION

El Mostafa Fadaili, Nadège Thirion-Moreau and Eric Moreau

STD, ISITV, University of Sud Toulon Var,
avenue George Pompidou, BP56,
F-83162 La Valette du Var, Cedex, FRANCE
e-mail: {fadaili,thirion,moreau}@univ-tln.fr

ABSTRACT

This paper is concerned with blind separation of source signals using time-frequency representations. We show that the separation can be realized through the non orthogonal joint zero diagonalization of spatial quadratic time frequency matrices. One advantage of the proposed method is that it does not require any whitening stage and thus it is intended to work even with a class of correlated signals.

1. INTRODUCTION

We consider the blind separation of instantaneous mixture of signals called sources. This problem has found numerous solutions in the past ten years. However, more recently, interest on solutions based on the use of time-frequency representations was growing [1]-[4]. One of the main reason is that it allows the possibility to consider a wider class of source signals rather than the classical one (statistically independent random source signals). One can distinguish two main classes of time-frequency representations: the linear one and the quadratic one. The use of quadratic time-frequency representations has lead to useful algorithms based on the joint diagonalization and/or the joint zero diagonalization of some particular sets of matrices [2].

One of the first approach [1] proposes to joint diagonalize a set of matrices that corresponds to spatial quadratic time-frequency representations calculated at some time-frequency points. These time-frequency points correspond to sources auto-terms only (there is no interference between source signals in such time-frequency points). However, a first whitening stage is required. Later in [4], it was shown that this whitening stage can be dropped which leads to advantages. Indeed better performances are generally obtained and the separation of "correlated" source signals can be considered.

On the other hand, the approach in [2] proposes to joint zero-diagonalize a set of matrices that corresponds again to spatial quadratic time-frequency representations calculated at some time-frequency points. Now these time-frequency points correspond only to sources interferences. Again a first whitening stage is required. Notice that solutions that combine joint diagonalization and joint zero-diagonalization were also proposed [3].

In this paper, we propose to generalize the joint zero-diagonalization approach to the case of non-orthogonal matrices. The proposed algorithm is based on the optimization of a quadratic criterion. The application of the algorithm to the source separation problem illustrates the usefulness of the proposed approach.

Since the proposed developments are based on the use of Spatial Quadratic Transforms (SQT) of signals and their properties [6][7][1], let us now briefly recall the important points related to our utilization.

Considering a real deterministic vectorial signal $\mathbf{z}(t)$, the SQT is given by a matrix $\mathbf{D}_z(t,\nu) = \left(D_{z_i,z_j}(t,\nu)\right)$ written as

$$\mathbf{D}_z(t,\nu) = \int_{\mathbb{R}^2} \mathbf{z}(\theta)\mathbf{z}^T(\theta')R(\theta,\theta';t,\nu)d\theta d\theta' \quad (1)$$

which is defined component-wise by

$$D_{z_i,z_j}(t,\nu) = \int_{\mathbb{R}^2} z_i(\theta)z_j(\theta')R(\theta,\theta';t,\nu)d\theta d\theta' \quad (2)$$

for all i and j. The diagonal terms of the SQT $\mathbf{D}_z(t,\nu)$ are called *auto-terms* while the off-diagonal ones are called *inter-terms*. The function $R(\theta,\theta';t,\nu)$ which is generally a complex function is referred to as the *kernel* of the transform. For physical reasons, this kernel is often constrained to satisfy the following property

$$R(\theta,\theta';t,\nu) = R^*(\theta',\theta;t,\nu) \quad (3)$$

where $(\cdot)^*$ stands for the complex conjugate operator. Then, the SQT satisfies an *hermitian symmetry* as

$$\mathbf{D}_z(t,\nu) = \mathbf{D}_z^H(t,\nu)$$

where $(\cdot)^H$ stands for the complex conjugate and transpose operator.

The auto-terms correspond to the same quadratic transform associated to different scalar deterministic signals. This quadratic transform is said *energetic* if its double integral over t and ν is equal to the energy of the considered signal, *i.e.* for a scalar signal $z(t)$ we have $\int\int D_{z,z}(t,\nu)dtd\nu = \int z^2(t)dt$. Such energetic transforms form the basis of diverse Time-Frequency Representations (TFR).

