

Séparation de sources doublement non stationnaire

Adrien MEYNARD

Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France

adrien.meynard@univ-amu.fr

Résumé – Les techniques de séparation aveugle de sources (*blind source separation*, BSS) visent à estimer conjointement des signaux sources et une matrice de mélange à partir des observations de mélanges. Cet article aborde un problème de BSS doublement non stationnaire, où la matrice de mélange dépend du temps et les sources sont non stationnaires. Plus précisément, ce sont des signaux stationnaires déformés, suivant le modèle de [1]. Nous présentons un algorithme joint pour la BSS et l'estimation des déformations brisant la stationnarité et des spectres. Il s'appuie sur des approximations appropriées du comportement de la transformée en ondelettes de ces signaux non stationnaires. Les performances cet algorithme sont évaluées sur des simulations numériques et comparées à d'autres algorithmes de BSS non stationnaires.

Abstract – Blind source separation (BSS) techniques aims at joint estimation of source signals and a mixing matrix from observations of mixtures. This paper addresses a doubly nonstationary BSS problem, where the mixing matrix is time dependent and sources are nonstationary, more precisely deformed stationary signals, following the model of [1]. An algorithm for joint BSS and estimation of stationarity-breaking deformations and spectra is introduced, that exploits suitable approximations for the behavior of the wavelet transform of such nonstationary signals. The performance of the approach is evaluated on numerical simulations, and compared with other nonstationary BSS algorithms.

1 Introduction

La plupart des algorithmes de BSS s'appuient sur une hypothèse de stationnarité des sources à estimer, et traitent souvent aussi des mélanges linéaires instantanés où la matrice de mélange est constante dans le temps (voir [2] pour une étude approfondie du sujet). Par exemple, SOBI [3] constitue une approche de référence qui exploite les statistiques de second ordre et dont l'estimation repose sur les diagonalisations approchées et jointes des matrices de covariance. Cependant, ces hypothèses sur le modèle de mélange limitent l'application de la BSS à un petit nombre de signaux et de situations.

Plusieurs extensions à la BSS de mélanges linéaires instantanés de signaux non stationnaires ont été proposées. La plupart d'entre elles reposent sur une analyse temps-fréquence quadratique (voir par ex. [4, 5]) et s'appuient sur une sélection de points du domaine temps-fréquence pour lesquels une seule des sources est active (se référer à [4] pour plus de détails sur la sélection des points dans le cas d'une distribution de Wigner modifiée). Nous mettrons en œuvre une de ces techniques pour comparaison, appelée QTF-BSS dans le document. Des algorithmes de BSS, fondés sur l'indépendance entre sources non stationnaires et l'indépendance temporelle de chaque source, ont également été proposés. Dans [6], l'idée est d'approcher les sources non stationnaires par des signaux stationnaires par morceaux. Ainsi, cette méthode s'approche d'une estimation SOBI par morceaux sur des sous-intervalles ne se chevauchant pas (elle est notée p-SOBI dans l'article).

En outre, afin de modéliser des situations physiques (par ex. : réverbération de signaux audio), des modèles de mélange non instantanés ont été proposés. Dans [7], les auteurs considèrent le cas très général d'un mélange convolutif variant dans le temps (ce qui inclue le mélange instantané non constant).

Néanmoins, la méthode de BSS associée repose sur l'hypothèse qu'il existe une représentation des sources dans laquelle elles sont parcimonieuses. Il s'agit alors, comme pour QTF-BSS, de sélectionner les points où une seule source est active. Concernant la BSS d'un mélange linéaire variant dans le temps, l'approche de différents algorithmes de BSS (voir par ex. [8]) consiste à effectuer, à l'instar de p-SOBI, une analyse en composantes indépendantes sur de courtes portions du mélange sur lesquelles la stationnarité du mélange peut être supposée.

