

Évaluation des algorithmes itératifs de séparation de sources

Introduction

Dans ce chapitre, nous présentons les performances des algorithmes itératifs de séparation de sources présentés dans les chapitres 4 et 5 à savoir :

- l’algorithme de séparation basé sur un critère de parcimonie $BSS-l_1$ que nous comparons à l’analyse en composantes indépendantes ICA ;
- l’algorithme de séparation basé sur un critère de parcimonie variable $BSS-l_p$ qui se traduit par la minimisation de la norme l_p paramétrée ;
- la classe d’algorithmes avec prétraitement de formation de voies avec les quatre différentes configurations de ce prétraitement à savoir :
 - formation de N voies vers les directions d’arrivées des sources BF_DOA ;
 - formation de K voies fixe BF_fixed ;
 - formation de voies fixe avec sélection des lobes ayant la plus grande énergie BF_fixed_BS ;
 - formation de voies fixe avec estimation du nombre de source BF_fixed_NbSrEstim.

Dans la suite, nous détaillerons les paramètres de ces algorithmes de séparation itératifs 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 itératifs

Les signaux de test des différentes bases de données ont une durée de 5s et sont échantillonnés à 16 kHz. La fenêtre d'analyse spectrale est de type Hanning et de longueur 128 ms (2048 échantillons) et le pas d'avancement est de 50%. Le pas de mise à jour des algorithmes itératifs est $\mu = 0.2$ choisi d'après l'état de l'art [29]. Nous choisissons un nombre d'itérations égal à 100 et nous verrons que les algorithmes présentés convergent bien avant d'atteindre ce nombre.

9.1 Comparaison entre le critère de parcimonie et le critère d'indépendance

Nous commençons cette évaluation des algorithmes de séparation de sources proposés par la comparaison des performances de BSS- l_1 , l'algorithme basé sur un critère de séparation de parcimonie qui est la minimisation de la norme l_1 ; et de l'algorithme basé sur un critère d'indépendance qui est l'analyse en composantes indépendantes avec minimisation de l'information mutuelle de Douglas et Gupta [29].

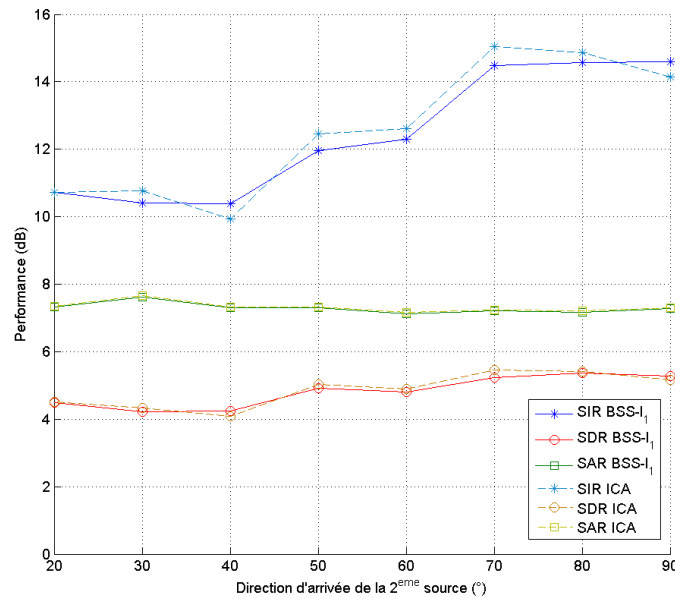


FIGURE 9.1 – Le rapport source-à-interférences SIR (courbes en bleu), le rapport source-à-distorsion SDR (courbes en rouge) et le rapport sources-à-artéfacts SAR (courbes en vert) des algorithmes BSS- l_1 (courbes continues) et ICA (courbes en tirets) : évaluation sur la base de données Theo-Rl-studio.