2. MODEL, ASSUMPTIONS AND SOLUTIONS

We consider the classical instantaneous blind sources separation problem where N sources signals are received on N sensors. In matrix and vector notations, the input/output relationship of the mixing model reads:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \tag{4}$$

with $\mathbf{A}$ the (N,N) real mixing matrix which is assumed invertible, $\mathbf{x}(t) = [x_1(t), \ldots, x_N(t)]^T$ the $(N,1)$ observations vector ($(\cdot)^T$ denotes the transposition operator) and $\mathbf{s}(t) = [s_1(t), \ldots, s_N(t)]^T$ the $(N,1)$ deterministic real sources vector.

The problem of blind sources separation consists in the estimation of a "separating" matrix, say $\mathbf{B}$, which applied to the observation as

$$\mathbf{y}(t) = \mathbf{B}\mathbf{x}(t) \tag{5}$$

yields an estimation of the source signals.

Defining $\mathbf{G} = \mathbf{B}\mathbf{A}$ as the matrix of the global system, the source separation problem is solved when one has found a separating matrix $\mathbf{B}$ in such a way that

$$\mathbf{G} = \mathbf{D}\mathbf{P} \tag{6}$$

where $\mathbf{D}$ is an invertible diagonal matrix which corresponds to arbitrary attenuations for the restored sources and $\mathbf{P}$ is a permutation matrix which corresponds to an arbitrary order of restitution of source signals.

Assumption A: Let us consider that we dispose of points in the time-frequency plane each of them corresponding to an inter-term and not to an auto-term. In other words, suppose that there exists couples (t_k, ν_k) such that

$$D_{s_i,s_j}(t_k,\nu_k) = (1-\delta_{i,j})D_{i,j,k} \tag{7}$$

where for all k there exists at least one couple (i,j) such that $D_{i,j,k} \neq 0$ and $\delta_{i,j} = 1$ if $i = j$ and 0 otherwise.

Notice that the above assumption for deterministic signals plays the role of the classical statistical independence assumption for random signals. It is clear that a discriminating property for source signals is always required in order to think about separation. Here we consider deterministic signals whose quadratic time-frequency representation do not overlap too much two by two in the above sense. In other words the signatures of the sources in the time-frequency plane are "sufficiently" different to be able to find time-frequency points satisfying the considered assumption.

3. JOINT ZERO-DIAGONALIZATION

Using (4) and because matrix $\mathbf{A}$ is real, it is easy to see that the SQT $\mathbf{D}_x(t,\nu)$ of the observation signals vector directly admits the following decomposition

$$\mathbf{D}_x(t,\nu) = \mathbf{A}\mathbf{D}_s(t,\nu)\mathbf{A}^T \tag{8}$$

where $\mathbf{D}_s(t,\nu)$ is the SQT of the source signals vector.

Notice that in general the matrix $\mathbf{D}_s(t,\nu)$ for any t and ν has no special structure. Nevertheless, there exists some time-frequency points for which the matrix $\mathbf{D}_s(t,\nu)$ has a specific structure. In particular, when assumption A is considered, $\mathbf{D}_s(t_k,\nu_k)$ are zero-diagonal. A zero-diagonal matrix being a matrix whose all diagonal components are zero. Our goal in the following consists in taking into consideration such a property.

Hence, we now briefly describe the problem of joint zero-diagonalization. Let us consider a set $\mathcal{D}$ of N_m matrices $\mathbf{D}_i$, $i \in \{1,\ldots,N_m\}$ which all admit the following decomposition: there exists a matrix $\mathbf{A}$ and N_m zero-diagonal matrices $\mathbf{\Lambda}_i$, $i \in \{1,\ldots,N_m\}$ such that

$$\mathbf{D}_i = \mathbf{A}\mathbf{\Lambda}_i\mathbf{A}^T\ , \quad \forall i \in \{1,\ldots,N_m\}\ .$$

The problem consists in estimating the matrix $\mathbf{A}$ and the zero-diagonal matrices $\mathbf{\Lambda}_i$, $i \in \{1,\ldots,N_m\}$ from the matrices set $\mathcal{D}$.