Dans ce qui suit, nous nous intéressons à un problème de BSS doublement non stationnaire et abordons le problème des signaux non stationnaires, mélangés par une matrice de mélange non stationnaire. Sur la base de travaux antérieurs sur la non-stationnarité et la déformation temporelle [9, 10], nous nous concentrons sur une classe spécifique de signaux non stationnaires, à savoir les signaux stationnaires composés avec une déformation temporelle. Dans [1], on analyse ce type de signaux individuellement. On étend ici cette analyse à des mélanges de tels signaux, et nous montrons que si les matrices de mélange et les fonctions de déformation temporelle sont suffisamment lisses et lentement variables, la vraisemblance approchée sur les transformées en ondelettes des observations peut être déduite sous une hypothèse de sources gaussiennes. Par conséquent, l'estimation conjointe de la matrice de mélange et des fonctions de déformation temporelle constitue un problème de maximum de vraisemblance. L'algorithme de BSS proposé consiste en une estimation alternée de la matrice de mélange et des fonctions de déformation. Cette dernière estimation est effectuée par l'algorithme JEFAS (*Joint Estimation of Frequency, Amplitude and Spectrum*) introduit dans nos travaux précédents (voir [1], code MATLAB disponible en ligne ¹).

1. <https://github.com/AdMeynard/JEFAS>

2 Modèle

2.1 Une classe de signaux non stationnaires

Dans ce qui suit, nous considérons les signaux non stationnaires comme des versions déformées de signaux stationnaires. Ces hypothèses sont appropriées pour décrire de nombreux signaux de la vie réelle. En effet, différentes classes d'opérateurs qui brisent la stationnarité sont pertinentes pour modéliser des phénomènes physiques (par ex. la modulation de fréquence, ou la modulation d'amplitude [1]). Nous nous concentrons ici sur l'opérateur de déformation temporelle. Un tel opérateur peut modéliser des phénomènes physiques non stationnaires aussi divers que l'effet Doppler, les variations de vitesse d'un moteur, la vocalisation d'animaux ou la parole [11, 1].

Introduisons maintenant quelques notations. Soit x un signal stationnaire, modélisé comme la réalisation d'un processus aléatoire stationnaire X dont le densité spectrale de puissance est notée \mathcal{S}_X . En agissant sur x avec l'opérateur de déformation temporelle noté \mathcal{D}_γ , on obtient un signal non stationnaire noté y . La déformation temporelle est définie par :

$$y(t) = \mathcal{D}_\gamma x(t) = \sqrt{\gamma'(t)} x(\gamma(t)), \quad (1)$$

où $\gamma \in C^2$ est une fonction strictement croissante.

La transformée en ondelettes est un outil naturel pour analyser de tels signaux. À partir d'une ondelette complexe ψ , on définit la transformée en ondelettes \mathcal{W}_x du signal x par :

$$\mathcal{W}_x(s, \tau) = \int_{\mathbb{R}} x(t) q^{-\frac{s}{2}} \bar{\psi} \left(\frac{t - \tau}{q^s} \right) dt \quad \text{avec } q > 1. \quad (2)$$

Dans ce cadre, on peut montrer [10, 1] que les transformées en ondelettes respectives \mathcal{W}_y et \mathcal{W}_x de y et x sont approximativement reliées par l'équation suivante :

$$\mathcal{W}_y(s, \tau) \approx \mathcal{W}_x(s + \log_q(\gamma'(\tau)), \gamma(\tau)). \quad (3)$$

Dans le cadre de processus aléatoires, on peut montrer [10, 1] que l'erreur d'approximation est non biaisée et que sa variance peut être contrôlée grâce aux propriétés de décroissance de l'ondelette ψ et aux variations de γ' .

2.2 Mélange instantané non stationnaire

Le problème que nous considérons est la BSS de signaux non stationnaires modélisés par l'équation (1). Les sources sont en outre supposées non corrélées. Nous supposons également que le nombre de sources M est égal au nombre de mélanges N . Dans le cas surdéterminé, on effectue classiquement une réduction de dimension pour revenir au cas $M = N$. Le cas sous-déterminé dépasse le cadre du présent travail.

