

**Séparation de sources dans le cas sous-déterminé :  
comparaison de deux approches basées sur des  
décompositions parcimonieuses**

Sylvain Lesage, Sacha Krstulovic, Rémi Gribonval

► **To cite this version:**

Sylvain Lesage, Sacha Krstulovic, Rémi Gribonval. Séparation de sources dans le cas sous-déterminé : comparaison de deux approches basées sur des décompositions parcimonieuses. XXe colloque GRETSI (traitement du signal et des images), 6-9 septembre 2005, Sep 2005, Louvain-la-Neuve, Belgique. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images, 2005. <inria-00564755>

**HAL Id: inria-00564755**

**<https://hal.inria.fr/inria-00564755>**

Submitted on 9 Feb 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Séparation de sources dans le cas sous-déterminé : comparaison de deux approches basées sur des décompositions parcimonieuses

Sylvain LESAGE, Sacha KRSTULOVIC, Rémi GRIBONVAL

Projet METISS, IRISA-INRIA  
Campus de Beaulieu, 35042 Rennes Cedex, France

sylvain.lesage@irisa.fr, sacha.krstulovic@irisa.fr, remi.gribonval@irisa.fr

**Résumé** – Cet article traite de la séparation de sources dans le cas sous-déterminé quand la matrice de mélange est connue. On se place dans le cadre des approches basées sur la décomposition parcimonieuse du mélange. Dans la nouvelle méthode proposée, on décompose le mélange par Matching Pursuit en introduisant une nouvelle classe de dictionnaires multi-canal, où les atomes sont caractérisés par une direction spatiale et une forme d’onde. La connaissance de la matrice de mélange est directement intégrée à la décomposition. Par rapport à la séparation par Matching Pursuit multi-canal suivie d’un clustering, le nouvel algorithme introduit moins d’artéfacts pour un niveau d’interférences résiduelles comparable. Nous comparons ces deux méthodes à l’algorithme de Bofill-Zibulevski et à DUET. Nous étudions également l’apport d’un lissage des décompositions et l’importance de la qualité de l’estimation de la matrice de mélange.

**Abstract** – This paper focuses on underdetermined source separation with a known mixing matrix. The approach is based on a sparse decomposition of the mixture. In the proposed method, the mixture is decomposed with Matching Pursuit by introducing a new class of multichannel dictionaries, where the atoms are given by a spatial direction and a waveform. The knowledge of the mixing matrix is directly integrated in the decomposition. Compared to the separation by multichannel Matching Pursuit followed by a clustering, the new algorithm introduces less artifacts whereas the level of residual interferences is about the same. These two methods are compared to Bofill-Zibulevsky’s algorithm and DUET. We also study the effect of smoothing the decompositions and the importance of the quality of the estimation of the mixing matrix.

## 1 Introduction

Le problème de la séparation de sources [1] consiste à retrouver des signaux inconnus (les sources) à partir de la seule connaissance de mélanges de ces signaux (les canaux). Chaque canal  $x_n$  est la somme des sources filtrées :

$$x_n(t) = \sum_{i=1}^I (a_{n,i} * s_i)(t)$$

avec  $a_{n,i}$  des filtres. Dans le cas d’un mélange linéaire instantané, les filtres correspondent simplement à une multiplication par une constante, si bien qu’en notation matricielle le mélange s’écrit  $\mathbf{x} = \mathbf{A}\mathbf{s}$ , où  $\mathbf{A}$  est la matrice de mélange et les matrices  $\mathbf{x}$  et  $\mathbf{s}$  ont pour lignes les signaux  $x_n$  et  $s_i$ . Dans le cas déterminé (resp. surdéterminé), lorsque le nombre de canaux observés est égal (resp. supérieur) à celui des sources, il est équivalent d’estimer la matrice de mélange et d’estimer les sources. Dans le cas sous-déterminé, au contraire, la connaissance ou l’estimation de la matrice de mélange n’est pas suffisante pour retrouver les sources et il faut généralement recourir à un modèle des sources pour les estimer [2]. En général, dans les performances d’un algorithme donné, il n’est pas facile de distinguer l’effet de la qualité de l’estimation de la matrice de mélange de celui de l’inadéquation du modèle choisi.

