

Évaluation des algorithmes adaptatifs de séparation de sources

Introduction

Dans ce chapitre, nous évaluons les performances des algorithmes adaptatifs de séparation de sources présentés dans le chapitre 6. Nous évaluons nos algorithmes adaptatifs avec prétraitement de formation de voies sur des cas de séparation de deux sources où le nombre réel de sources peut être fixe ou variable. Nous proposons les trois scénarios d'évaluation suivant :

1. le nombre de sources à estimer est connu ; nous utilisons l'algorithme `BF_fixed[5°]_BS+BSS- l_1` avec le vrai nombre de sources N ;
2. le nombre de sources à estimer est inconnu ; nous utilisons l'algorithme `BF_fixed[5°]_BS+BSS- l_1` avec un nombre de sources à estimer fixe et égal à $N_{hyp} = 5$;
3. le nombre de source est estimé ; nous utilisons l'algorithme `BF_fixed_NbSrEstim+BSS- l_1` qui nous permet d'estimer le nombre de sources dans l'étape de prétraitement par formation de voies.

Dans la suite, nous détaillerons les paramètres de ces algorithmes de séparation adaptatifs et nous procéderons à une évaluation de leurs performances en utilisant les outils d'évaluation présentés dans la section précédente : `BSS_EVAL` et `PEASS`.

Paramètres des algorithmes de séparation adaptatifs

Les signaux de test ont une durée de 5s pour le cas où le nombre réel de sources est fixe et 15s pour le cas où le nombre réel de sources est variable (*cf.* figure 10.1). Ces

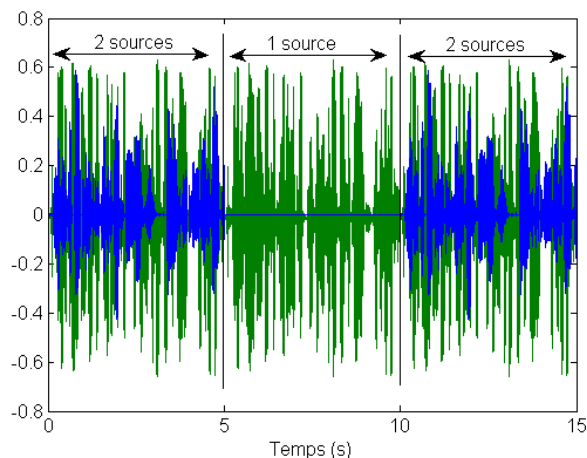


FIGURE 10.1 – Exemple d’un cas de séparation avant le mélange (sources bruts) avec un nombre de sources variant entre 1 et 2.

signaux sont échantillonnés à 16 kHz et sont extraits de la base de données *Theo-RI-studio*. Vingt cas de séparation ont été testés pour chacune des deux configurations. La fenêtre d’analyse spectrale est de type Hanning et de longueur 64 ms (1024 échantillons) et le pas d’avancement est de 50%. Le pas de mise à jour des algorithmes itératifs est $\mu = 0.05$. Nous choisissons un nombre d’itérations par fenêtre d’analyse temporelle longue égal à 2. Les résultats que nous présenterons dans la prochaine section sont la moyenne des résultats obtenus pour les 20 cas de séparation sur chaque fenêtre d’analyse longue.

10.1 Évaluation des algorithmes adaptatifs de séparation sans estimation du nombre de sources

10.1.1 Nombre de sources connu

Nous évaluons le rapport source-à-interférences SIR, le rapport source-à-distorsion SDR et le rapport sources-à-artéfacts SAR de l’algorithme adaptatif de séparation de sources avec un prétraitement par formation de voies et sélection de lobes `BF_fixed[5°]_BS+BSS-l1` pour deux configurations du nombre de sources : le nombre de sources est connu et fixe au cours du temps (*cf.* figures 10.2, 10.3 et 10.4) et le nombre de sources est connu et variable au cours du temps (*cf.* figures 10.5, 10.6 et 10.7). Le SIR et le SDR augmentent avec l’augmentation de l’angle séparant les deux sources et ceci pour toutes les itérations. Les rapports sources-à-artéfacts relatifs aux différentes directions d’arrivées quant à eux restent proches dans un intervalle de

5dB, comme observé dans le cas de la séparation itérative dans les figures 9.26 et 9.30. Pour ce cas de séparation de sources où le nombre de sources est connu et invariable au cours du temps, l'algorithme adaptatif `BF_fixed[5°]_BS+BSS- l_1` présente de bonnes performances de séparation.

Dans le cas où le nombre de sources est variable, quand une seule source est active (typiquement entre les fenêtres d'analyses longues 157 et 299) et que cette source est bien extraite, le rapport source-à-interférences SIR estimé par `BSS_EVAL` est très élevé. Nous l'avons donc limité à 30dB dans la figure 10.5 pour des raisons pratiques. Nous remarquons que les performances de séparation baissent juste avant et après le changement du nombre de sources (*cf.* figures 10.5, 10.6 et 10.7). En effet, la source qui s'évanouit est de moins en moins présente dans la fenêtre d'analyse glissante, et quand seulement quelques échantillons relatifs à cette source sont présents dans cette fenêtre, l'algorithme de séparation a des difficultés à l'extraire.

Pour ce cas de séparation de sources où le nombre de sources est connu et variable au cours du temps, l'algorithme adaptatif `BF_fixed[5°]_BS+BSS- l_1` suit bien le changement du nombre de sources et donne de bonnes performances de séparation.

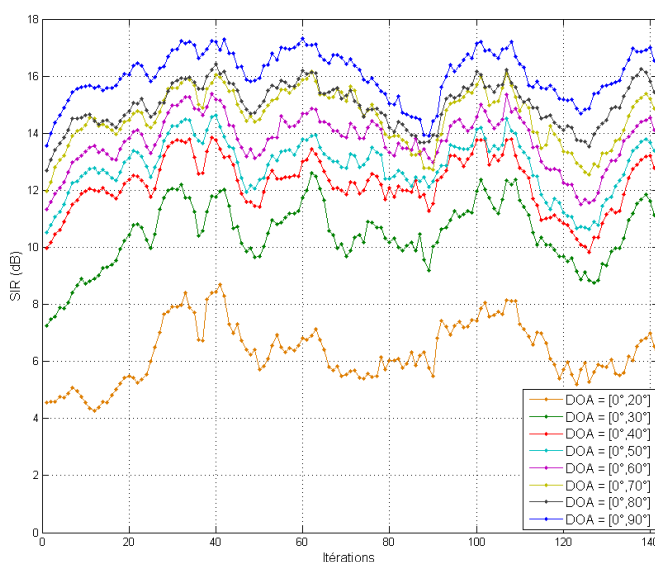


FIGURE 10.2 – Le rapport source-à-interférences SIR de `BF_fixed[5°]_BS+BSS- l_1` au cours des fenêtres d'analyse longues : nombre de sources réel fixe et connu (2 sources).

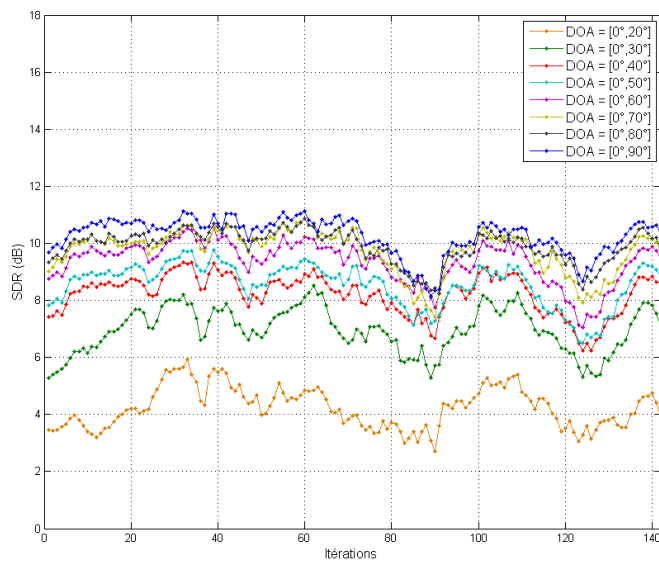


FIGURE 10.3 – Le rapport source-à-distorsion SDR de $\text{BF_fixed}[5^\circ]_BS+BSS-l_1$ au cours des fenêtres d'analyse longue : nombre de sources réel fixe et connu (2 sources).