Soit $\mathbf{y}(t), \mathbf{z}(t) \in \mathbb{R}^N$ les vecteurs colonnes contenant respectivement toutes les sources et observations à l'instant t . Le mélange s'écrit alors

$$\mathbf{z}(t) = \mathbf{A}(t)\mathbf{y}(t), \quad (4)$$

où $\mathbf{A}(t) \in \mathbb{R}^{N \times N}$ désigne la matrice de mélange variant dans le temps, supposée être inversible. Ce modèle généralise le modèle de modulation d'amplitude correspondant au cas où $\mathbf{A}(t)$ est diagonale.

Notre objectif est de déterminer conjointement la matrice de mélange $\mathbf{A}(t)$, les fonctions de déformation temporelle $\gamma_i(t)$ et le spectre des sources stationnaires \mathcal{S}_{X_i} pour $i \in \{1, \dots, N\}$ à partir des observations $\mathbf{z}(t)$.

Plaçons nous à un instant fixe τ . Puis, pour chaque observation z_i , on note $\mathbf{w}_{z_i, \tau} = \mathcal{W}_{z_i}(\mathbf{s}, \tau)$ le vecteur ligne contenant les valeurs de la transformée en ondelettes à l'instant τ pour un vecteur d'échelles \mathbf{s} (dont la taille est notée M_s). Ensuite, tous ces vecteurs sont rassemblés dans une matrice de taille $N \times M_s$ notée $\mathbf{w}_{\mathbf{z}, \tau}$ telle que $\mathbf{w}_{\mathbf{z}, \tau} = (\mathbf{w}_{z_1, \tau}^T \cdots \mathbf{w}_{z_N, \tau}^T)^T$. La même notation est utilisée pour la transformée en ondelettes des sources $\mathbf{w}_{\mathbf{y}, \tau}$. On suppose que la matrice $\mathbf{A}(t)$ varie lentement vis-à-vis des oscillations des signaux. On peut alors montrer que la relation linéaire (4) devient (sous cette hypothèse) une relation entre les transformées en ondelettes de \mathbf{y} et \mathbf{z} de la forme

$$\mathbf{w}_{\mathbf{z}, \tau} \approx \mathbf{A}(\tau)\mathbf{w}_{\mathbf{y}, \tau}. \quad (5)$$

Le théorème suivant (où $\mathbf{M}^{\circ 2}$ désigne le carré élément par élément d'une matrice \mathbf{M}) donne une borne quantitative sur l'erreur d'approximation dans l'équation (5).

Théorème 1. *Avec les notations ci-dessus, soit $\boldsymbol{\epsilon}_\tau \in \mathbb{C}^{N \times M_s}$ l'erreur d'approximation dans le domaine des ondelettes, définie comme suit :*

$$\boldsymbol{\epsilon}_\tau = \mathbf{w}_{\mathbf{z}, \tau} - \mathbf{A}(\tau)\mathbf{w}_{\mathbf{y}, \tau}. \quad (6)$$

Supposons que les sources stationnaires sous-jacentes X_i ($i = 1, \dots, N$) soient des processus aléatoires stationnaires au second ordre, de moyenne nulle, de puissance σ_X^2 . Alors l'erreur d'approximation $\boldsymbol{\epsilon}_\tau$ est une matrice aléatoire complexe du second ordre, de moyenne nulle et circulaire. De plus, la variance des coefficients de la matrice d'erreur est bornée comme suit :

$$\mathbb{E} \{ |\boldsymbol{\epsilon}_\tau|^{\circ 2} \} \leq \sigma_X^2 k_\psi^2 \mathbf{A}'^{\circ 2} \gamma'_\infty (q^{3s})^T, \quad (7)$$

où

$$\begin{aligned} k_\psi \in \mathbb{R}_+ &: & k_\psi &= \int_{\mathbb{R}} |t \psi(t)| dt, \\ \mathbf{A}'_\infty \in \mathbb{R}_+^{N \times N} &: & (\mathbf{A}'_\infty)_{ij} &= \sup_t |\mathbf{A}'_{ij}(t)|, \\ \gamma'_\infty \in \mathbb{R}_+^{N \times 1} &: & (\gamma'_\infty)_i &= \sup_t |\gamma'_i(t)|. \end{aligned}$$

La démonstration repose sur les mêmes arguments que celle du théorème 1 dans [1]. Un développement de Taylor de \mathbf{A} en τ permet la construction de la borne donnée ci-dessus.