Dans cet article, nous nous intéressons au cas sous-déterminé. Notre approche emploie des modèles reposant sur l’existence de représentations parcimonieuses des sources [3], et suppose que la matrice de mélange est parfaitement connue. Nous comparons deux algorithmes de séparation basés sur des variantes de Matching Pursuit (MP) [4]. La première variante consiste à décomposer le mélange multi-canal sans connaissance de la matrice de mélange, puis à utiliser la matrice de

mélange pour classifier les composantes obtenues et les affecter aux sources à estimer [5, 6]. La seconde variante consiste à utiliser la matrice de mélange dès l’étape de décomposition parcimonieuse, sans qu’aucune étape de classification ultérieure ne soit nécessaire. Les performances de ces deux algorithmes sont comparées au meilleur séparateur linéaire (MSL) [7], à l’algorithme de Bofill-Zibulevski (BZ) [8] et à l’algorithme DUET [6].

Cet article est organisé comme suit : dans la section 2, on rappelle la définition générale du Matching Pursuit. Les algorithmes de séparation qui en découlent sont décrits dans la section 3, puis nous détaillons les conditions expérimentales et les résultats obtenus dans la section 4.

## 2 Matching Pursuit

Un signal  $x$  (considéré comme un vecteur de l’espace de Hilbert  $\mathcal{H}$  des signaux à énergie finie) admet une décomposition parcimonieuse dans le dictionnaire  $\mathcal{D} = \{g_k\}$  d’atomes  $g_k$  – ou signaux élémentaires  $g_k$  – s’il peut s’écrire comme une combinaison linéaire  $x = \sum_k c_k g_k$  où peu de coefficients  $\{c_k\}$  sont non négligeables. L’algorithme Matching Pursuit calcule itérativement des approximations parcimonieuses de la forme  $x = \sum_{m=1}^M c_{k_m} g_{k_m} + R^M$  où  $R^M$  est un résidu qui tend vers zéro lorsque le nombre d’itérations  $M$  tend vers l’infini. Son principe est de sélectionner à chaque étape l’atome le plus corrélé avec le résidu, puis de mettre à jour le résidu en ôtant la contribution de cet atome.

Les critères d’arrêt les plus courants sont basés sur le niveau absolu ou relatif de l’énergie du résidu et/ou sur un nombre fixe d’itérations à effectuer. Par ailleurs, le dictionnaire de Gabor est

classiquement utilisé pour décomposer de façon parcimonieuse les signaux audio. Il représente une collection d'atomes temps-fréquence de Gabor de la forme :

$$g_{s,u,\xi}(t) = w \left( \frac{t-u}{s} \right) \cdot \exp(2j\pi\xi(t-u)).$$

Ces atomes sont définis par le choix d'une fenêtre  $w$  d'énergie unitaire (Hanning, Gaussienne, ...), d'un facteur d'échelle  $s$ , d'une localisation temporelle  $u$ , et d'une fréquence  $\xi$ . Un tel dictionnaire permet un calcul rapide des produits scalaires d'un signal avec les atomes par application de FFT fenêtrées.

### 3 Séparation par Matching Pursuit

Pour résoudre des problèmes de séparation de sources dans le cas multi-canal, une méthode à base de décompositions parcimonieuses sur un dictionnaire a été proposée [5, 9]. Elle utilise le Matching Pursuit multi-canal, suivi d'un clustering. Après avoir rappelé le principe de cette méthode, nous proposons une variante où la définition du dictionnaire inclut des connaissances sur la matrice de mélange  $\mathbf{A}$ .

#### 3.1 Matching Pursuit multi-canal

Pour la décomposition parcimonieuse de signaux multi-canal, on utilise un dictionnaire  $\mathcal{D}$  comportant des atomes multi-canal  $\mathbf{g}$ , définis par  $\mathbf{g} = (c_1g, c_2g, \dots, c_Ng)$ , où  $g \in \mathcal{D}$  est un atome mono-canal d'un dictionnaire  $\mathcal{D}$  et où les coefficients  $c_1, \dots, c_N$  satisfont  $\sum_{n=1}^N c_n^2 = 1$ . Le Matching Pursuit multi-canal conduit à une décomposition du type :

$$(x_1, \dots, x_N) = \hat{\mathbf{x}}^M + (R_1^M, \dots, R_N^M).$$

avec  $\hat{\mathbf{x}}^M := \sum_{m=1}^M (c_{1,k_m}g_{k_m}, \dots, c_{N,k_m}g_{k_m})$ . L'algorithme se déroule de la façon suivante :