When $\mathbf{A}$ is orthogonal, the above problem have been reported in e.g. [2] where solutions can be found. For the non-orthogonal case, we propose to consider the following objective function

$$\mathcal{C}(\mathbf{B}) = \sum_{i=1}^{N_m} \|\mathsf{Diag}\{\mathbf{B}^T\mathbf{D}_i\mathbf{B}\}\|^2 \tag{9}$$

where the operator $\mathsf{Diag}\{\cdot\}$ is defined as the diagonal matrix built from the diagonal components of the matrix argument. In fact we are looking for the argument of the minimization of $\mathcal{C}(\mathbf{B})$. In that case, this optimal matrix argument plays directly the role of a separating matrix. The rationale of this objective function is to look for a matrix in order to yield signals whose quadratic time-frequency representations are zero-diagonal and thus corresponds to source signals.

For the optimization of $\mathcal{C}(\mathbf{B})$, let us remark that it can be written as

$$\begin{aligned}\mathcal{C}(\mathbf{B}) &= \sum_{i=1}^{N_m}\sum_{\ell=1}^{N}\left|\mathbf{b}_\ell^T\mathbf{D}_i\mathbf{b}_\ell\right|^2 \\ &= \sum_{\ell=1}^{N}\mathbf{b}_\ell^T\left(\sum_{i=1}^{N_m}\mathbf{D}_i\mathbf{b}_\ell\mathbf{b}_\ell^T\mathbf{D}_i^H\right)\mathbf{b}_\ell \\ &= \sum_{\ell=1}^{N}\mathbf{b}_\ell^T\mathbf{Q}_\ell(\mathbf{b}_\ell)\mathbf{b}_\ell \qquad (10)\end{aligned}$$

where $\mathbf{b}_\ell$, $\ell = 1, \dots, N$ are the column vectors of matrix $\mathbf{B}$, $(.)^H$ is the matrix conjugate transpose operator and $\mathbf{Q}_\ell(\mathbf{b}_\ell) = \sum_{i=1}^{N_m}\mathbf{D}_i\mathbf{b}_\ell\mathbf{b}_\ell^T\mathbf{D}_i^H$ is a quadratic form. Notice that the minimization of $\mathcal{C}(\mathbf{B})$ can be realized column by column. For a given column, an optimum of the quadratic form in (10) can be found by calculating the eigenvector of $\mathbf{Q}_\ell(\mathbf{b}_\ell)$ associated with the lowest eigenvalues. However since matrix $\mathbf{Q}_\ell(\mathbf{b}_\ell)$ for a given ℓ also depends on the vector $\mathbf{b}_\ell$, it is necessary to consider an iterative procedure. We propose to consider the following one:

For each ℓ, given $\mathbf{b}_\ell^{(0)}$ an initial unit norm vector with $i \in \mathbb{N}_*$, do *(a)* and *(b)*

(a) Calculate $\mathbf{Q}_\ell(\mathbf{b}_\ell^{(i-1)})$

(b) Find the lowest eigenvalue $\lambda^{(i)}$ and the associated eigenvector $\mathbf{b}_\ell^{(i)}$ of matrix $\mathbf{Q}_\ell(\mathbf{b}_\ell^{(i-1)})$

Stop when $|\lambda^{(i)} - \lambda^{(i-1)}| \leq \varepsilon$ where ε is a given small positive threshold.

4. COMPUTER SIMULATIONS

We consider $N = 3$ real synthetic source signals of 128 time samples. The first one is a sinusoïdal signal, the second one is a sinusoidal frequency modulation signal and the third one is a linear frequency modulation signal. In these simulations, the Spatial Pseudo Wigner-Ville (SPWV, [6][7]) representation is used. The real part (resp. the imaginary part) of the source SPWV, *i.e.* $\mathbf{D}_{\text{PWV},s}(t,\nu)$ is given on the top (resp. the bottom) of Fig. 1. Notice that it is computed over 64 frequency bins and with a Hamming window of length 33. One can observe that the diagonal terms of the SPWV are real to the extent that they correspond to the quadratic time-frequency representations of each of the 3 sources and because the kernel of the used SQTFR exhibits hermitian symmetry. With regard to the off-diagonal terms, they are complex: they correspond to the bilinear time-frequency representations of couples of different sources.