Notons que, mis à part les termes contrôlant la borne sur l'erreur dans (3), c'est-à-dire la résolution temporelle de l'ondelette et les variations des fonctions de déformation temporelle, la borne sur l'erreur dans (5) est également contrôlée par les variations des coefficients de la matrice de mélange.

3 Procédure d'estimation

La procédure d'estimation repose sur les relations approchées (3) et (5), que nous supposons valides, c'est-à-dire que \mathbf{A} et γ' varient suffisamment lentement. Cette hypothèse nous permet d'écrire une vraisemblance approchée sur les transformées en ondelettes des observations dans le cas gaussien. L'estimation s'appuie sur des transformées en ondelettes discrètes,

les paramètres évoluant dans le temps sont donc estimés sur une grille de temps discrète D . Dans ce qui suit, nous décrivons la procédure d'estimation pour un $\tau \in D$ donné. Par souci de simplicité, nous introduisons les notations suivantes : $\mathbf{B}_\tau = \mathbf{A}(\tau)^{-1}$, $\theta_{i,\tau} = \log_q(\gamma'_i(\tau))$ et $\boldsymbol{\theta}_\tau = (\theta_{1,\tau} \cdots \theta_{N,\tau})^T$.

3.1 Cadre probabiliste

Dorénavant, les X_i sont supposés gaussiens. Il découle de cette hypothèse que la transformée en ondelettes de la i^e source Y_i est une matrice aléatoire gaussienne, complexe circulaire (cf. la Proposition 3 dans [10]). Alors $\mathbf{w}_{y_i,\tau} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_i(\theta_{i,\tau}))$, où

$$[\boldsymbol{\Sigma}_i(\theta_{i,\tau})]_{kk'} = q^{\frac{s_k + s_{k'}}{2}} \int_{\mathbb{R}} \mathcal{S}_{X_i}(q^{-\theta_{i,\tau}} \xi) \overline{\psi}(q^{s_k} \xi) \psi(q^{s_{k'}} \xi) d\xi.$$

Soit p_V , de manière générique, la densité de probabilité d'une variable aléatoire V . Alors, l'hypothèse d'indépendance des sources conduit à l'opposée de la log-vraisemblance suivante :

$$\begin{aligned} \ell_\tau(\mathbf{B}_\tau, \boldsymbol{\theta}_\tau) &\triangleq -\log(p_{\mathbf{w}_{z,\tau} | (\mathbf{B}_\tau, \boldsymbol{\theta}_\tau)}(\mathbf{w}_{z,\tau}; \mathbf{B}_\tau, \boldsymbol{\theta}_\tau)) + c \\ &= -M_s \log |\det(\mathbf{B}_\tau)| + \frac{1}{2} \sum_{i=1}^N \log |\det \boldsymbol{\Sigma}_i(\theta_{i,\tau})| \\ &\quad + \frac{1}{2} \sum_{i=1}^N [\mathbf{B}_\tau \mathbf{w}_{z,\tau}]_i \cdot \boldsymbol{\Sigma}_i(\theta_{i,\tau})^{-1} [\mathbf{B}_\tau \mathbf{w}_{z,\tau}]_i^H, \end{aligned}$$

où $[\mathbf{M}]_i$ indique la i^e ligne de la matrice \mathbf{M} , et \mathbf{M}^H sa transposée conjuguée. Les estimations du maximum de vraisemblance (MV), c'est-à-dire les minimiseurs de $\ell_\tau(\mathbf{B}_\tau, \boldsymbol{\theta}_\tau)$, peuvent être évalués numériquement pour chaque τ , ils sont notés respectivement $\tilde{\mathbf{B}}_\tau$ et $\tilde{\boldsymbol{\theta}}_\tau$.