1. Initialisation :  $M = 1$ ,  $R_n^0 = x_n$ ,  $c_{n,k} = 0$ ,  $\forall n, \forall k$ ;
2. Calcul du produit scalaire entre chaque canal du résidu  $R_n^{M-1}$  et chaque atome  $g_k$  du dictionnaire mono-canal.
3. Sélection de  $k_M = \arg \max_k \sum_{n=1}^N |\langle R_n^{M-1}, g_k \rangle|^2$
4. Pour chaque canal  $n$ , mise à jour du résidu :  $R_n^M = R_n^{M-1} - \langle R_n^{M-1}, g_{k_M} \rangle g_{k_M}$  et des coefficients :  $c_{n,k_M}^M = c_{n,k_M}^{M-1} + \langle R_n^{M-1}, g_{k_M} \rangle$
5. Si le critère d'arrêt n'est pas atteint,  $M \leftarrow M + 1$ , puis retour à 2

A partir du signal  $\hat{\mathbf{x}}^M$ , approché par Matching Pursuit multi-canal, chaque source  $s_i$  est estimée en utilisant les atomes de la décomposition qui lui sont attribués de la manière suivante : la matrice de mélange  $\mathbf{A}$  étant supposée connue et ses colonnes  $\mathbf{a}_i = (a_{n,i})_{1 \leq n \leq N}$  normées ( $\|\mathbf{a}_i\|_2 = 1$ ), on attribue l'atome  $k_M$  à la source d'indice

$$\hat{i}_M = \arg \max_i |\langle \mathbf{c}_{k_M}, \mathbf{a}_i \rangle|.$$

Cela revient à partitionner l'espace des coefficients  $\mathbf{c} = (c_n)_{1 \leq n \leq N}$  en  $I$  sous-ensembles correspondants aux colonnes  $\mathbf{a}_i$  de  $\mathbf{A}$  ( $I$  étant le nombre de sources). La source  $s_i$  est alors reconstituée par :

$$\hat{s}_i = \sum_{M \text{ tq } \hat{i}_M = i} \langle \mathbf{c}_{k_M}, \mathbf{a}_i \rangle g_{k_M}. \quad (1)$$

On nomme MPC1 cet algorithme, et MPC2 une variante qui consiste à attribuer chaque atome aux deux sources qui l'entourent. Cette seconde sélection, également utilisée dans l'algorithme BZ [8], correspond à la minimisation de la norme  $l_1$  de la projection des coefficients  $\mathbf{c}_{k_M}$  sur deux directions de la matrice de mélange :

$$(\hat{i}_M, \hat{j}_M) = \arg \min_{(i,j)} \|[\mathbf{a}_i \ \mathbf{a}_j]^{-1} \mathbf{c}_{k_M}\|_1.$$

#### 3.2 Dictionnaire multi-canal directionnel

En combinant l'expression des mélanges  $x_n = \sum_{i=1}^I a_{n,i} \cdot s_i$  et celle d'une possible décomposition parcimonieuse  $s_i = \sum_{k=1}^K c_{i,k} g_k$  de chaque source  $s_i$  sur le dictionnaire mono-canal  $\mathcal{D}$ , on peut écrire  $x_n = \sum_{i,k} a_{n,i} c_{i,k} g_k$ . Cela se traduit matriciellement par  $\mathbf{x} = \mathbf{A} \mathbf{C} \Phi^T$ , avec  $\Phi^T$  la matrice dont les lignes sont les atomes mono-canal  $g_k$ , et  $\mathbf{C} = \{c_{i,k}\}_{i,k}$  une matrice de coefficients parcimonieux. Cette décomposition peut aussi s'écrire  $\mathbf{x} = \sum_{i,k} c_{i,k} \mathbf{a}_i g_k$ , c'est-à-dire que  $\mathbf{x}$  admet une décomposition parcimonieuse sur le dictionnaire multi-canal "directionnel" constitué des atomes  $\mathbf{a}_i g_k = (a_{1,i} g_k, \dots, a_{N,i} g_k)$ . Le produit scalaire se calcule alors par :