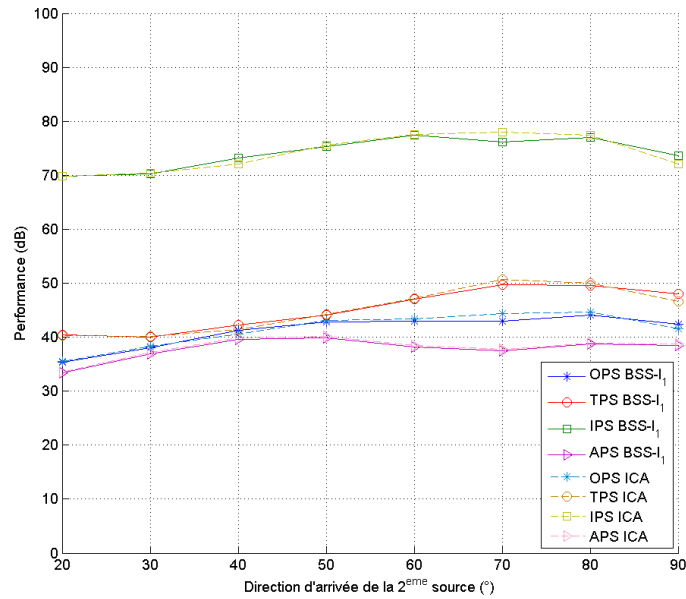


FIGURE 9.2 – Le score perceptuel global OPS (courbes en bleu), le score perceptuel relatif à la cible TPS (courbes en rouge), le score perceptuel relatif aux interférences IPS (courbes en vert) et le score perceptuel relatifs aux artéfacts APS (courbes en mauve) des algorithmes BSS- l_1 (courbes continues) et ICA (courbes en tirets) : évaluation sur la base de données Theo-RI-studio.

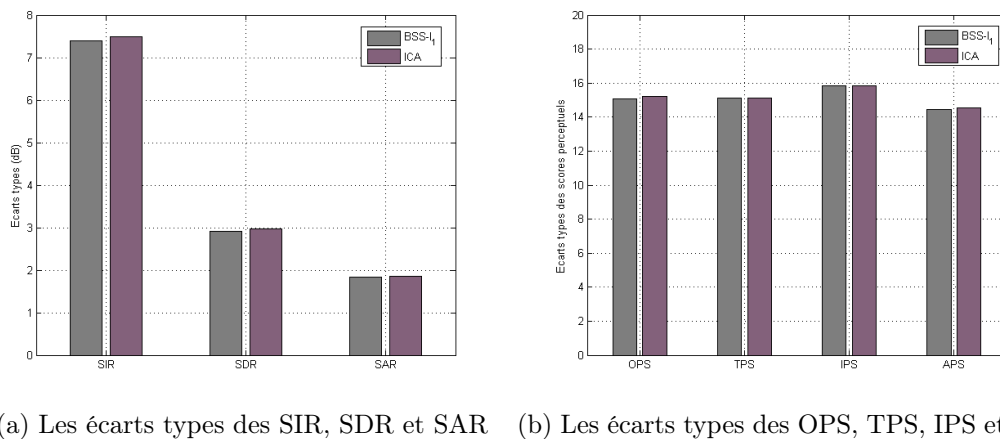


FIGURE 9.3 – Les écarts types des résultats de séparations des algorithmes BSS- l_1 (barres grises) et ICA (barres violettes) : évaluation sur la base de données Theo-RI-studio.

La figure 9.1 montre une comparaison du rapport signal-sur-bruit SIR, le rapport signal-à-distorsion SDR et le rapport signal-à-artéfacts SAR de ces deux algorithmes évalués sur la base de données **Theo-RI-studio**. Les courbes montrent que les performances de ces deux approches sont comparables. Nous retrouvons ce même constat dans les performances de l'évaluation perceptuelle du résultat de séparation de ces deux algorithmes (*cf.* figure 9.2). En effet, 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 de **BSS- l_1** et d'**ICA** sont proches comme le confirme la figure 9.3 des écarts types. Ceci peut s'expliquer par le fait que l'analyse en composantes indépendantes conduit à la minimisation de la parcimonie des sources à estimer à partir des mélanges convolutifs, et ceci de la même façon que la minimisation de la norme l_1 .

9.2 Minimisation de la norme l_p paramétrée

Avant le développement de l'algorithme basé sur la minimisation de la norme l_p paramétrée **BSS- l_p -param** [5], nous avons essayé de savoir comment se comporte un algorithme basé sur la minimisation de la norme l_p avec différentes valeurs du paramètre p , donc différentes contraintes de parcimonie. Nous avons considéré un paramètre p entre 0.1 (contrainte de parcimonie la plus dure) et 1 (contrainte de parcimonie la moins dure). La figure 9.4 montre le résultat de la séparation en termes de rapport signal-à-interférences de quatre cas de séparation de deux sources. Nous remarquons que pour la pseudo norme l_p avec un paramètre $p < 1$, nous pouvons avoir de meilleurs résultats de séparation que pour la norme l_1 . Cependant, la valeur du paramètre p pour laquelle le SIR est le plus élevé est différente d'un cas de séparation à un autre. L'écart entre les rapports source-à-interférences (SIR) obtenus en utilisant différentes valeurs du paramètre p peut s'expliquer par la convergence vers des minima locaux. Ceci peut être vérifié en remplaçant la méthode d'optimisation actuelle par une méthode qui peut éviter la convergence vers un minimum local, par exemple en utilisant le recuit simulé. Nous n'avons pas pu tester ce genre de méthodes mais ceci reste dans nos perspectives à court terme.