3.2 Algorithme d'estimation

La stratégie d'estimation est d'alterner les estimations de \mathbf{B}_τ , $\boldsymbol{\theta}_\tau$ et des spectres. L'algorithme 1 (nommé JEFAS-BSS) synthétise toutes les étapes d'estimation. En ce qui concerne l'initialisation de la BSS, nous utilisons le résultat de p-SOBI qui fournit une méthode de base donnant une matrice de mélange initiale variant dans le temps. L'algorithme JEFAS (qui est détaillé dans [1]) permet l'estimation des déformations temporelles et des spectres. Remarquons également que l'estimation de \mathbf{B}_τ s'effectue avec un pas temporel de Δ_τ . La séparation s'effectue alors en considérant \mathbf{B}_τ constante sur l'intervalle $I_\tau = [\tau - \Delta_\tau/2, \tau + \Delta_\tau/2]$.

Enfin, la convergence est contrôlée en utilisant le rapport source à interférence (SIR) introduit dans [12]. Pour une source estimée donnée \tilde{y}_i , le SIR, noté $\text{SIR}(\tilde{y}_i, \mathbf{y})$, quantifie la présence d'interférences dans \tilde{y}_i provenant des autres sources y_j , $j \neq i$. Moins il y a d'interférences, plus le SIR est grand. Comme nous n'avons pas accès aux véritables sources, nous utilisons comme critère d'arrêt le SIR entre $\tilde{\mathbf{y}}^{(k-1)}$ et $\tilde{\mathbf{y}}^{(k)}$ (au lieu de \mathbf{y}) qui donne une évaluation de la mise à jour de la BSS et constitue donc une évaluation de convergence pertinente. Pratiquement, le critère de convergence est testé sur le SIR moyenné entre les sources. Il est défini comme suit :

$$\frac{1}{N} \sum_{i=1}^N \text{SIR}(\tilde{y}_i^{(k-1)}, \tilde{\mathbf{y}}^{(k)}) > \Lambda, \quad (8)$$

Algorithme 1 JEFAS-BSS

Initialisation : Evaluer $\tilde{\mathbf{B}}_\tau^{(0)}$ au moyen de p-SOBI. Calculer les sources estimées $\tilde{\mathbf{y}}^{(0)}(\tau) = \tilde{\mathbf{B}}_\tau^{(0)} \mathbf{z}(\tau)$.

• $k \leftarrow 1$

tant que le critère d'arrêt (8) est faux **et** $k \leq k_{max}$ **faire**

• Pour $i = 1, \dots, N$, estimer les paramètres $\tilde{\theta}_{i,\tau}^{(k)}$, $\forall \tau \in D$ et le spectre $\tilde{\mathcal{S}}_{X_i}^{(k)}$ en appliquant JEFAS à $\tilde{y}_i^{(k-1)}$.

pour $\tau = 0, \Delta_\tau, \dots, T$ **faire**

• Estimer $\tilde{\mathbf{B}}_\tau^{(k)}$ par MV en remplaçant $\boldsymbol{\theta}_\tau$ et \mathcal{S}_X par leurs estimations actuelles $\tilde{\boldsymbol{\theta}}_\tau^{(k)}$ et $\{\tilde{\mathcal{S}}_{X_i}^{(k)}\}_{i=1,\dots,N}$.

fin pour

• Estimer les sources $\tilde{\mathbf{y}}^{(k)}(\tau) = \tilde{\mathbf{B}}_\tau^{(k)} \mathbf{z}(\tau)$.

• $k \leftarrow k + 1$

fin tant que

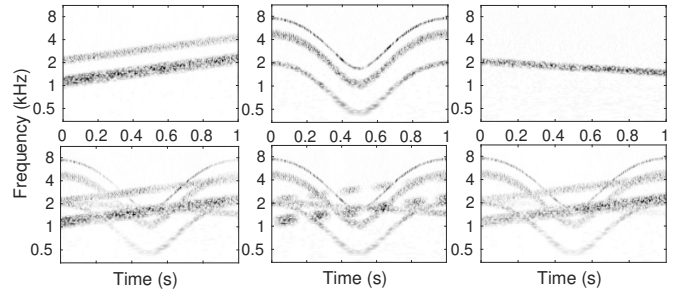


FIGURE 1 – En haut : scalogrammes des trois sources non stationnaires. En bas : scalogrammes des trois observations.

où Λ est choisi suffisamment grand pour que l'amélioration de la qualité de la BSS entre la $(k-1)^e$ et la k^e itération de JEFAS-BSS ne soit pas significative.