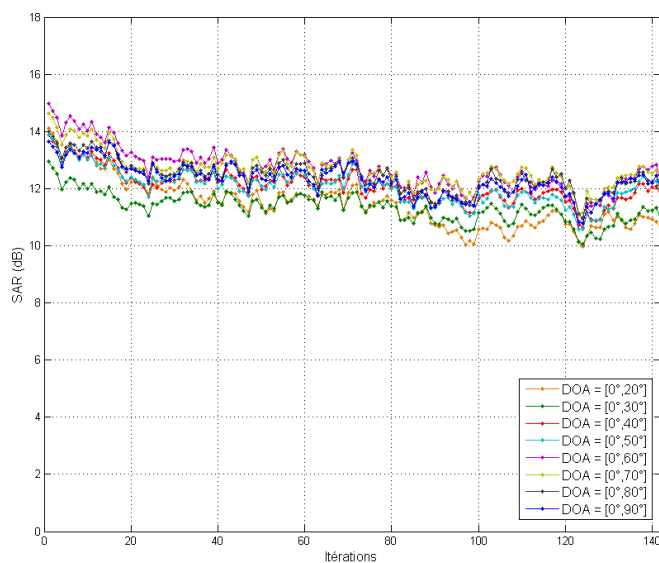


FIGURE 10.4 – Le rapport sources-à-artéfacts SAR de $\text{BF_fixed}[5^\circ]_BS+BSS-l_1$ au cours des fenêtres d'analyse longue : nombre de sources réel fixe et connu (2 sources).

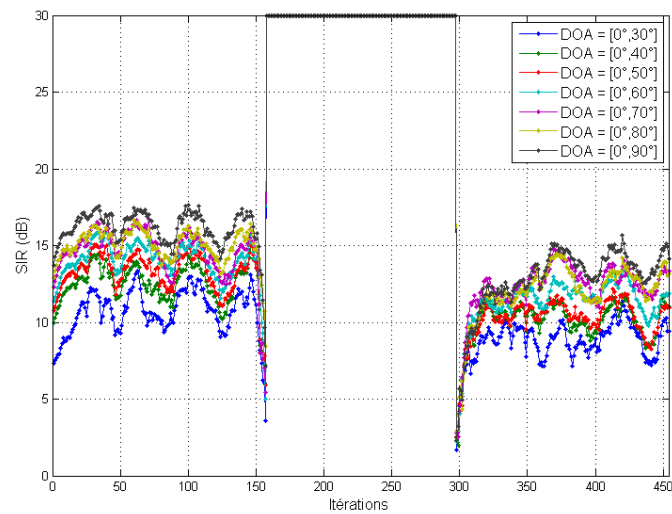


FIGURE 10.5 – Le rapport source-à-interférences SIR de $\text{BF_fixed}[5^\circ]_BS+BSS-l_1$ au cours des fenêtres d'analyse longues : nombre de sources réel connu et variable entre 1 et 2 sources.

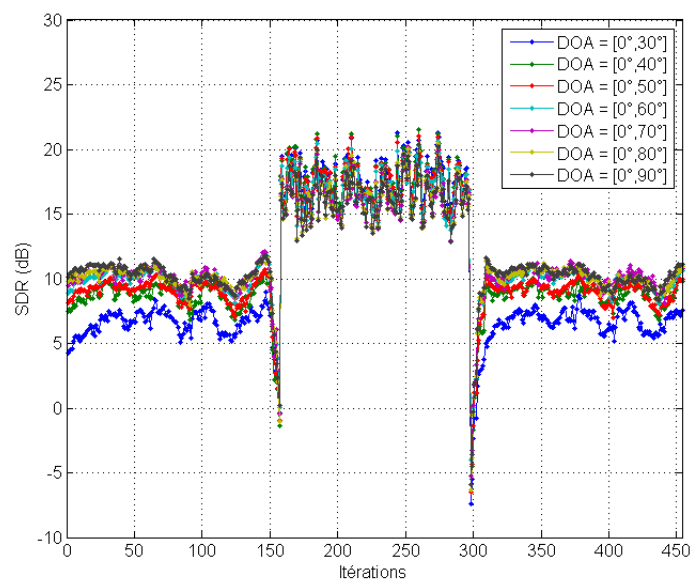


FIGURE 10.6 – Le rapport source-à-distorsion SDR de $\text{BF_fixed}[5^\circ]_BS+BSS-l_1$ au cours des fenêtres d'analyse longues : nombre de sources réel connu et variable entre 1 et 2 sources.

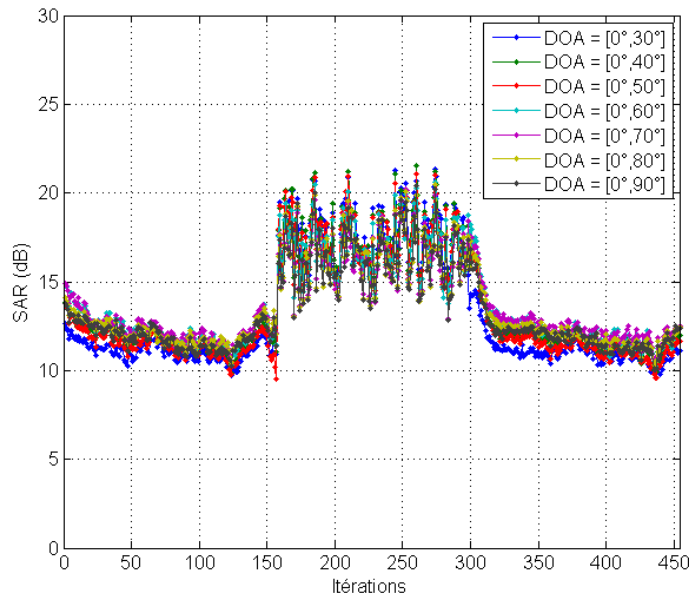


FIGURE 10.7 – Le rapport source-à-artéfacts SAR de $\text{BF_fixed}[5^\circ]_BS+BSS-l_1$ au cours des fenêtres d'analyse longues : nombre de sources réel connu et variable entre 1 et 2 sources.

10.1.2 Nombre de sources fixé *a priori*

Nous évaluons ici le SIR, le SDR et le SAR de l'algorithme de séparation adaptatif avec prétraitement par formation de voies $\text{BF_fixed}[5^\circ]_BS+BSS-l_1$ pour le cas où le nombre de sources est fixe (*cf.* figures 10.8, 10.9 et 10.10) ou variable (*cf.* figures 10.11, 10.12 et 10.13). Pour chacune de ces deux configurations, nous séparons un nombre de sources fixe $N_{hyp} = 5$, supérieur au nombre de sources réel.

Que ce soit pour le cas où le nombre de sources est fixe (*cf.* figures 10.8, 10.9 et 10.10) ou variable (*cf.* figures 10.11, 10.12 et 10.13), les performances de la séparation sont proches de celles obtenues en connaissant le nombre réel de sources.

L'avantage de cette méthode de séparation adaptative est qu'elle ne nécessite pas une estimation du nombre de sources actives. Cependant, le nombre de sources à extraire doit être fixé de telle sorte qu'il soit supérieur au vrai nombre de sources. Ceci suppose une certaine connaissance de l'environnement dans lequel se déroule la séparation de sources et du nombre maximal de sources actives.

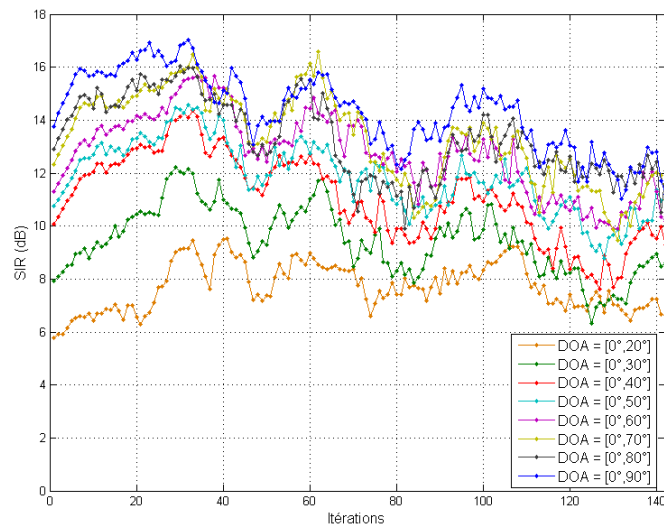


FIGURE 10.8 – Le rapport source-à-interférences SIR de $\text{BF_fixed}[5^\circ]_BS+BSS-l_1$ au cours des fenêtres d'analyse : nombre de sources réel est inconnu et fixé *a priori*.

