

Algorithme de zéro-diagonalisation conjointe pour la séparation de sources déterministes

El Mostafa FADAILI, Nadège THIRION-MOREAU et Eric MOREAU

STD, ISITV, Université du Sud Toulon Var
avenue G. Pompidou, BP56, 83162 La Valette du Var Cedex
{fadaili, thirion, moreau}@univ-tln.fr

Résumé – Nous considérons le problème de la séparation aveugle de sources déterministes par zéro-diagonalisation conjointe d’un ensemble particulier de matrices issues des représentations temps-fréquence spatiales quadratiques des observations. Nous proposons un nouvel algorithme de zéro-diagonalisation conjointe sous transformation non unitaire, obtenu par optimisation paramétrique d’un critère que nous donnons. L’avantage d’une telle approche est de ne requérir aucun blanchiment préalable des observations. Nous proposons également un détecteur de points temps-fréquence permettant de construire l’ensemble des matrices à zéro-diagonaliser. Enfin, des simulations informatiques illustrent l’amélioration des performances en comparant l’algorithme proposé à d’autres algorithmes basés soit sur de la diagonalisation conjointe sous transformation unitaire ou non unitaire soit sur de la zéro-diagonalisation sous transformation unitaire.

Abstract – The problem of blind separation of deterministic sources based on joint zero-diagonalization of a particular set of spatial quadratic time-frequency matrices is considered. We propose a new non unitary joint zero-diagonalization algorithm. It is obtained thanks to the parametric optimization of a criterium we give. The advantage of such an approach is that it does not require a pre-whitening stage any more. We also detail a time-frequency points selection procedure used in order to build the set of matrices to be joint zero-diagonalized. Finally, computer simulations illustrate the effectiveness of the proposed algorithm by comparing it with other ones based on unitary or non unitary joint-diagonalization and on unitary joint zero-diagonalization.

1 Introduction

Dans ce papier, nous nous intéressons au problème de la séparation aveugle de mélanges instantanés de signaux sources basée sur l’utilisation de distributions temps-fréquence spatiales quadratiques (DTFSQ). Ces dernières années, ces méthodes ont fait l’objet de nombreux travaux [1]-[8][10] du fait qu’elles permettent de considérer une classe plus large de sources que la classe habituelle des sources aléatoires et indépendantes. Elles ont notamment permis d’aborder le cas de sources non stationnaires, même s’il existe d’autres manières de traiter ce problème [9]. Une partie des algorithmes proposés [1][3][7] reposent, après blanchiment, sur la diagonalisation conjointe d’un ensemble particulier de matrices.

Plus récemment, on s’est intéressé à ces deux problèmes que sont la suppression de l’étape de blanchiment [4]-[6][10] et la prise en compte d’un nouveau type de matrices temps-fréquence (t - f) se devant cette fois d’être zéro-diagonalisées conjointement [2][10]. Dans tous les cas, ces méthodes requièrent une étape préalable de sélection de points t - f “utiles” en ce sens qu’ils permettent la construction d’ensembles de matrices devant être diagonalisées conjointement et/ou zéro-diagonalisées conjointement. Il a d’abord été proposé dans [1] de diagonaliser conjointement les matrices issues des DTFSQ calculées en des points t - f ne correspondant qu’à des auto-termes sur les sources, cette méthode s’appliquant après blanchiment des observations. Par la suite, il a été montré dans [4]-[6], qu’en utilisant des algorithmes de diagonalisation conjointe (DC) non unitaire l’étape de blanchiment pouvait être supprimée ce qui présente de nombreux avantages parmi lesquels le fait que l’on obtienne généralement de meilleures performances et que l’on

puisse considérer le cas de sources “corrélées”.

Dans le même temps, une approche complémentaire fut proposée dans [2], consistant cette fois à zéro-diagonaliser conjointement (encore après une étape préalable de blanchiment) un autre type de matrices issues des DTFSQ : celles calculées en des points t - f ne correspondant qu’à des inter-termes sur les sources. Dans [10], nous avons proposé un premier algorithme de zéro-diagonalisation conjointe (ZDC) sous transformation non unitaire basé sur un schéma d’optimisation de type algébrique, l’avantage d’une telle approche étant à nouveau de ne plus requérir de blanchiment préalable des observations. Dans ce papier, nous en proposons un nouveau fondé cette fois sur un schéma d’optimisation de type Newton et basé sur une modélisation paramétrique des colonnes de la matrice de séparation recherchée du même type que celle utilisée dans [12]. Au préalable, nous présentons le critère sur lequel est fondé l’algorithme. Nous détaillons également une procédure de sélection automatique des points t - f permettant la construction d’un ensemble de matrices devant être zéro-diagonalisées conjointement.