$$\langle \mathbf{R}^M, \mathbf{a}_i g_k \rangle = \mathbf{a}_i^T \mathbf{R}^M g_k^T.$$

et la reconstruction de la source  $s_i$  se fait par :

$$\hat{s}_i = \sum_k c_{i,k} g_k. \quad (2)$$

Ce nouvel algorithme est appelé MPD et ses propriétés théoriques ont été étudiées dans [10]. L'utilisation d'un dictionnaire directionnel revient à appliquer Matching Pursuit multi-canal en contraignant les coefficients  $\mathbf{c}_{k_M}$  de la section 3.1 à être proportionnels à une colonne  $\mathbf{a}_i$  de  $\mathbf{A}$ .

## 4 Expériences

Nous comparons MPC1, MPC2 et MPD avec les algorithmes de référence MSL, BZ et DUET. Les expériences portent sur le mélange stéréo linéaire instantané de trois sources musicales. La première est un enregistrement de violoncelle, la deuxième de percussions, et la troisième de piano. La fréquence d'échantillonnage des signaux est de 8000 Hz pour une durée de 2.4s, soit 19200 échantillons. La matrice de mélange est la suivante :

$$\begin{bmatrix} \cos(\pi/8) & \cos(\pi/4) & \cos(3\pi/8) \\ \sin(\pi/8) & \sin(\pi/4) & \sin(3\pi/8) \end{bmatrix}$$

Les sources  $y$  sont séparées d'un angle de  $\pi/8$ . L'énergie des percussions, situées au milieu, est environ deux fois plus faible que les énergies du piano et du violoncelle qui sont du même ordre.

Nous utilisons les mesures de performance de séparation proposées dans [7], qui permettent d'analyser finement l'origine des distorsions distinguant la source estimée de la source réelle. Ces mesures, exprimées en décibels, sont le rapport entre le signal et a) la distorsion relative totale (SDR), b) celle due aux interférences (SIR) et c) celle due aux artefacts de calcul (SAR). Pour ces trois mesures, de la même famille que le classique rapport signal à bruit, les performances sont d'autant meilleures que le rapport est élevé.

## 4.1 Algorithmes de référence

Les performances des algorithmes basés sur Matching Pursuit sont comparées à celles de trois algorithmes : le meilleur séparateur linéaire (MSL) [7], DUET [6] et l’algorithme de Bofill-Zibulevski (BZ) [8].

Le premier consiste uniquement en l’application d’une matrice  $\mathbf{B}$  au signal.  $\mathbf{B}$  est telle que les sources estimées  $\hat{\mathbf{s}} = \mathbf{B}\mathbf{x}$  minimisent la distorsion due aux interférences [7]. Si on suppose les sources mutuellement orthogonales, et si on connaît la matrice de mélange  $\mathbf{A}$  et la matrice diagonale  $\mathbf{D}$  des normes des sources, alors en notant  $\hat{\mathbf{A}} = \mathbf{A}\mathbf{D}$ , la matrice  $\mathbf{B}$  est donnée par  $\hat{\mathbf{A}} : \mathbf{B} = \mathbf{D}\hat{\mathbf{A}}^H(\hat{\mathbf{A}}\hat{\mathbf{A}}^H)^{-1}$ .

L’algorithme DUET [6] applique une transformée de Fourier à court terme (TFCT) à chaque canal du signal, puis applique un masquage en supposant une seule source active pour chaque “case” temps-fréquence, et enfin inverse la TFCT pour construire la source estimée.

L’algorithme de Bofill-Zibulevski [8] repose sur le même principe que DUET, à la différence près que chaque case temps-fréquence est attribuée aux deux sources les plus proches. Cette attribution est déterminée par une minimisation de norme  $l_1$ .

Dans toutes les expériences, ces deux algorithmes sont appliqués avec une fenêtre de Hanning, de taille 4096 échantillons, avec un décalage de 2048 échantillons (50% de la taille de la fenêtre). Les performances de DUET et BZ sont fortement dépendantes du dimensionnement de la fenêtre, et nous avons observé qu’une taille plus grande et surtout plus petite les dégrade fortement dans les cas étudiés.