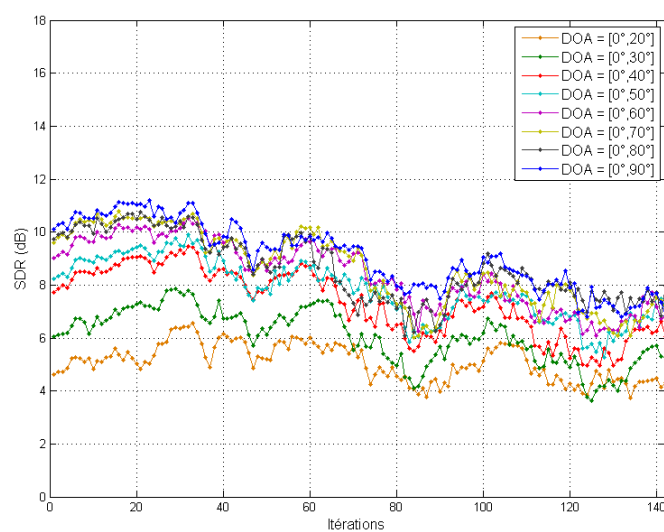


FIGURE 10.9 – Le rapport source-à-distorsion SDR de $\text{BF_fixed}[5^\circ]_BS+BSS-l_1$ au cours des fenêtres d'analyse : nombre de sources réel est inconnu et fixé *a priori*.

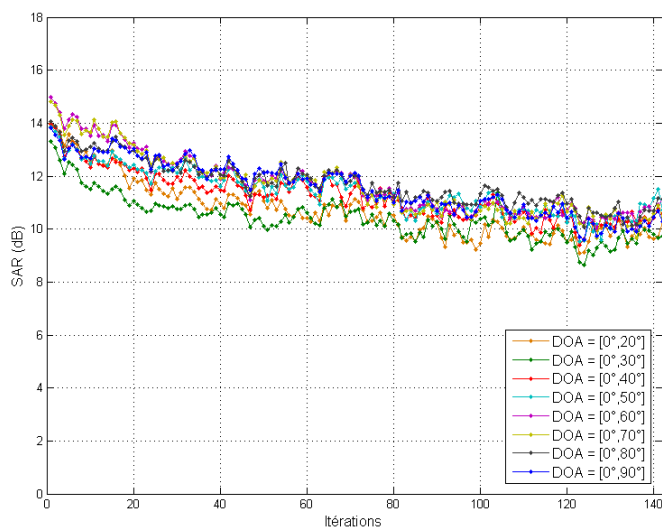


FIGURE 10.10 – Le rapport sources-à-artéfacts SAR de $\text{BF_fixed}[5^\circ]_BS+BSS-l_1$ au cours des fenêtres d'analyse : nombre de sources réel connu et variable entre 1 et 2 sources.

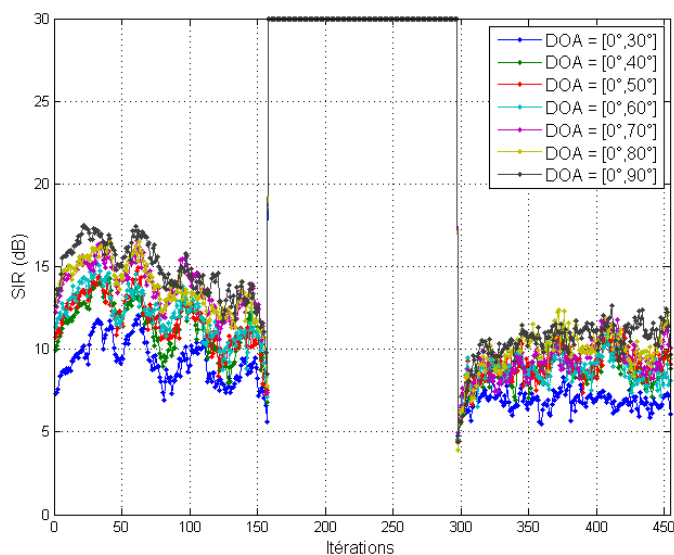


FIGURE 10.11 – Le rapport source-à-interférences SIR de $\text{BF_fixed}[5^\circ]_BS+BSS-l_1$ au cours des fenêtres d'analyse : le nombre de sources réel est inconnu et fixé *a priori*.

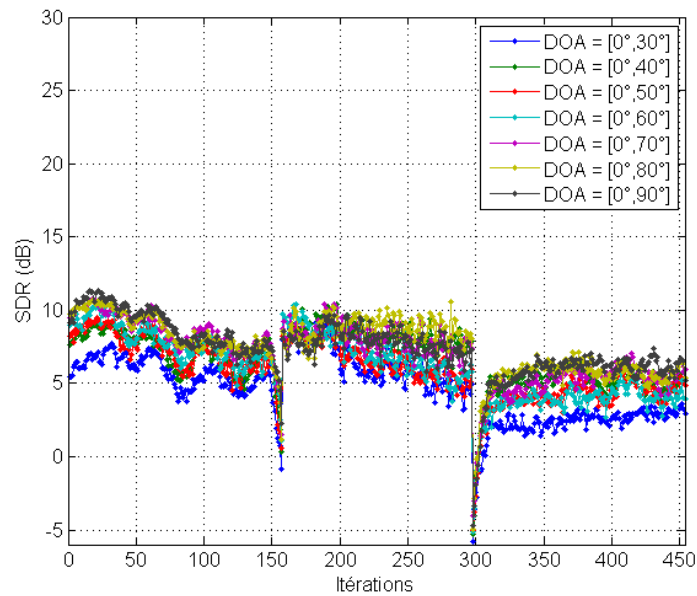


FIGURE 10.12 – Le rapport source-à-distorsion SDR de $\text{BF_fixed}[5^\circ]_BS+BSS-l_1$ au cours des fenêtres d'analyse : le nombre de sources réel est inconnu et fixé *a priori*.

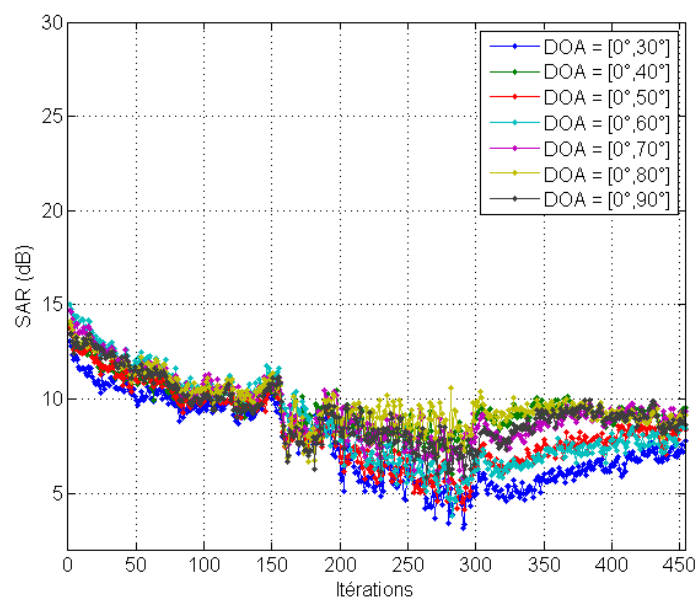


FIGURE 10.13 – Le rapport source-à-artéfacts SAR de $\text{BF_fixed}[5^\circ]_BS+BSS-l_1$ au cours des fenêtres d'analyse : le nombre de sources réel est inconnu et fixé *a priori*.

10.2 Evaluation des algorithmes de séparation avec estimation du nombre de sources

Dans cette section, nous évaluons notre algorithme de séparation en estimant le nombre de sources [2].

Les figures 10.14 et 10.15a montrent le nombre moyen de sources estimé au cours des fenêtres d'analyse longues où le nombre de sources à estimer est fixe et variable respectivement. Ces résultats ont été obtenus avec le prétraitement par formation de voies fixe et sélection des lobes plus grandes énergies en fixant l'angle interlobes à 5° . Nous comparons notre méthode d'estimation du nombre de sources à deux méthodes EIG1 [48] et EIG2 basée sur un simple seuillage des valeurs propres ordonnées des matrices de covariance des signaux reçus dans le domaine temps-fréquence. Nos résultats sont proches de EIG2 mais dans notre cas, l'estimation du nombre de sources est un résultat direct du prétraitement par formation de voies, elle est simple à implémenter et ne demande pas plus de calcul que celui de l'estimation de pics. La méthode EIG2 nécessite plus de temps de traitement que notre méthode à cause du calcul des matrices de covariances et de la décomposition en valeurs propres.