4 Résultats

Nous construisons un exemple synthétique avec $N = 3$ pour évaluer les performances de l'algorithme JEFAS-BSS. La durée des signaux est de 1 seconde et ils sont échantillonnés à $F_s = 44,1$ kHz. Les deux sources gaussiennes sont non stationnaires selon le modèle (1), leurs spectres de puissance sous-jacents \mathcal{S}_{X_i} sont constitués de différentes fenêtres de Hann ne se chevauchant pas. Les coefficients de la matrice de mélange évoluent sinusoidalement dans le temps (avec des fréquences différentes). Les transformées en ondelettes des sources et des observations sont affichées sur la figure 1. Leurs supports étant superposés, un algorithme de BSS s'appuyant sur la parcimonie d'une représentation des signaux ne serait pas efficace ici.

JEFAS-BSS est appliqué aux observations, il converge en 5 itérations. Les scalogrammes des sources estimées sont affichés sur la figure 2. Par manque de place, les fonctions de déformation temporelle et les spectres estimés par JEFAS ne sont pas tracés ici. Pour une étude complète des performances de JEFAS, nous renvoyons le lecteur à [1]. Dans ce qui suit, nous nous concentrons sur les performances de la BSS.

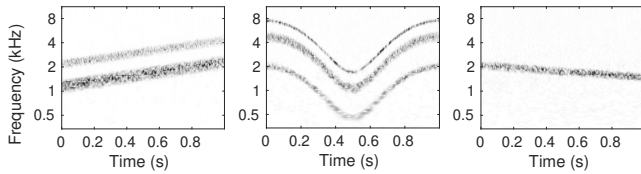


FIGURE 2 – Scalogrammes des trois sources estimées.

TABLE 1 – Comparaison du SIR et de l’indice d’Amari moyen pour quatre algorithmes de BSS.

| Algorithme | SIR (dB) | | ρ moyen (dB) | |
|------------|----------|-------|-------------------|-------|
| | Moyenne | ET | Moyenne | ET |
| SOBI | 12, 14 | 3, 59 | -6, 62 | 0, 82 |
| p-SOBI | 3, 56 | 1, 75 | -9, 03 | 0, 23 |
| QTF-BSS | 0, 69 | 3, 94 | -3, 42 | 0, 42 |
| JEFAS-BSS | 30, 82 | 0, 63 | -15, 47 | 0, 58 |

Afin d’évaluer la BSS obtenue via JEFAS-BSS, nous la comparons à d’autres algorithmes de BSS (SOBI, p-SOBI, QTF-BSS). Nous appliquons ces algorithmes à 20 réalisations de l’exemple synthétique ci-dessus. Pour évaluer la qualité des algorithmes BSS, nous calculons le SIR entre les vraies sources et leurs estimations en moyenne sur 20 simulations. Les résultats sont donnés dans le tableau 1 ainsi que les écarts types (ET) correspondants. Néanmoins, cette quantité est un indicateur global de la qualité de la BSS. Afin de suivre l’évolution temporelle de la qualité de la BSS, nous introduisons l’indice d’interférence inter-symboles normalisé ρ (introduit dans [13] pour l’indice non normalisé). Cet indice est d’autant plus faible que la matrice $\tilde{\mathbf{B}}(t)\mathbf{A}(t)$ est proche de l’identité. De plus, on a toujours $\rho(t) \in [0, 1]$. Par manque de place, l’évolution de l’indice d’Amari (en décibels) pour les algorithmes BSS que nous évaluons n’est pas tracé. Cependant, les moyennes et écarts-types temporels de $\rho(t)$ sont mesurés. Les valeurs moyennes obtenues sur 20 simulations sont données dans le tableau 1.