These source signals are mixed by the following mixing matrix:

$$\mathbf{A} = \begin{pmatrix} 1 & 0.9 & -0.5 \\ 0.3 & 1 & 0.4 \\ 0.2 & 0.1 & 1 \end{pmatrix}.$$

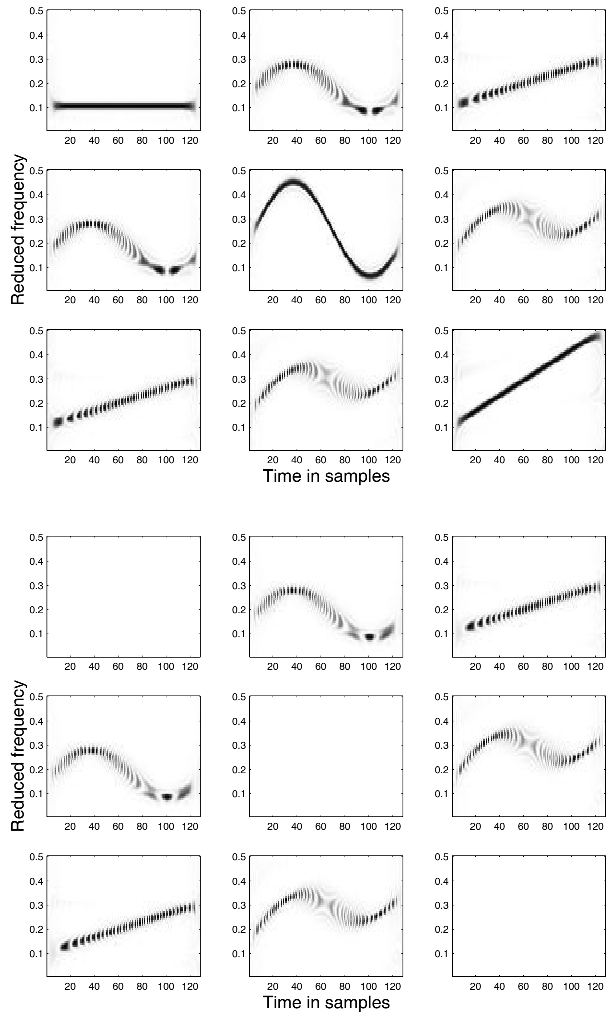

Fig. 1. Top: the real part of the SPWV of the sources vector, bottom: its imaginary part

Hence we consider that they are received on $M = 3$ sensors. The real part of the observation SPWV, *i.e.* $\mathbf{D}_{\text{PWV},x}(t,\nu)$ is given on Fig. 2

The used time-frequency points are displayed on Fig. 3 with a "plus". On the left, time-frequency points are those which are used for the building of a set of matrices to be joint-diagonalized. For an easier interpretation they are superimposed with the trace of the source SPWV. On the right, time-frequency points are those which are used for the building of a set of matrices to be joint zero-diagonalized. For an easier interpretation they are superimposed with the sum of the off-diagonal terms of the source SPWV. All those points have been obtained in an automatic mode using the selection procedures already proposed in [3] and [4]. They have led to the selection of 683 matrices to be joint-diagonalized and 987 to be joint zero-diagonalized.

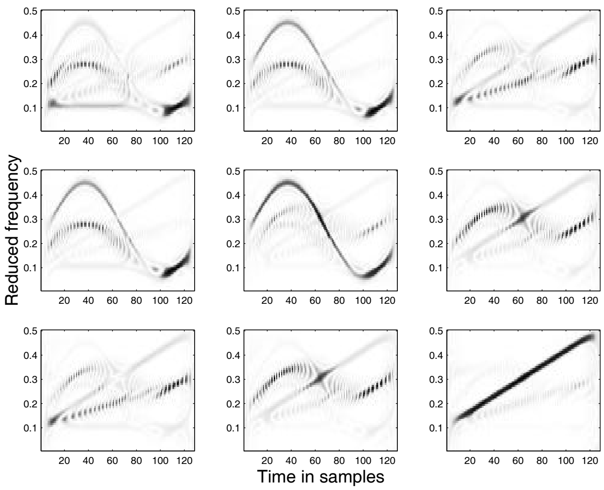

Fig. 2. The real part of the SPWV of the observations.