2 Modèle & hypothèses

Considérons le problème de la séparation aveugle de sources dans lequel $N \in \mathbb{N}^+ \setminus \{0, 1\}$ signaux baptisés sources sont reçus au niveau d’une antenne de $P \in \mathbb{N}^+ \setminus \{0, 1\}$ capteurs. Plaçons-nous de plus dans le cas $P \geq N$, *i.e.* sur-déterminé (pour le cas sous-déterminé [8]). En notations matricielles, la relation qui lie l’entrée à la sortie du système mélangeant, est

donnée dans le cas non bruité par :

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

où \mathbf{A} est la matrice de mélange de dimensions $P \times N$, $\mathbf{x}(t) = (x_1(t), \dots, x_P(t))^T$ le vecteur $P \times 1$ des observations et $\mathbf{s}(t) = (s_1(t), \dots, s_N(t))^T$ le vecteur $N \times 1$ des sources et $(\cdot)^T$ l'opérateur de transposition. Dans toute la suite on se placera dans le cas de signaux sources réels et une matrice de mélange réelle.

Nous considérons les hypothèses suivantes :

- La matrice de mélange instantané \mathbf{A} est inconnue mais de rang plein.
- Le nombre de sources N est connu.
- Les sources $s_i(t)$, $i = 1, \dots, N$ sont des signaux déterministes.

Supposant les signaux sources inobservables et la matrice de mélange inconnue, l'objectif est d'estimer une "inverse" de la matrice de mélange afin de retrouver les contributions des différentes sources. En général, celles-ci ne sont restituées qu'à une amplitude et une permutation près. C'est pourquoi, on considère généralement que la matrice de mélange \mathbf{A} a été estimée quand la matrice obtenue, notée $\hat{\mathbf{A}}$, se factorise sous la forme $\hat{\mathbf{A}} = \mathbf{A} \mathbf{D} \mathbf{P}$ avec \mathbf{D} matrice diagonale inversible et \mathbf{P} matrice de permutation.

En ce qui concerne finalement la séparation de sources à proprement parler, quand $P \geq N$, *i.e.* dans le cas sur-déterminé, une estimée des sources $\hat{\mathbf{s}}(t)$ s'obtient au moyen de $\hat{\mathbf{A}}^\dagger \mathbf{x}(t)$ ($\hat{\mathbf{A}}^\dagger$ désignant la pseudo-inverse de la matrice $\hat{\mathbf{A}}$). Quand $P < N$, *i.e.* dans le cas sous-déterminé, il n'y a plus une solution unique et la séparation effective des sources nécessite des hypothèses et des développements supplémentaires.

3 Transformée spatiale quadratique

Une transformée spatiale quadratique (TSQ) est définie [11] comme :

$$\begin{aligned} (\mathbf{x}(t), \mathbf{x}(t)) &\stackrel{\text{TSQ}}{\longleftrightarrow} \mathbf{D}_{xx}(t, \nu; R) = \mathbf{D}_x(t, \nu) \\ &= \int_{\mathbb{R}^2} \mathbf{x}(\theta) \mathbf{x}^H(\theta') R(\theta, \theta'; t, \nu) d\theta d\theta' \end{aligned} \quad (2)$$

où R est le noyau de la transformation et où $(\cdot)^H$ désigne l'opérateur de transposition conjuguaison. Les termes diagonaux de cette matrice sont appelés "auto-termes", ils correspondent aux transformées quadratiques associées à chacune des composantes du vecteur $\mathbf{x}(t)$. Les termes hors diagonale sont quant à eux appelés "inter-termes", ils correspondent aux transformations bilinéaires associées à des couples de composantes distinctes de ce même vecteur. Pour des raisons physiques, on contraint souvent le noyau de la TSQ à satisfaire la propriété suivante