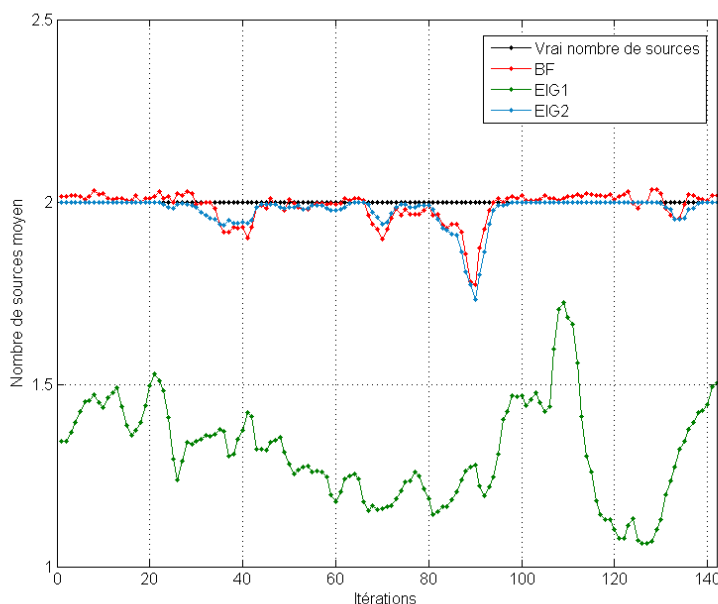


FIGURE 10.14 – Estimation du nombre de source au cours des fenêtres d'analyse avec `BF_fixed[5°]_NbSrEstim` : le nombre réel de sources est fixe.

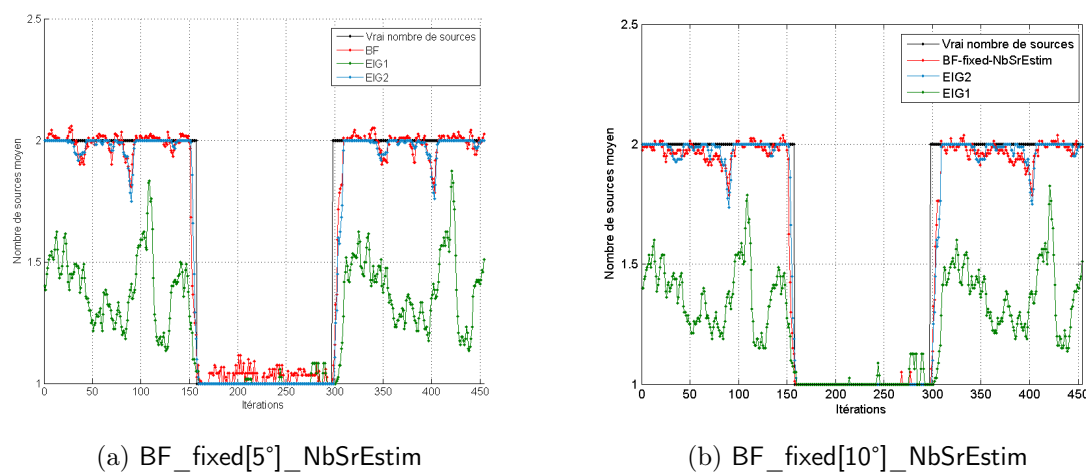


FIGURE 10.15 – Estimation du nombre de source au cours des fenêtres d’analyse avec `BF_fixed_NbSrEstim` : le nombre réel de sources varie entre 1 et 2.

Si dans le prétraitement par formation de voies fixe nous choisissons un angle inter-lobes égal à 10° , nous remarquons qu’il y a moins de sur-estimation du nombre de sources avec la méthode du prétraitement par formation de voies `BF_fixed[10°]_NbSrEstim` par comparaison à `BF_fixed[5°]_NbSrEstim` et ceci dans le cas où une seule source est active (*cf.* figure 10.15b). Rappelons que pour l’estimation du nombre de sources, après le filtrage par formation de voies, nous sélectionnons N_{max} directions de visée qui correspondent aux N_{max} lobes ayant les plus grandes énergies. Ensuite, un histogramme qui correspond au nombre d’occurrence des directions d’arrivées sélectionnées est construit et le nombre de pics dont la valeur est supérieure à un seuil donné correspond au nombre de sources. La résolution angulaire de la formation de voies obtenue avec un angle inter-lobes égal à 5° est plus fine que celle obtenue avec un angle inter-lobes de 10° . Dans le cas où l’angle inter-lobes est égal à 5° , nous voyons donc apparaître des pics relatifs à la réverbération par exemple, ce qui peut conduire à une sur-estimation du nombre de sources actives. Cette sur-estimation du nombre de sources peut être évitée en utilisant le « bon » seuillage.

Les figures 10.16, 10.17 et 10.4 montrent le rapport source-à-interférences SIR, le rapport source-à-distorsion SDR et le rapport sources-à-artéfacts SAR moyens dans les cas de séparation où le nombre de sources ne change pas au cours du temps. Les performances de séparation de l’algorithme adaptatif avec l’estimation du nombre de sources dans ce cas restent proches de celles où le nombre de sources à séparer est connu.

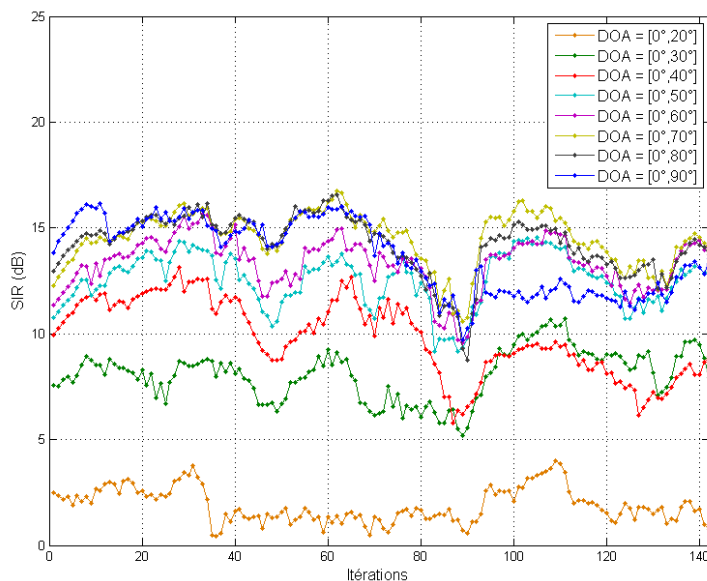


FIGURE 10.16 – Le rapport source-à-interférences SIR de BF_fixed[5°]_NbSrEstim+BSS- l_1 au cours des fenêtres d'analyse : le nombre de sources est estimé (nombre de sources réel égal à 2).

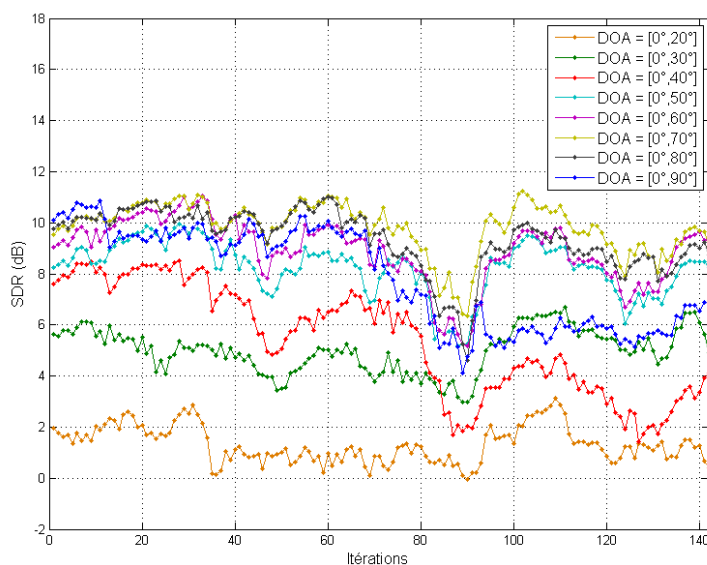


FIGURE 10.17 – Le rapport source-à-distorsion SDR de BF_fixed[5°]_NbSrEstim+BSS- l_1 au cours des fenêtres d'analyse : le nombre de sources est estimé (nombre de sources réel égal à 2).