## 4.2 Variantes des algorithmes MP

Dans cette expérience, on étudie l’influence du nombre d’itérations, de la composition du dictionnaire, de l’exploitation du résidu et d’un post-traitement par lissage dans le cadre de l’algorithme MPD. Deux dictionnaires peuvent être utilisés pour la décomposition :

- un “petit” dictionnaire constitué d’atomes de Gabor de taille  $s = 4096$  qui se recouvrent à moitié ( $u = ns/2$ ,  $n \in \mathbb{N}$ ). Il correspond à la TFCT utilisée par les algorithmes DUET et BZ.
- un “grand” dictionnaire constitué d’atomes de Gabor de taille  $s = 64$  à 16384 (par puissance de deux). Le décalage entre atomes est également de 50% de la taille d’un atome.

La figure 1 représente les rapports SDR, SIR et SAR de la source “piano” estimée par les différents algorithmes, en fonction du nombre d’itérations (les résultats étant similaires pour les deux autres sources).

Premièrement, on remarque que tout au long des itérations l’utilisation du grand dictionnaire conduit à une meilleure séparation que le petit dictionnaire. En effet, dans le cas du grand dictionnaire, l’algorithme MP choisit lui-même la taille de fenêtre optimale. Ainsi la nécessité d’optimiser la taille de fenêtre *a priori* est supprimée, contrairement au cas des algorithmes BZ et DUET.

Par ailleurs, on peut remarquer que l’amélioration des performances est monotone lorsque le nombre d’itérations augmente. Dans le détail, les artefacts, qui dominent la distorsion, sont importants lorsque les sources sont reconstruites avec peu

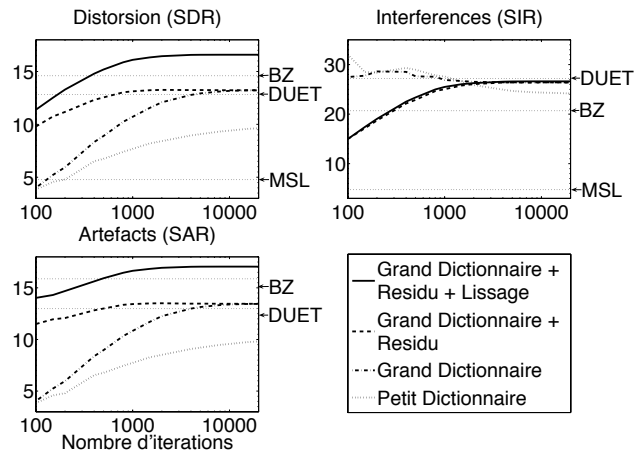


FIG. 1 – Distorsions (dB), source “piano” estimée par MPD

d’atomes, et ils diminuent au fil des itérations grâce à l’apport de nouveaux atomes. Après suffisamment d’itérations, MPD devient meilleur que DUET sur le plan des taux d’artefacts et de distorsion.

Pour compenser la distorsion due au faible nombre d’atomes, on peut séparer le résidu de la décomposition  $\mathbf{R}^M$  en utilisant le séparateur linéaire  $\mathbf{A}^H(\mathbf{A}\mathbf{A}^H)^{-1}\mathbf{R}^M$  et l’ajouter aux sources estimées. Le séparateur utilisé suppose que tous les résidus des sources ont la même énergie. Asymptotiquement, cette hypothèse se vérifie, les sources les plus énergétiques étant celles dont les atomes sont les premiers sélectionnés. L’addition du résidu augmente largement la performance globale (SDR) pour un faible nombre d’itérations. Cet effet se décompose en une forte diminution des artefacts, au détriment d’une augmentation des interférences. Pour un plus grand nombre d’itérations, l’énergie du résidu approche de zéro et l’amélioration apportée sur le SDR diminue. La réduction des artefacts permet d’obtenir de meilleures performances globales que DUET à partir d’un nombre réduit d’itérations (moins de 1000).