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

où $(\cdot)^*$ est l'opérateur de conjuguaison. Alors, la TSQ possède la propriété de symétrie hermitienne suivante :

$$\mathbf{D}_x(t, \nu) = \mathbf{D}_x^H(t, \nu). \quad (4)$$

Il s'avère en fait que cette condition n'est pas très restrictive dans la mesure où elle est satisfaite par de nombreuses transformations dont les distribution de Wigner-Ville (WV), de Choï-Williams (CW), de Pseudo Wigner-Ville (PWV), Pseudo Wigner-Ville lissée (moyennant l'adjonction de conditions sur

l'une des deux fenêtres qui doit être réelle), celle de Born-Jordan (BJ), etc...

On suppose également dans ce qui suit que les représentations t - f de chaque signaux sources ne se recouvrent pas trop ceci dans le but de pouvoir sélectionner des points t - f en vue de la construction des ensembles de matrices à diagonaliser ou à zéro-diagonaliser conjointement.

On montre enfin [13] que la DTFSQ $\mathbf{D}_x(t, \nu)$ du vecteur des observations admet naturellement la décomposition suivante :

$$\mathbf{D}_x(t, \nu) = \mathbf{A} \mathbf{D}_s(t, \nu) \mathbf{A}^T \quad (5)$$

où $\mathbf{D}_s(t, \nu)$ est la DTFSQ du vecteur des signaux sources.

4 Zéro-diagonalisation non unitaire

Soit un ensemble $\{\mathbf{Z}_i, i = 1, \dots, N_m\}$ de matrices hermitiennes tel qu'il existe une matrice \mathbf{B} telle que, pour tout i ,

$$\mathbf{B} \mathbf{Z}_i \mathbf{B}^T$$

soit une matrice zéro-diagonale, *i.e.* une matrice dont les éléments diagonaux sont tous nuls. Dans le but d'estimer la matrice \mathbf{B} à partir de la factorisation précédente, on propose l'utilisation de la fonction objectif suivante :

$$\mathcal{C}(\mathbf{B}) = \sum_{i=1}^{N_m} \|\text{Diag}\{\mathbf{B} \mathbf{Z}_i \mathbf{B}^T\}\|^2 \quad (6)$$

où l'opérateur $\text{Diag}\{\cdot\}$ renvoie une matrice diagonale construite à partir de la diagonale de la matrice en argument. Il s'agit donc de trouver une matrice en minimisant par rapport à \mathbf{B} la somme des carrés des éléments de la diagonale de $\mathbf{B} \mathbf{Z}_i \mathbf{B}^T$.

Afin d'effectuer cette minimisation, nous proposons d'utiliser un algorithme de gradient ou de Newton. Cependant, avant de calculer le gradient de \mathcal{C} , il convient de remarquer que l'on peut toujours écrire $\mathcal{C}(\mathbf{B})$ sous la forme

$$\mathcal{C}(\mathbf{B}) = \sum_{i=1}^{N_m} \sum_{\ell=1}^N |\mathbf{b}_\ell \mathbf{Z}_i \mathbf{b}_\ell^T|^2 = \sum_{\ell=1}^N \mathcal{C}_i(\mathbf{b}_\ell) \quad (7)$$

où les vecteurs \mathbf{b}_ℓ , $\ell = 1, \dots, N$ sont les vecteurs lignes de la matrice \mathbf{B} et où

$$\mathcal{C}_i(\mathbf{b}) = \sum_{i=1}^{N_m} |\mathbf{b} \mathbf{Z}_i \mathbf{b}^T|^2. \quad (8)$$

Ainsi la minimisation de $\mathcal{C}(\mathbf{B})$ peut se faire ligne par ligne à partir de $\mathcal{C}_i(\mathbf{b})$. Donc considérons un vecteur ligne quelconque noté \mathbf{b} en lui imposant une norme égale à 1 et supposons que celui-ci dépende de paramètres regroupés dans un $(K, 1)$ vecteur $(K = N - 1)$ noté $\boldsymbol{\theta} = (\theta_k)$. Une telle paramétrisation peut par exemple se trouver dans [12]. Comme les matrices \mathbf{Z}_i sont hermitiennes, le terme $\mathbf{b} \mathbf{Z}_i \mathbf{b}^T$ est réel et on a

$$\frac{\partial \mathcal{C}_i(\mathbf{b})}{\partial \theta_k} = 2 \sum_{i=1}^{N_m} (\mathbf{b} \mathbf{Z}_i \mathbf{b}^T) \frac{\partial (\mathbf{b} \mathbf{Z}_i \mathbf{b}^T)}{\partial \theta_k}. \quad (9)$$