Comme on pouvait s’y attendre, l’indice d’Amari et le SIR montrent clairement que JEFAS-BSS offre une meilleure BSS que les autres algorithmes. p-SOBI permet d’améliorer l’indice d’Amari par rapport à SOBI mais le SIR est altéré en raison des effets de bords dus à la segmentation de la BSS. Plus surprenant, les performances QTF-BSS sont inférieures à celles de SOBI. Cela peut être dû au fait que les estimations QTF-BSS et SOBI reposent sur un modèle matriciel de mélange constant bien que QTF-BSS soit adapté à la BSS de signaux non stationnaires, contrairement à SOBI.

5 Conclusion

Dans cet article, nous avons présenté JEFAS-BSS, un algorithme pour la BSS de mélanges instantanés non stationnaires d’une classe de signaux non stationnaires. JEFAS-BSS a été évalué sur un exemple synthétique sur lequel il surpasse nettement les méthodes de BSS existantes. JEFAS-BSS a également appliqué à un mélange de signaux audios réels. Les résultats numériques de cet exemple sont disponibles en ligne ¹.

Références

- [1] A. MEYNARD et B. TORRÉSANI, « Spectral Analysis for Nonstationary Audio », *IEEE/ACM Transactions on Audio, Speech and Language Processing*, vol. 26, p. 2371 – 2380, déc. 2018.
- [2] P. COMON et C. JUTTEN, *Handbook of Blind Source Separation, Independent Component Analysis and Applications*. Academic Press (Elsevier), fév. 2010.
- [3] A. BELOUCHRANI, K. ABED-MERAÏM, J.-F. CARDOSO et E. MOULINES, « A blind source separation technique using second-order statistics », *IEEE Transactions on Signal Processing*, vol. 45, p. 434–444, fév. 1997.
- [4] N. THIRION-MOREAU et M. AMIN, « Chapter 11 - quadratic time-frequency domain methods », in *Handbook of Blind Source Separation* (P. COMON et C. JUTTEN, édés), p. 421 – 466, Oxford : Academic Press, 2010.
- [5] A. BELOUCHRANI, M. AMIN, N. THIRION-MOREAU et Y. D. ZHANG, « Source separation and localization using time-frequency distributions : An overview », *IEEE Signal Processing Magazine*, vol. 30, p. 97–107, nov. 2013.
- [6] D.-T. PHAM et J.-F. CARDOSO, « Blind separation of instantaneous mixtures of nonstationary sources », *IEEE Transactions on Signal Processing*, vol. 49, p. 1837–1848, sept. 2001.
- [7] R. KAFTORY et Y. Y. ZEEVI, « Blind separation of time/position varying mixtures », *IEEE Transactions on Image Processing*, vol. 22, p. 104–118, jan. 2013.
- [8] R. E. PRIETO et P. JINACHITRA, « Blind source separation for time-variant mixing systems using piecewise linear approximations », in *Proceedings. IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, vol. 5, p. v/301–v/304, mars 2005.
- [9] M. CLERC et S. MALLAT, « Estimating deformations of stationary processes », *Ann. Statist.*, vol. 31, p. 1772–1821, 12 2003.
- [10] H. OMER et B. TORRÉSANI, « Time-frequency and time-scale analysis of deformed stationary processes, with application to non-stationary sound modeling », *Applied and Computational Harmonic Analysis*, vol. 43, no. 1, p. 1 – 22, 2017.
- [11] D. STOWELL, *Computational Analysis of Sound Scenes and Events*, chap. Computational Bioacoustic Scene Analysis, p. 303–333. Springer, 2018.
- [12] E. VINCENT, R. GRIBONVAL et C. FÉVOTTE, « Performance measurement in blind audio source separation », *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, p. 1462–1469, juil. 2006.
- [13] E. MOREAU et O. MACCHI, « A one stage self-adaptive algorithm for source separation », in *Proceedings of ICASSP ’94. IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. iii, p. III/49–III/52, avril 1994.