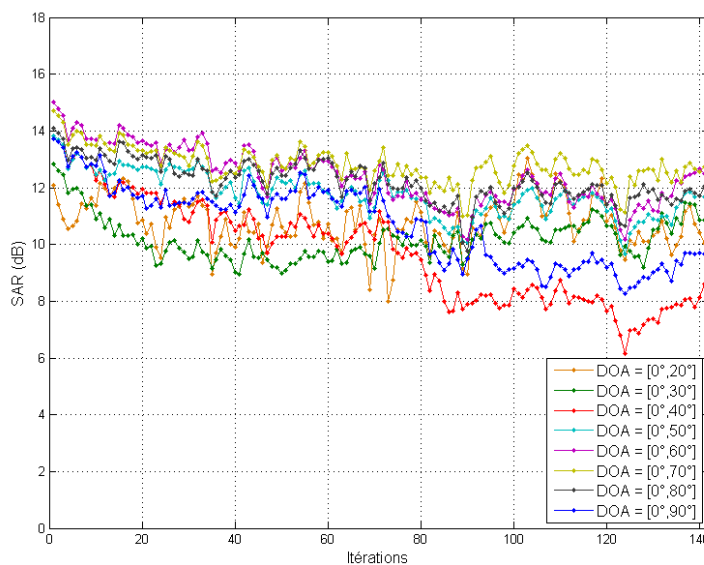


FIGURE 10.18 – Le rapport sources-à-artéfacts SAR de BF_fixed[5°]_NbSrEstim+BSS- l_1 au cours des fenêtres d'analyse : le nombre de sources est estimé (nombre de sources réel égal à 2).

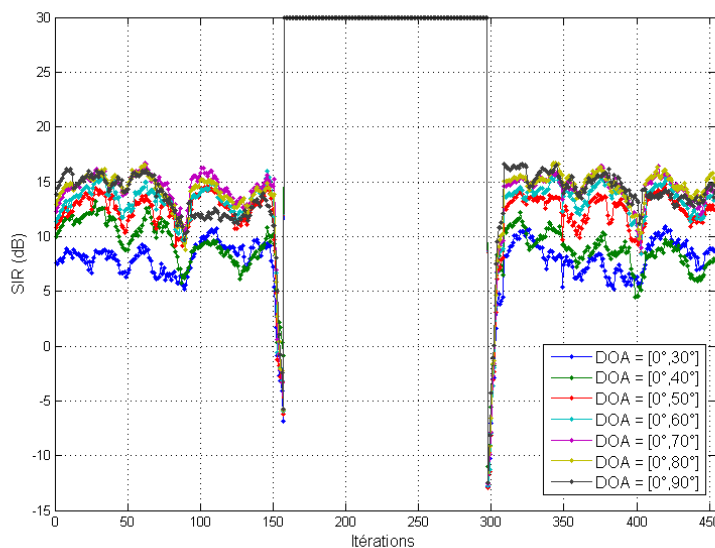


FIGURE 10.19 – Le rapport source-à-interférences SIR de BF_fixed[5°]_NbSrEstim+BSS- l_1 au cours des fenêtres d'analyse : le nombre de sources est estimé et variable entre 1 et 2 sources.

Les figures 10.19, 10.20 et 10.21 montrent le SIR, le SDR et le SAR moyens dans les cas de séparation où le nombre de sources est variable au cours du temps et pour

différentes directions d'arrivées.

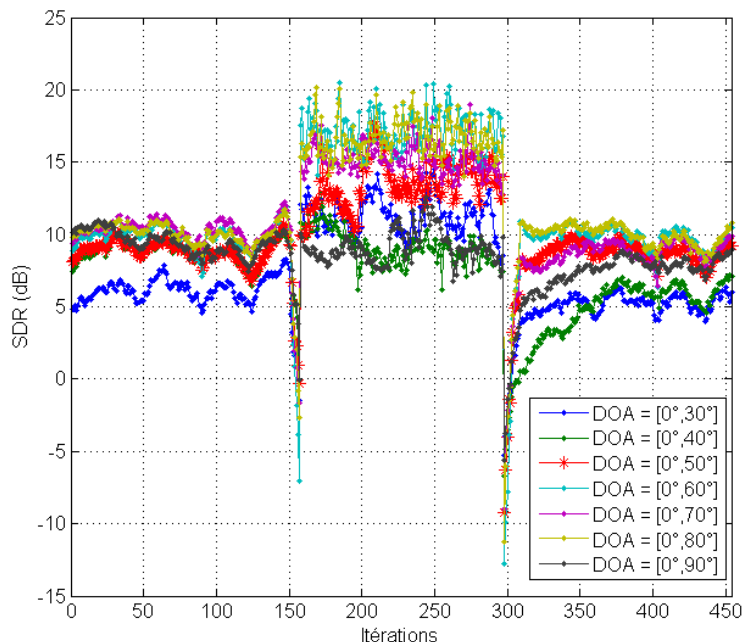


FIGURE 10.20 – Le rapport source-à-distorsion SDR de $\text{BF_fixed}[5^\circ]_\text{NbSrEstim+BSS-}l_1$ au cours des fenêtres d'analyse : le nombre de sources est estimé et variable entre 1 et 2 sources.

Analysons le SIR des sources estimées dans la figure 10.19. Entre la première et la 150^{ème} itérations (le terme itération fait référence à la fenêtre d'analyse longue glissante), le SIR estimé pour les directions d'arrivées considérées varie entre 8 et 16 dB et converge correctement. Ensuite, entre les itérations 150 et 157, nous observons une phase de transition se traduisant par une baisse du SIR due à la disparition progressive de la source qui va s'éteindre. Entre les itérations 158 et 299, une seule source est active ce qui se traduit par un fort rapport source-à-interférences. Quand la source qui a disparue pendant les 141 dernières itérations s'active de nouveau, nous assistons à une nouvelle phase de transition qui se traduit par une baisse des performances de séparation. Cette baisse de SIR est due au nombre d'échantillons encore peu nombreux de la source qui vient de s'activer. Le SIR augmente au fur et à mesure que plus d'échantillons de la source qui vient de s'activer sont disponibles, ce qui est la conséquence d'une meilleure séparation. Cette augmentation du SIR se stabilise après la 350 itération pour atteindre la même performance qu'entre la première et la 150^{ème} itérations. Notre algorithme suit donc bien le changement dynamique du nombre de sources et converge assez rapidement. Nous rappelons

que la matrice de séparation est initialisée une seule fois et que l'adaptation est totalement automatique et dépend du nombre de sources estimé.

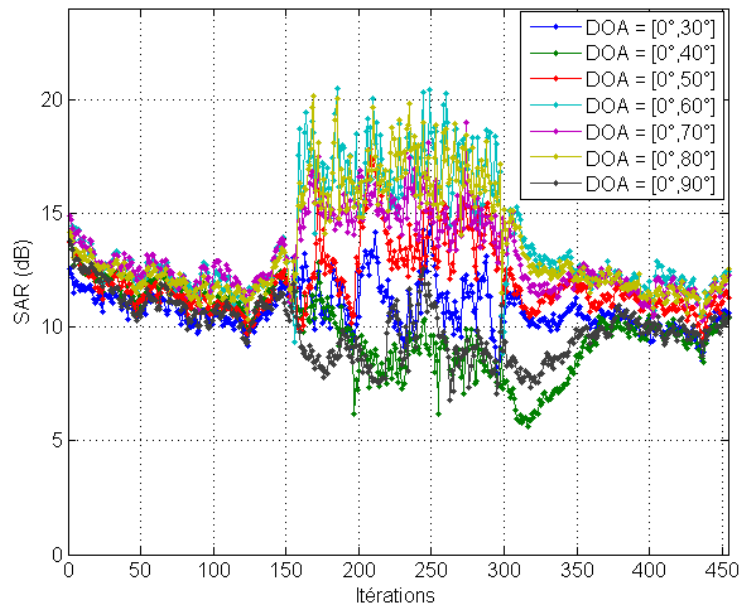


FIGURE 10.21 – Le rapport sources-à-artéfacts SAR de `BF_fixed[5°]_NbSrEstim+BSS-l1` au cours des fenêtres d'analyse : le nombre de sources est estimé et variable entre 1 et 2 sources.

10.3 Résumé des résultats

Pour résumer les résultats obtenus avec notre algorithme de séparation adaptative de sources audio, nous avons calculé la moyenne sur toutes les fenêtres d'analyse et pour chaque direction d'arrivées du rapport source-à-interférences SIR, du rapport source-à-distorsion SDR et du rapport sources-à-artéfacts SAR quand le nombre de sources est fixe (*cf.* figures 10.22, 10.23 et 10.24) et variable (*cf.* figures 10.25, 10.26 et 10.27) au cours du temps. Dans chaque figure, nous avons tracé les performances moyennes des trois cas que nous avons étudiés : le nombre de sources à séparer est connu (courbes bleues), le nombre de sources à séparer est fixé *a priori* ($N_{hyp} = 5$, courbes rouges) et le nombre de sources à séparer est estimé (courbes vertes).