Comme par ailleurs on a

$$\frac{\partial (\mathbf{b} \mathbf{Z}_i \mathbf{b}^T)}{\partial \theta_k} = 2 \frac{\partial \mathbf{b}}{\partial \theta_k} \mathbf{Z}_i \mathbf{b}^T \quad (10)$$

en introduisant la matrice $\frac{\partial \mathbf{b}}{\partial \boldsymbol{\theta}}$ définie par

$$\left(\frac{\partial \mathbf{b}}{\partial \boldsymbol{\theta}} \right)_{l,m} = \frac{\partial b_m}{\partial \theta_l} \quad (11)$$

on obtient

$$\frac{\partial (\mathbf{b} \mathbf{Z}_i \mathbf{b}^\top)}{\partial \boldsymbol{\theta}} = 2 \frac{\partial \mathbf{b}}{\partial \boldsymbol{\theta}} \mathbf{Z}_i \mathbf{b}^\top \quad (12)$$

et donc en utilisant en plus (10) dans (9)

$$\mathbf{g}_{c_l}(\boldsymbol{\theta}) \stackrel{\text{def}}{=} \frac{\partial C_l(\mathbf{b})}{\partial \boldsymbol{\theta}} = 4 \sum_{i=1}^{N_m} (\mathbf{b} \mathbf{Z}_i \mathbf{b}^\top) \frac{\partial \mathbf{b}}{\partial \boldsymbol{\theta}} \mathbf{Z}_i \mathbf{b}^\top. \quad (13)$$

Le calcul de $\frac{\partial \mathbf{b}}{\partial \boldsymbol{\theta}}$ dépend de la paramétrisation considérée. Nous considérons dans la suite celle explicitée dans [12]. Un algorithme de gradient s'écrit alors

$$\boldsymbol{\theta}^i = \boldsymbol{\theta}^{i-1} - \lambda \mathbf{g}_{c_l}(\boldsymbol{\theta}^{i-1}) \quad (14)$$

où λ est une constante positive.

Cependant dans notre cas le Hessien peut aussi être calculé afin d'obtenir un algorithme de Newton. D'abord, à partir de (9), on a dans le cas scalaire

$$\frac{\partial^2 C_l(\mathbf{b})}{\partial \theta_k \partial \theta_\ell} = 2 \sum_{i=1}^{N_m} \frac{\partial (\mathbf{b} \mathbf{Z}_i \mathbf{b}^\top)}{\partial \theta_k} \frac{\partial (\mathbf{b} \mathbf{Z}_i \mathbf{b}^\top)}{\partial \theta_\ell} + (\mathbf{b} \mathbf{Z}_i \mathbf{b}^\top) \frac{\partial^2 (\mathbf{b} \mathbf{Z}_i \mathbf{b}^\top)}{\partial \theta_k \partial \theta_\ell}. \quad (15)$$

Ecrivons maintenant l'expression précédente sous une forme vectorielle. A cette fin, on pose

$$\frac{\partial^2 \mathbf{b}}{\partial \boldsymbol{\theta}^2} = \left(\left(\frac{\partial^2 \mathbf{b}}{\partial \boldsymbol{\theta} \partial \theta_1} \right)^\top, \dots, \left(\frac{\partial^2 \mathbf{b}}{\partial \boldsymbol{\theta} \partial \theta_K} \right)^\top \right)^\top \quad (16)$$

et

$$\frac{\partial^2 C_l(\mathbf{b})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} = \left(\frac{\partial^2 C_l(\mathbf{b})}{\partial \theta_k \partial \theta_\ell} \right) \quad (17)$$