Néanmoins, les performances de MPD avec séparation du résidu restent inférieures à celles de l’algorithme BZ avec une fenêtre optimisée. Partant de l’hypothèse que le lissage introduit par le recouvrement des fenêtres de la TFCT dans l’algorithme BZ joue un rôle dans sa bonne performance [11], nous avons appliqué un lissage des sources estimées par MPD avec ajout du résidu. Ce lissage consiste à effectuer plusieurs estimations des sources à partir de versions décalées du dictionnaire et à produire la moyenne de ces estimations. L’amélioration apportée par le lissage, très nette sur le plan des artefacts ( $\sim 3$ dB), mais pas systématique pour les interférences, permet d’obtenir de meilleures performances en SDR et SAR que DUET et BZ.

Pour MPC1 et MPC2, le changement de dictionnaire, l’ajout du résidu et le lissage produisent des effets du même type que ceux observés pour MPD. Le caractère itératif de ces algorithmes implique un fort coût de calcul, compensé par la mise en œuvre d’une implémentation rapide de l’algorithme [12].

## 4.3 Comparaison des performances

L’expérience suivante est destinée à évaluer la capacité des différents algorithmes à maintenir une bonne séparation lorsque la matrice de mélange n’est pas connue, mais seulement approchée. Une imprécision volontaire est introduite par

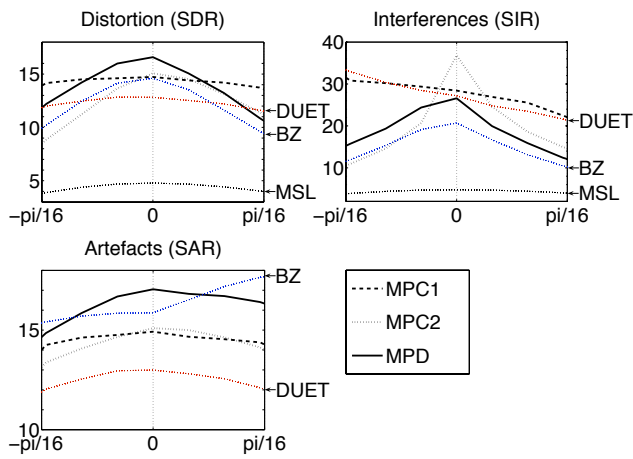


FIG. 2 – Distorsions (dB), source “piano” en fonction de l’angle de perturbation

rotation de la vraie matrice. Les directions des trois sources sont décalées d’un angle égal, qui varie entre  $-\pi/16$  et  $\pi/16$ . Les expériences sont effectuées avec le “grand” dictionnaire. Elles incluent la séparation du résidu et le lissage, et utilisent 5000 itérations. Les performances sont données sur la figure 2, en fonction de l’angle de perturbation, pour le piano.

*Évolution du SAR* – Les méthodes étudiées gardent un niveau d’artéfacts à peu près constant en fonction de la perturbation. MPD et BZ introduisent le moins d’artéfacts, suivies par MPC1 et MPC2 qui présentent des performances équivalentes, elles-mêmes meilleures que DUET. Ces niveaux d’artéfacts sont intrinsèques aux modèles sous-jacents à chaque méthode.

*Évolution du SIR* – Pour les méthodes MPC1 et DUET, les atomes temps-fréquence sont attribués à une seule source. Ainsi, ces méthodes produisent le plus faible niveau d’interférences et restent robustes à la perturbation de la matrice de mélange. Dans le cas de MPC2, MPD et BZ, le fait d’attribuer les atomes temps-fréquence à plusieurs sources introduit une plus grande sensibilité à la perturbation de la matrice de mélange. Pour une matrice de mélange bien estimée, MPC2 produit le moins d’interférences.

*Évolution du SDR* – Par définition, la distorsion globale (SDR) est dominée par le minimum du SAR ou du SIR. Pour une matrice de mélange bien estimée, les méthodes basées sur Matching Pursuit obtiennent de meilleures performances globales que les méthodes de référence. Par ordre décroissant de performance, les méthodes s’échelonnent de la manière suivante : MPD, MPC1, MPC2, BZ, DUET et MSL. En revanche, lorsqu’une perturbation est introduite sur la matrice de mélange, les méthodes MPC1 et DUET (attribution à une seule direction) s’avèrent plus robustes que les méthodes MPC2, BZ (attribution à deux directions, ce qui entraîne une plus grande sensibilité en matière d’interférences) et que MPD (sélection des atomes par Matching Pursuit seulement sur les directions estimées des sources).