Dans le cas où le nombre de sources est fixe au cours des itérations, les performances de séparation quand le nombre de sources est connu sont les meilleures (*cf.* figures 10.22, 10.23 et 10.24). Cependant les performances de l'algorithme de séparation avec estimation du nombre de sources ne sont pas loin derrière et nous

remarquons qu'il y a entre 1 et 2 dB d'écart entre ces deux algorithmes. Cet écart est dû à une sur-estimation ou une sous-estimation du nombre de sources.

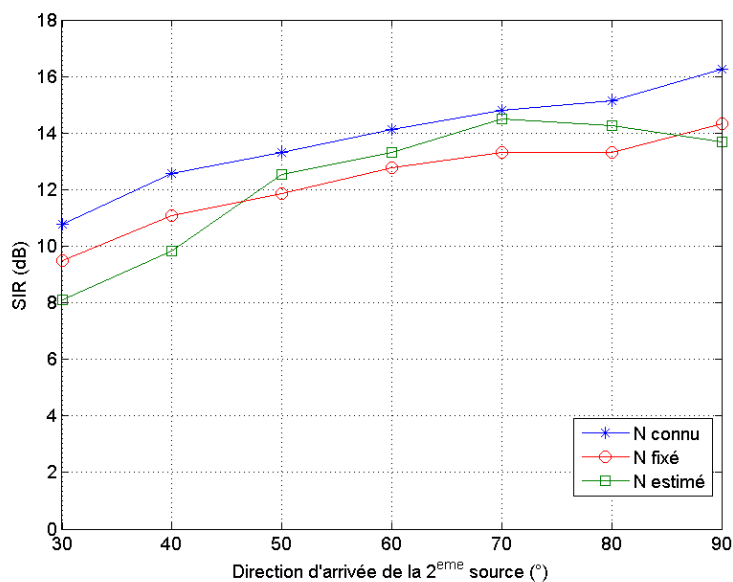


FIGURE 10.22 – Le rapport source-à-interférences SIR moyen sur toutes les fenêtres d'analyse longues : nombre de sources fixe au cours des itérations.

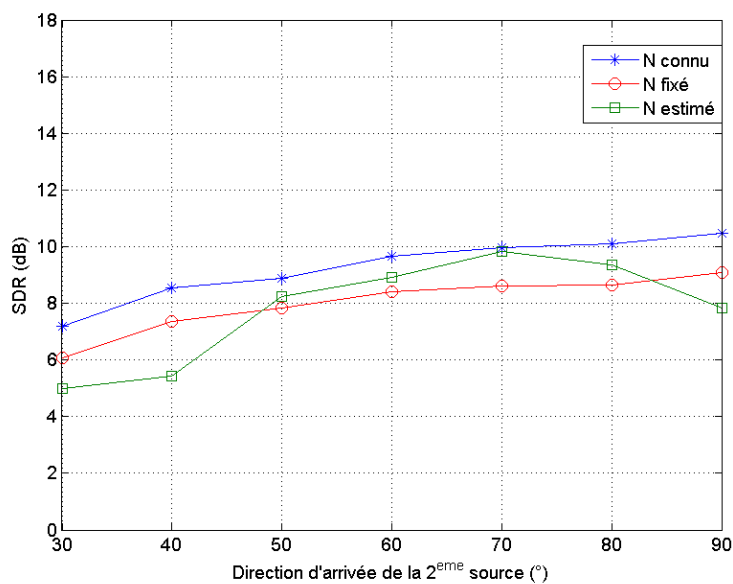


FIGURE 10.23 – Le rapport source-à-distorsion SDR moyen sur toutes les fenêtres d'analyse longues : nombre de sources fixe au cours des itérations.

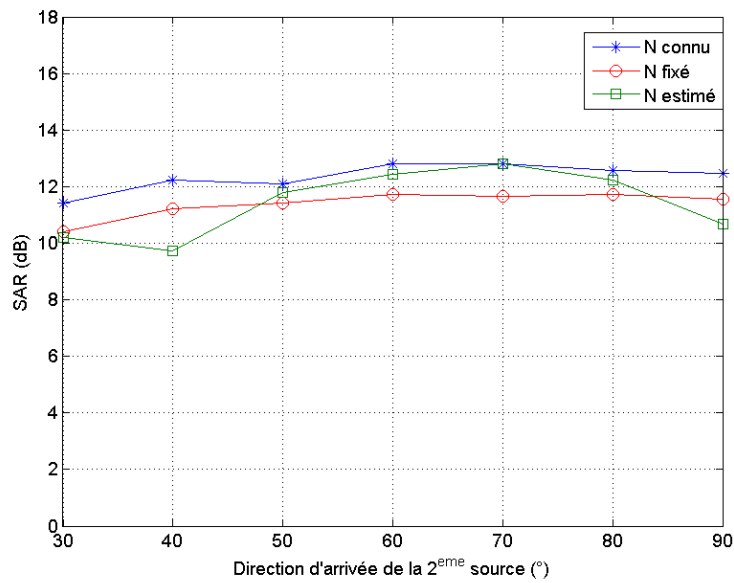


FIGURE 10.24 – Le rapport sources-à-artéfacts SAR moyen sur toutes les fenêtres d’analyse longues : nombre de sources fixe au cours des itérations.

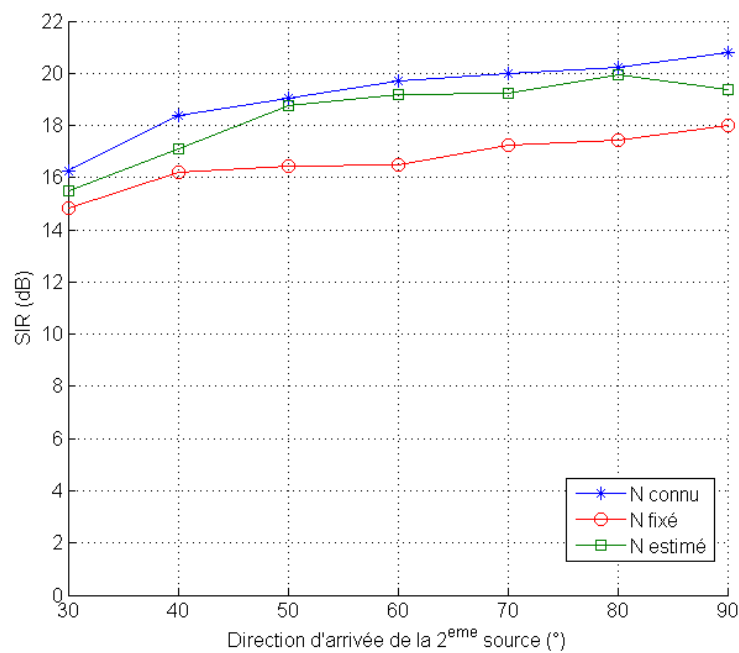


FIGURE 10.25 – Le rapport source-à-interférences SIR moyen sur toutes les fenêtres d’analyse longues : nombre de sources variable au cours des itérations

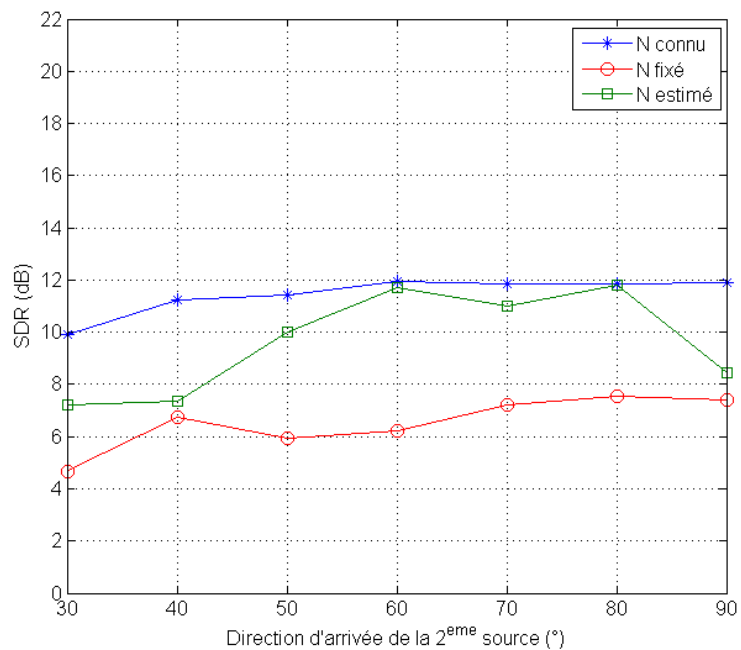


FIGURE 10.26 – Le rapport source-à-distorsion SDR moyen sur toutes les fenêtres d’analyse longues : nombre de sources variable au cours des itérations