On montre alors que

$$\mathbf{H}_{c_l}(\boldsymbol{\theta}) \stackrel{\text{def}}{=} \frac{\partial^2 C_l(\mathbf{b})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} = 4 \sum_{i=1}^{N_m} \left(C_i(\boldsymbol{\theta}) + (\mathbf{b} \mathbf{Z}_i \mathbf{b}^\top) \mathbf{D}_i(\boldsymbol{\theta}) \right) \quad (18)$$

où

$$C_i(\boldsymbol{\theta}) = 2 \frac{\partial \mathbf{b}}{\partial \boldsymbol{\theta}} \mathbf{Z}_i \mathbf{b}^\top \mathbf{b} \mathbf{Z}_i \left(\frac{\partial \mathbf{b}}{\partial \boldsymbol{\theta}} \right)^\top \quad (19)$$

et

$$\mathbf{D}_i(\boldsymbol{\theta}) = \text{Unvec} \left\{ \frac{\partial^2 \mathbf{b}}{\partial \boldsymbol{\theta}^2} \mathbf{Z}_i \mathbf{b}^\top \right\} + \frac{\partial \mathbf{b}}{\partial \boldsymbol{\theta}} \mathbf{Z}_i \left(\frac{\partial \mathbf{b}}{\partial \boldsymbol{\theta}} \right)^\top \quad (20)$$

et où l'opérateur $\text{Unvec}\{\cdot\}$ renvoie une matrice carrée à partir de son argument vectoriel. Finalement l'algorithme de Newton s'écrit sous la forme

$$\boldsymbol{\theta}^i = \boldsymbol{\theta}^{i-1} - \lambda \mathbf{H}_{c_l}^{-1}(\boldsymbol{\theta}^{i-1}) \mathbf{g}_{c_l}(\boldsymbol{\theta}^{i-1}) \quad (21)$$

où λ est une constante positive. Notons que dans le cas rectangulaire, pour éviter des problèmes d'inversion du Hessien, on passe à un schéma d'optimisation de type Levenberg-Marquardt.

5 Sélection automatique des points t - f

Il existe 4 types distincts de points t - f : 1- ceux qui ne correspondent qu'à des auto-termes sur les sources (en ces points la

DTFSQ des sources est diagonale, propriété perdue sous l'effet du mélange), 2- ceux qui ne correspondent qu'à des intertermes sur les sources (en ces points la DTFSQ des sources est zéro-diagonale, propriété perdue sous l'effet du mélange), 3- ceux qui correspondent à la fois à des inter- et à des auto-termes sur les sources (en ces points la DTFSQ des sources ne possède pas de structure algébrique particulière), et enfin 4- ceux où il n'y a ni inter- ni auto-termes sur les sources; seuls les deux premiers s'avèrent intéressants d'un point de vue de la séparation de sources. Dans cette partie, nous expliquons la sélection automatique des points t - f permettant la construction de l'ensemble des matrices à zéro-diagonaliser. En outre, pour pouvoir comparer les algorithmes de ZDC unitaire et non unitaire sur des ensembles identiques de points, nous présentons une procédure de sélection des points t - f à partir des DTFSQ des observations blanches. Pour la sélection automatique des points t - f permettant la construction de l'ensemble des matrices à diagonaliser conjointement, on procède comme dans [3] car les points t - f sont sélectionnés, pour les mêmes raisons que précédemment, à partir des DTFSQ des observations blanches (dans [4] ou [6], on décrit deux autres procédures de sélection automatique à partir des DTFSQ des observations non blanches cette fois).

Etant donné les hypothèses faites (\mathbf{A} réelle, DTFSQ à symétrie hermitienne), on montre que les DTFSQ des sources, $\mathbf{D}_s(t, \nu)$, et des observations blanches, $\mathbf{D}_x(t, \nu)$, en des points t - f ne correspondant qu'à des auto-termes sur les sources sont des matrices réelles et non nulles alors qu'elles sont généralement complexes en des points t - f ne correspondant qu'à des intertermes sur les sources. La trace ($\text{trace}\{\cdot\}$) étant invariante sous transformation unitaire, on a par ailleurs $\text{trace}\{\mathbf{D}_x(t, \nu)\} = \text{trace}\{\mathbf{D}_s(t, \nu)\}$. L'une des manières les plus naturelles de sélectionner les matrices issues des DTFSQ (des observations blanches) à zéro-diagonaliser consiste à choisir celles qui sont telles que la valeur absolue de leur trace ($|\text{trace}\{\cdot\}|$) soit faible (inférieure à un seuil ε_2) tandis que la norme euclidienne de leur partie imaginaire ($\|\Im\{\cdot\}\|$) soit grande (supérieure à un seuil ε_1). Ce qui donne en résumé :