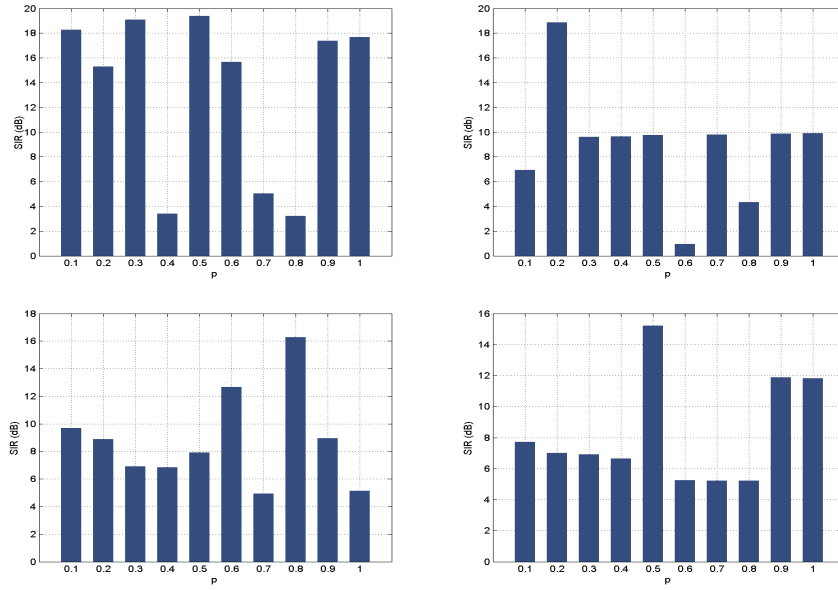


FIGURE 9.4 – Variation du rapport signal-à-interférences SIR de l’algorithme $BSS-l_p$ en fonction du paramètre p : exemple de séparation de 4 paires de sources de la base de données Theo-RI-studio.

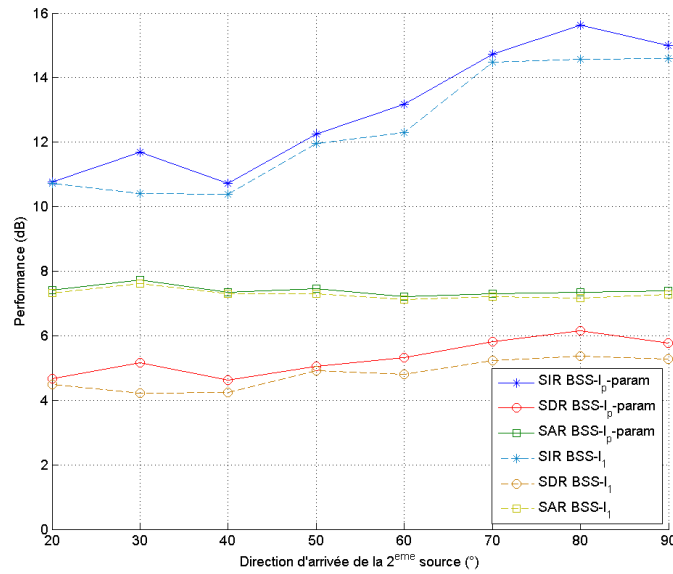


FIGURE 9.5 – Le rapport source-à-interférences SIR (courbes en bleu), le rapport source-à-distorsion SDR (courbes en rouge) et le rapport sources-à-artéfacts SAR (courbes en vert) de l’algorithme $BSS-l_p\text{-param}$ (courbes continues) en comparaison avec $BSS-l_1$ (courbes en tirets) : évaluation sur la base de données Theo-RI-studio.

Nous avons donc essayé d’exploiter la norme l_p dans la séparation de source sous

une autre forme. Nous avons tenté de durcir la contrainte de parcimonie de la norme l_p au fur et à mesure que l'algorithme avance dans les itérations en faisant décroître le paramètre p selon une fonction sigmoïde comme nous l'avons présenté dans la section 4.4. Nous appelons cet algorithme **BSS- l_p -param**. La figure 9.5 montre le rapport source-à-interférences SIR, le rapport source-à-distorsion SDR et le rapport sources-à-artéfacts SAR de **BSS- l_1** et **BSS- l_p -param**. Les performances de **BSS- l_p -param** sont légèrement supérieures à celles de l'algorithme **BSS- l_1** . L'analyse perceptuelle (*cf.* figure 9.6) montre que le score perceptuel relatif aux artéfacts de **BSS- l_p -param** est légèrement inférieur à celui de **BSS- l_1** . L'algorithme **BSS- l_p -param** a l'avantage d'avoir des écarts types inférieurs à ceux de **BSS- l_1** pour les scores perceptuels et le rapport source-à-interférences (*cf.* figure 9.7).