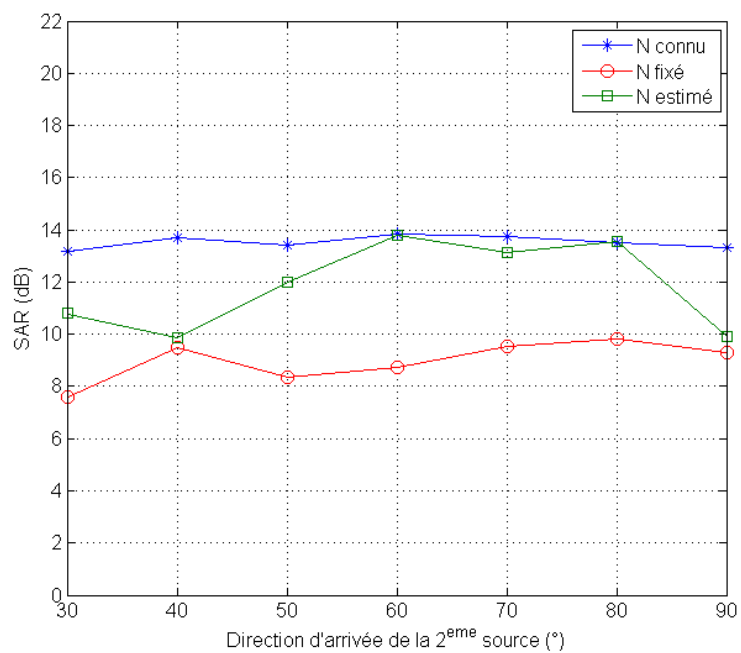


FIGURE 10.27 – Le rapport sources-à-artéfacts SAR moyen sur toutes les fenêtres d’analyse longues : nombre de sources variable au cours des itérations

10.4 Comparaison avec HARK

10.4.1 HARK : principe

Dans le chapitre consacré à l'état de l'art de la séparation de sources audio, nous avons cité le logiciel d'audition des robots HARK [55]. HARK, pour HRI-JP Audition for Robots with Kyoto University, qui veut dire « listen » en anglais médiéval, est un logiciel open source d'audition des robots. HARK a été réalisé en 2008 et la dernière version 1.0.0 a été introduite en novembre 2010, il se base sur l'environnement de programmation par flot de données Flowdesigner, fonctionne en temps-réel, comporte un certain nombre de modules pour l'audition des robots et supporte les convertisseurs analogiques/numériques multicanaux. Ce logiciel complet d'audition des robots comporte les modules suivants :

1. localisation de sources avec MUSIC (Multiple Signal Classification) ;
2. séparation de sources avec une décorrélation géométrique d'ordre supérieur et un pas d'adaptation adaptatif, cet algorithme de séparation s'appelle GHDSS-AS (Geometric High-order Decorrelation-based Source Separation with Adaptive Stepsize) ;
3. amélioration de la qualité de la parole (*Speech enhancement*) avec HRLE (Histogram-based Recursive Level Estimation) ;
4. reconnaissance automatique de la parole avec MFT-ASR.

Nous nous intéressons au module de séparation de sources GHDSS-AS afin de le comparer à notre algorithme de séparation adaptatif avec estimation du nombre de sources BF_fixed[5°]_NbSrEstim+BSS- l_1 .

GHDSS-AS [61] est un algorithme de séparation de sources hybride qui combine formation de voies et séparation de sources, mais d'une manière complètement différente de la nôtre, et qui utilise un pas d'adaptation adaptatif.

Soient un signal audio de N sources $\mathbf{s}(t) = [s_1(t), \dots, s_N(t)]^T$ et d'un réseau de M microphones. Les sorties du réseau de capteurs sont notés $\mathbf{x}(t) = [x_1(t), \dots, x_M(t)]^T$, t étant l'indice de temps. Soit $\mathbf{h} = [\mathbf{h}(0), \dots, \mathbf{h}(L-1)]$ le vecteur des réponses impulsionnelles tronquées à la longueur L , avec

$\mathbf{h}(l) = [h_{ij}]_{1 \leq i \leq N, 1 \leq j \leq M}$, une matrice de dimension $N \times M$ contenant

les $l^{\text{ème}}$ coefficients des réponses impulsionnelles des différents chemins acoustiques entre les N sources et M capteurs. Les mélanges à la sortie des capteurs sont la somme des convolutions entre les signaux sources et les réponses impulsionnelles

des différents chemins de propagation entre les sources et les capteurs. Dans le domaine temps-fréquence, les signaux du mélange à l'indice fréquentiel f et à la trame temporelle k peuvent être approximés par :

$$\mathbf{X}(f, k) \simeq \mathbf{H}(f) \mathbf{S}(f, k) \quad (10.1)$$

$\mathbf{X}(f, k) = [X_1(f, k), \dots, X_M(f, k)]^H$ (respectivement $\mathbf{S}(f, k) = [S_1(f, k), \dots, S_N(f, k)]^H$) est la TFCT de $\{\mathbf{x}(t)\}_{1 \leq t \leq T}$ (respectivement de $\{\mathbf{s}(t)\}_{1 \leq t \leq T}$) à la fréquence $f \in [1, \frac{N_f}{2} + 1]$ et la fenêtre d'analyse $k \in [1, N_T]$. \mathbf{H} est la TFCT des filtres de mélanges $\{\mathbf{h}(l)\}_{0 \leq l \leq L-1}$. Le but est toujours de trouver, à chaque fréquence, une matrice de séparation $\mathbf{W}(f)$ qui conduira à l'estimation des sources originales dans le domaine temps-fréquence :

$$\mathbf{Y}(f, k) = \mathbf{W}(f) \mathbf{X}(f, k) \quad (10.2)$$

avec $\mathbf{Y}(f, k) = [Y_1(f, k), \dots, Y_N(f, k)]^H$. Cette matrice de séparation est estimée en minimisant une fonction de coût $J(\mathbf{W})$ par une méthode de descente de gradient par exemple, l'équation de mise à jour s'écrit :

$$\mathbf{W}_{n+1} = \mathbf{W}_n - \mu \nabla J(\mathbf{W}_n) \quad (10.3)$$

où \mathbf{W}_n représente \mathbf{W} à l'itération n et μ est le pas d'avancement.

La méthode de séparation de sources proposée par Nakajima et *al.* [61] consiste en la combinaison d'un algorithme de séparation de sources basé sur la décorrélation d'ordre supérieur (HDSS : High-Order Decorrelation based Source Separation) et d'une contrainte géométrique typiquement par formations de voies. La fonction de coût globale J_{GHDS} s'écrit alors comme suit :

$$J_{GHDS} = \alpha J_{HDSS}(\mathbf{W}) + \beta J_{GC}(\mathbf{W}) \quad (10.4)$$

où J_{HDSS} est la fonction de coût relative à la décorrélation d'ordre supérieur HDSS, J_{GC} est la fonction de coût relative à la contrainte géométrique et α et β sont deux facteurs de poids tel que $\alpha + \beta = 1$.

HDSS est un algorithme qui utilise une matrice de corrélation d'ordre supérieur $\mathbf{E}_\phi = \phi(\mathbf{Y}(f, :)) \mathbf{Y}^H(f, :) - \text{diag}(\phi(\mathbf{Y}(f, :)) \mathbf{Y}^H(f, :))$ comme fonction de coût, où ϕ est une fonction non linéaire. $J(\mathbf{W}_n)$ pour HDSS est défini comme suit :

$$J_{HDSS}(\mathbf{W}) = \|E(\mathbf{E}_\phi)\|^2 \quad (10.5)$$

La contrainte géométrique est basée sur une formation de voies de type *Delay-and-Sum* et s'écrit comme suit :

$$J_{GC}(\mathbf{W}_n) = \|E(\mathbf{E}_{GC})\|^2 \quad (10.6)$$

$$\mathbf{E}_{GC} = \text{diag}(\mathbf{W}\mathbf{D} - \mathbf{I}) \quad (10.7)$$

Dans [61], les auteurs proposent aussi d'utiliser des pas de mise à jour μ_{HDSS} et μ_{GC} relatifs à chacune des deux fonctions de coût 10.5 et 10.6. Ces pas de mise à jour sont calculés d'une manière adaptative de la manière suivante :

$$\mu_{HDSS} = \frac{\|\mathbf{E}_\phi\|^2}{2 \left\| 2\mathbf{E}_\phi \tilde{\phi}(\mathbf{Y}(f, :)) \mathbf{X}^H(f, :) \right\|^2} \quad (10.8)$$

$$\mu_{GC} = \frac{\|\mathbf{E}_{GC}\|^2}{2 \|\mathbf{E}_{GC}\mathbf{D}^H\|^2} \quad (10.9)$$

où $\tilde{\phi}(Y_i(f, k)) = \phi(Y_i(f, k)) + Y_i(f, k) \frac{\partial \phi(Y_i(f, k))}{\partial Y_i(f, k)}$.