• pour la ZDC choisir (t, ν) tels que :

$$\begin{cases} \|\Im\{\mathbf{D}_x(t, \nu)\}\| > \varepsilon_1 \\ |\text{trace}\{\mathbf{D}_x(t, \nu)\}| < \varepsilon_2 \end{cases} \quad (22)$$

6 Simulations informatiques

Dans cette partie, nous illustrons au moyen de simulations informatiques l'intérêt de cet algorithme en comparant les performances obtenues avec celles d'autres algorithmes basés sur de la DC (unitaire ou non) et de la ZDC (unitaire). On utilise la distribution de Pseudo-Wigner Ville spatiale. Dans cet exemple, on considère le cas de 2 sources (une somme de sinusoides, une modulation sinusoidale de fréquence) reçues sur 2 capteurs. La partie réelle de la distribution de PWV spatiale des sources est représentée au niveau de la figure 1, celle de la distribution de PWV spatiale des observations au niveau de la figure 2. Les 651 points t - f détectés et utilisés par tous les algorithmes que nous comparons sont représentés par un "plus" au niveau de la figure 3 (haut), à gauche pour la DC et à droite pour la ZDC. Pour faciliter la compréhension, ils sont superposés à la trace de la DTFSQ des sources pour la DC et à la

somme des termes hors diagonale pour la ZDC. Ils ont été obtenus en fixant $\varepsilon_1 = 5.9$ et $\varepsilon_2 = 4.5$. En figure 3 (bas) les points ont été choisis à la main. En utilisant un indice de performance classique [6], les résultats obtenus sont résumés au niveau du Tableau 1. On compare la DC (JD) et la ZDC (JZD) des matrices sélectionnées après une étape de blanchiment. On donne aussi le résultat obtenu par l’algorithme classique JADE. On compare finalement la DC (JD_{NO,YER}) et l’algorithme proposé basé sur la ZDC (JZD_{NO,PAR}) des matrices sélectionnées, ces deux derniers algorithmes opérant sans étape de blanchiment. Dans cet exemple, les résultats obtenus sont meilleurs avec l’algorithme proposé et dans tous les cas bien meilleurs que lorsqu’une étape de blanchiment est considérée.

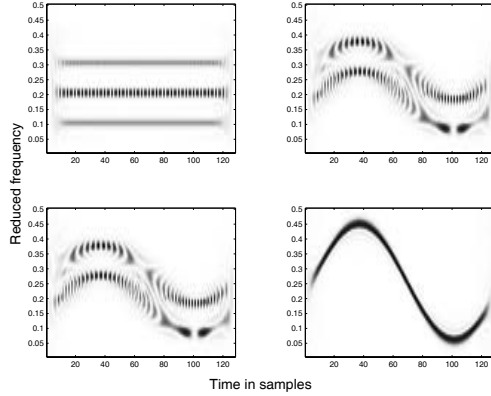


FIG. 1: Partie réelle de la distribution de PWV spatiale des sources

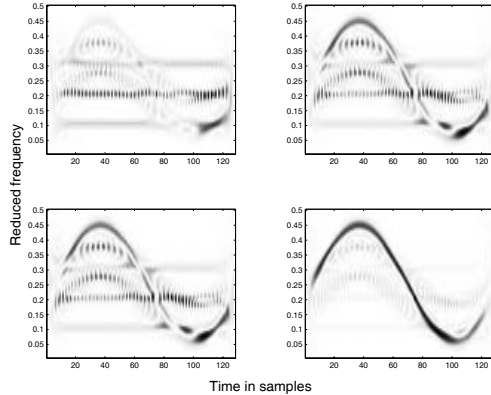


FIG. 2: Partie réelle de la distribution de PWV spatiale du mélange

TAB. 1: Comparaison des performances obtenues au moyen de : la méthode proposée (JZD_{NO,PAR}), JADE, DC non orthogonale (JD_{NO,YER}), DC orthogonale (JD), ZDC orthogonale (JZD) sur des ensembles de points obtenus en mode automatique ou choisis à la main.