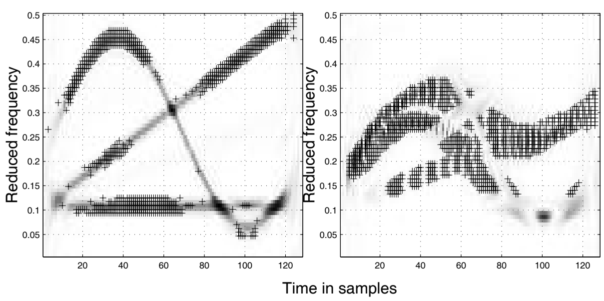

Fig. 3. The time-frequency points selected in an automatic mode. They are represented by a "+". Left: time-frequency points used for joint-diagonalization superimposed on the trace of the sources SPWV, right: time-frequency points used for joint zero-diagonalization superimposed on the sum of the off-diagonal terms of the sources SPWV.

We compare our proposed algorithm denoted by $\mathtt{JZD}_{\mathrm{NO}}$ with the unitary joint zero-diagonalization (JZD) algorithm proposed in [2] using the same set of 987 matrices. We also compare it with the non unitary joint-diagonalization ($\mathtt{JD}_{\mathrm{NO}}$) algorithm proposed in [4] and with the unitary joint-diagonalization (JD) algorithm proposed in [1] using the same set of 683 matrices.

To evaluate the performances of the separating algorithms we use the performance index I proposed in [5]. This index is given in dB defined by I dB $= 10\log(I)$. The resulting performances indexes are summed up in Table 1.

It is clear with the sight of these results, that in this case, the proposed non orthogonal joint zero-diagonalization method performs better than the others, and, more generally, that non orthogonal methods perform better than orthogonal ones.

Method	Performances (dB)
$\mathtt{JZD}_{\mathrm{NO}}$	-40.94
$\mathtt{JD}_{\mathrm{NO}}$	-35.55
JZD	-27.86
JD	-27.77

Table 1. A comparison of the performance indexes reached thanks to JD and $\mathtt{JD}_{\mathrm{NO}}$ performed on the same set of 683 time-frequency matrices and thanks to JZD and $\mathtt{JZD}_{\mathrm{NO}}$ performed on the same set of 987 time-frequency matrices.

5. DISCUSSION & CONCLUSION

In this paper, we have shown that blind sources separation based on spatial quadratic time-frequency representations can be performed without a preliminary whitening stage of the observations. To that aim, we propose a non-orthogonal joint zero-diagonalization procedure. One of the main advantage of such an approach is to apply even to potentially correlated sources.

6. REFERENCES

[1] A. Belouchrani and M. G. Amin, "Blind source separation based on time-frequency signal representations", *IEEE Trans. Signal Processing*, Vol. 46, No. 11, pp 2888-2897, Nov. 1998.

[2] A. Belouchrani, K. Abed-Meraim, M. G. Amin and A. M. Zoubir, "Joint anti-diagonalisation for blind source separation", in Proc *Int. Conf. on Accoustics, Speech and Signal Processing (ICASSP'2001)*, Salt Lake City, USA, pp 2789-2792, May 2001.

[3] L. Giulieri, N. Thirion-Moreau and P.-Y. Arquès, "Blind sources separation based on bilinear time-frequency representations: a performance analysis", in Proc *Int. Conf. on Accoustics, Speech and Signal Processing (ICASSP 2002)*, Orlando, USA, pp 1649-1652, May 2002.

[4] L. Giulieri, H. Ghennioui, N. Thirion-Moreau and E. Moreau, "Non orthogonal joint diagonalization of spatial quadratic time frequency matrices for source separation", accepted to *IEEE Signal Processing Letters*, 2005.

[5] E. Moreau, "A generalization of joint-diagonalization criteria for source separation", *IEEE Trans. Signal Processing*, Vol. 49, No 3, pp 530-541, March 2001.

[6] P. Flandrin, *Time-Frequency/Time-Scale Analysis*, Academic Press, 1999.

[7] L. Cohen, *Time-Frequency Analysis*, Prentice Hall, 1995.