10.4.2 Evaluation des résultats

Nous avons comparé les performances de séparation de notre algorithme adaptatif de séparation de source avec prétraitement par formation de voies fixe suivi de la sélection de lobes et estimation du nombre de sources BF_fixed[5°]_NbSrEstim+BSS- l_1 avec celle obtenues par le module de séparation de source GHDSS-AS de HARK (nommé HARK dans les figures 10.28 et 10.29). Cette évaluation des performances de séparation a été effectuée sur 40 cas de séparation de deux sources de la base de données Theo-RI-studio, la première source est placée à 0° et la deuxième à 60° , le nombre de source est fixe au cours du temps. Pour une source estimée, les mesures de qualité ont été évaluées sur la totalité de ce signal, donc après sa reconstruction, et non pas sur les fenêtres d'analyse comme nous l'avons effectué dans les deux sections précédentes. Les mesures qui seront présentées dans la suite sont la moyenne et l'écart type des 40 cas de séparation de deux sources.

La figure 10.28 montre que les performances de séparation de notre algorithme BF_fixed[5°]_NbSrEstim+BSS- l_1 en terme de rapport source-à-interférences SIR, de rapport source-à-distorsion SDR et de rapport sources-à-artéfacts SAR sont au même niveau que celles obtenues par HARK.

La figure 10.29 montre les performances de séparation perceptuelles en terme

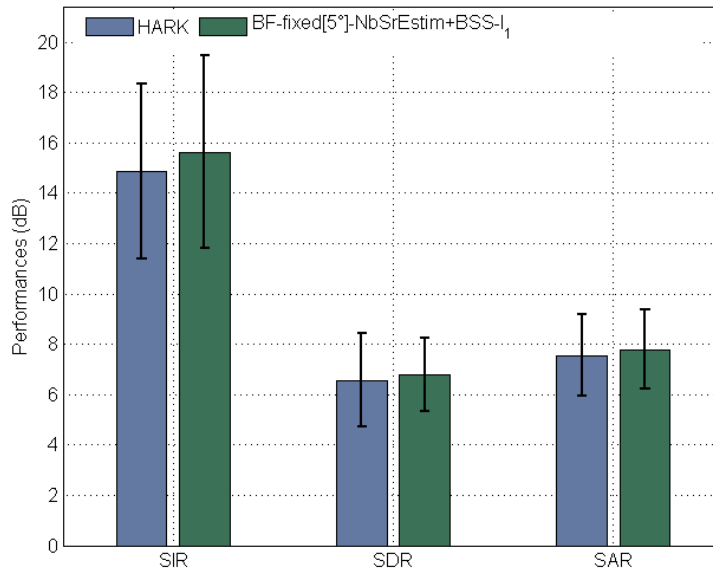


FIGURE 10.28 – Le rapport source-à-interférences (SIR), le rapport source-à-distorsion (SAR) et le rapport sources-à-artéfacts (SAR) des sources estimées avec HARK (en bleu) et BF_fixed[5°]_NbSrEstim+BSS- l_1 (en vert) et leurs écarts types (barres verticales) : les moyennes pour la séparation de 40 paires de sources de la base de données Theo-RI-studio, les directions d’arrivées sont 0° et 60° .

de score perceptuel global OPS, de score perceptuel relatif à la source cible TPS, de score perceptuel relatif aux interférences IPS et de score perceptuel relatif aux artéfacts APS de BF_fixed[5°]_NbSrEstim+BSS- l_1 et de HARK. Si nous analysons ces mesures, nous remarquons que :

- en terme de préservation de la source cible (TPS), nous avons un score proche de celui obtenu par HARK ;
- en terme de suppression des autres sources, celles qui sont différentes de la source cible (IPS), nous avons un score supérieur à celui de HARK d’environ 40 points ;
- en terme d’absence de bruit artificiel additionnel dans les sources estimées (APS), nous avons un score inférieur à celui de HARK de 25 points.

Les observations précédentes nous amènent à un score perceptuel global de notre algorithme BF_fixed[5°]_NbSrEstim+BSS- l_1 inférieur à celui de HARK de 15 points.

Ces premiers résultats obtenus en comparant notre algorithme de séparation aveugle de sources avec estimation du nombre de sources BF_fixed[5°]_NbSrEstim+BSS- l_1 à l’algorithme référence de l’état de l’art relatif à la séparation de sources pour l’audition des robots sont très prometteurs. Ils montrent que les résultats en terme

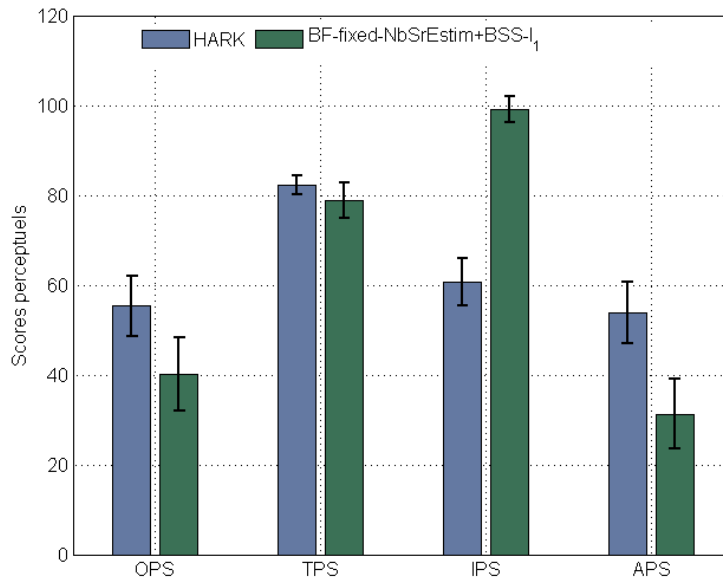


FIGURE 10.29 – Le score perceptuel global (OPS), le score perceptuel relatif à la cible (TPS), le score perceptuel relatif aux interférences (IPS) et le score perceptuel relatif aux artéfacts (APS) des sources estimées avec HARK (en bleu) et BF_fixed[5°]_NbSrEstim+BSS- l_1 (en vert) et leurs écarts types (barres verticales) : les moyennes pour la séparation de 40 paires de sources de la base de données TheoRI-studio, les directions d’arrivées sont 0° et 60° .

de séparation de sources sont comparables à ceux obtenus par HARK et qu’un travail reste à faire afin d’améliorer le score perceptuel relatif aux artéfacts.

Conclusion

Dans ce chapitre, nous nous sommes intéressés à la séparation de sources adaptative avec un nombre de sources fixe ou variable au cours du temps. Nous avons testé trois configurations : le nombre de sources est connu, le nombre de sources est inconnu mais fixé *a priori* à une valeur maximale choisie (*i.e.* sur-estimé) et le nombre de sources est estimé. Les conclusions suivantes peuvent être tirées :

- les performances de l’algorithme de séparation avec estimation du nombre de sources BF_fixed[5°]_NbSrEstim+BSS- l_1 sont proches de celles où le vrai nombre de sources est connu BF_fixed[5°]_BS+BSS- l_1 de 2dB en moyenne ;
- Dans le cas où nous estimons un nombre de sources supérieur au nombre de sources réel, les performances de séparation sont affectées par des distorsions et des artéfacts ;

- notre algorithme `BF_fixed[5°]_NbSrEstim+BSS-l1` a des performances très proches de celles de l'algorithme référence de l'état de l'art de l'audition des robots HARK, les performances de séparation ont été évaluées pour la séparation de 40 paires de sources de la base de données Theo-RI-studio, les directions d'arrivées sont 0° et 60°.
-