## 5 Conclusions

Nous avons comparé plusieurs méthodes de séparation de sources sous-déterminée par décomposition parcimonieuse, en supposant que la matrice de mélange est connue. Dans les algorithmes MPC1 et MPC2, la matrice de mélange est uti-

lisée *a posteriori* pour classer et regrouper les atomes issus de la décomposition par Matching Pursuit. Dans l’algorithme MPD, la connaissance de la matrice de mélange est incluse *a priori* dans la définition du dictionnaire. La version de MPD avec séparation du résidu et lissage donne de meilleures performances, en matière de distorsion globale et d’artéfacts, que les méthodes DUET et BZ. Lorsque la matrice de mélange est bien estimée, MPD et MPC2 donnent les meilleurs résultats. En revanche, MPC1 et DUET semblent plus robustes à une erreur d’estimation de la matrice de mélange.

Le formalisme proposé permet d’envisager l’estimation conjointe de la matrice de mélange et des sources, et par ailleurs l’apprentissage de dictionnaire multi-canal directionnel à partir du mélange.

## Références

- [1] J.-F. Cardoso, “Blind signal separation : statistical principles,” *Proc. IEEE, Special issue on blind identification and estimation*, vol. 9, no. 10, pp. 2009–2025, Oct. 1998.
- [2] O. Bermond and J.-F. Cardoso, “Méthodes de séparation de sources dans le cas sous-déterminé,” in *Proc. GRETSI, Vannes, France, 1999*, pp. 749–752.
- [3] M. Zibulevsky and B. Pearlmutter, “Blind source separation by sparse decomposition in a signal dictionary,” *Neural Computations*, vol. 13, no. 4, pp. 863–882, 2001.
- [4] S. Mallat and Z. Zhang, “Matching pursuit with time-frequency dictionaries,” *IEEE Trans. on Signal Processing*, vol. 41, pp. 3397–3415, Dec. 1993.
- [5] R. Gribonval, “Sparse decomposition of stereo signals with matching pursuit and application to blind separation of more than two sources from a stereo mixture,” in *Proc. Int. Conf. Acoust. Speech Signal Process. (ICASSP’02), Orlando, Florida, USA, May 2002, 2002*.
- [6] O. Yilmaz and S. Rickard, “Blind separation of speech mixtures via time-frequency masking,” *IEEE Transactions on Signal Processing*, vol. 52, no. 7, pp. 1830–1847, July 2004.
- [7] R. Gribonval, L. Benaroya, E. Vincent, and C. Févotte, “Proposals for performance measurement in source separation,” in *Proc. 4th Int. Symp. on Independent Component Analysis and Blind Signal Separation (ICA2003)*, Nara, Japan, Apr. 2003, pp. 763–768.
- [8] P. Bofill and M. Zibulevsky, “Blind separation of more sources than mixtures using sparsity of their short-time fourier transform,” in *Proc. ICA2000, Helsinki, June 2000*, pp. 87–92.
- [9] R. Gribonval, “Piecewise linear source separation,” in *Proc. SPIE’03 – Wavelets : Applications in Signal and Image Processing*, San Diego, California, USA, vol. 5207, August 2003, pp. 297–310.
- [10] R. Gribonval and M. Nielsen, “Beyond sparsity : recovering structured representations by  $\ell_1$ -minimization and greedy algorithms. – application to the analysis of sparse underdetermined ICA-,” IRISA, Tech. Rep. 1684, Jan. 2005, <http://www.irisa.fr/metiss/gribonval/>.
- [11] S. Araki, S. Makino, H. Sawada, and R. Mukai, “Reducing musical noise by a fine-shift overlap-add method applied to source separation using a time-frequency mask,” in *Proceedings in Acoustics, Speech, and Signal Processing (ICASSP’ 05)*, vol. 3, march 2005, pp. 81–84.
- [12] R. Gribonval and S. Krstulović, “The Matching Pursuit ToolKit”, voir <http://www.irisa.fr/metiss/matchp/>.