Méthodes	Perf. (mode auto.)	Perf. (main)
JD (dB)	-29.3	-28.76
JZD (dB)	-29.44	-29.49
JD _{NO,YER} (dB)	-37.56	-42.01
JZD _{NO,PAR} (dB)	-40.31	-42.1
JADE (dB)	-21.48	

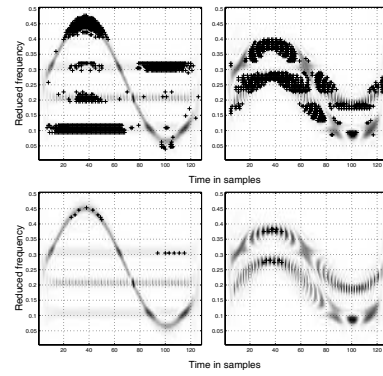


FIG. 3: Haut : les 651 points *t-f* sélectionnés de manière automatique pour la DC (gauche) ; les 651 retenus pour la ZDC (droite) ; Bas : points choisis à la main (14 pour la DC et 13 pour la ZDC)

Références

- [1] A. Belouchrani, M. G. Amin, “Blind source separation based on time-frequency signal representations,” *IEEE Trans. Signal Processing*, Vol. 46, No. 11, pp. 2888-2897, November 1998.
- [2] A. Belouchrani, K. Abed-Meraim, M. G. Amin, A. M. Zoubir, “Joint anti-diagonalisation for blind source separation”, In Proc. ICASSP’2001, Salt Lake City, USA, pp. 2789-2792, May 2001.
- [3] L. Giulieri, N. Thirion-Moreau, P.-Y. Arquès, “Blind sources separation based on bilinear time-frequency representations: a performance analysis”, In Proc. ICASSP’2002, Orlando, USA, pp. 1649-1652, May 2002.
- [4] L. Giulieri, N. Thirion-Moreau and P.-Y. Arquès, “Blind sources separation based on quadratic time-frequency representations: a method without pre-whitening”, In Proc. ICASSP’2003, Hong-Kong, pp. 289-292, April 2003.
- [5] A. Bousbia-Salah, A. Belouchrani and H. Bousbia-Salah, “A one step time-frequency blind identification”, in Proc. ISSPA’2003, Paris, France, July 2003.
- [6] L. Giulieri, H. Ghennioui, N. Thirion-Moreau and E. Moreau, “Non-Orthogonal Joint-Diagonalization of Spatial Quadratic Time Frequency Matrices for Source Separation”, *IEEE Signal Processing Letters*, Vol. 12, n°5, pp. 415-418, May 2005.
- [7] C. Févotte and C. Doncarli, “Two contributions to blind source separation using time-frequency distributions”, *IEEE Signal Processing Letters*, Vol. 11, No. 3, pp. 386-389, March 2004.
- [8] N. Linh-Trung, A. Belouchrani, K. Abed-Meraim and B. Boashash, “Separating more sources than sensors using time-frequency distributions”, in Proc. ISSPA’2001, Kuala-Lumpur, Malaysia, August 2001.
- [9] Pham, D.-T. and Cardoso, J.-F, “Blind separation of instantaneous mixtures of non-stationary sources”, In Proc. ICA’2000, Helsinki, Finland, pp. 187-193, 2000.
- [10] E. M. Fadaili, N. Thirion-Moreau and E. Moreau, “Non-Orthogonal Zero-Diagonalization for Source Separation Based on Time-Frequency Representation”, in Proc. ICASSP’2005, Philadelphia PA, USA, vol. IV, pp. 297-300, March 2005.
- [11] P. Flandrin, *Time-Frequency/Time-Scale Analysis*, Academic Press, 1999.
- [12] A.J. Van Der Veen, “Joint diagonalization via subspace fitting techniques”, In Proc. ICASSP’2001, pp. 2773-2776, Salt Lake City, USA, May 2001.
- [13] Y. Zhang and M. Amin, “Subspace analysis of spatial time-frequency distribution matrices”, in *IEEE Trans. on Signal Processing*, Vol.49, No. 4, pp.747-759, April 2001.