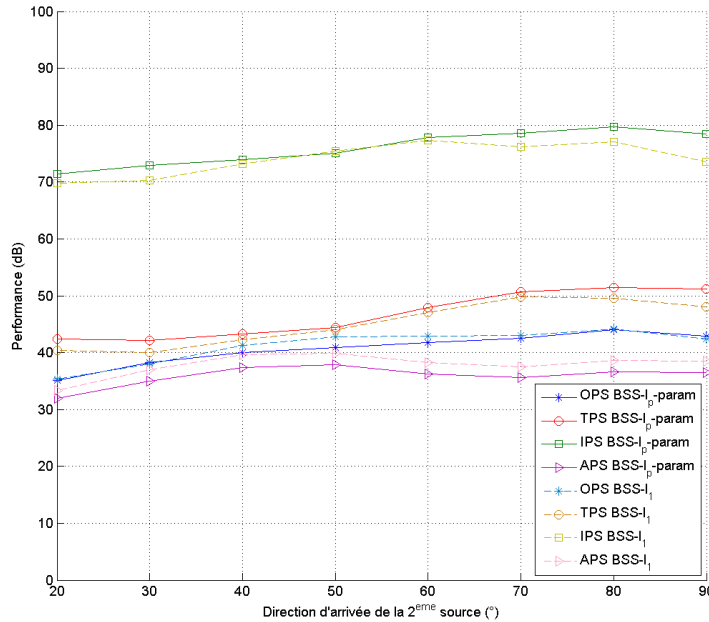


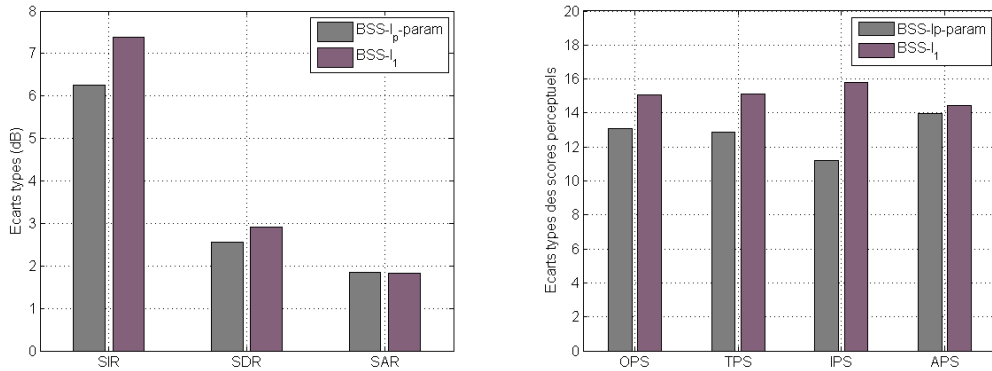
FIGURE 9.6 – Le score perceptuel global OPS (courbes en bleu), le score perceptuel relatif à la cible TPS (courbes en rouge), le score perceptuel relatif aux interférences IPS (courbes en vert) et le score perceptuel relatifs aux artéfacts APS (courbes en mauve) de l'algorithme **BSS- l_p -param** (courbes continues) en comparaison avec **BSS- l_1** (courbes en tirets) : évaluation sur la base de données Theo-RI-studio.

La figure 9.8 montre les courbes de la parcimonie moyenne des sources estimées au cours des itérations mesurées par l'indice de Gini et pour les différentes directions d'arrivées considérées. L'indice de Gini est une bonne mesure de parcimonie pour la parole [69], il est compris entre 0 et 1 : plus le signal est parcimonieux, plus son indice de Gini est proche de 1. Nous remarquons qu'après une centaine d'itérations, ce qui correspond à un p inférieur à 0.5, les courbes de parcimonie décroissent légèrement.

Quand le paramètre p est inférieur à 0.5, nous remarquons que les performances de chaque paire de sources estimée sont chahutées (exemple du rapport source-à-interférences SIR de la figure 9.9 au cours des itérations et pour différentes directions d'arrivées), ce qui implique une baisse de la moyenne générale de la parcimonie.

Ceci peut être dû au fait que la contrainte de parcimonie devient très dure pour permettre une convergence stable de l'algorithme (si nous considérons le cas de chaque source estimée individuellement). Suite à cette constatation, deux pistes sont possibles : la première consiste à arrêter les itérations à 100 et donc à $p = 0.5$, la deuxième consiste à faire décroître p de 1 à 0.5 selon le même type de sigmoïde en gardant le même nombre d'itérations. Ceci fera l'objet de nos futurs travaux. Dans la même figure 9.8, nous remarquons aussi que la parcimonie des sources estimées augmente à travers les itérations et converge rapidement vers la valeur optimale (moins de 30 itérations dans notre contexte).

L'algorithme de séparation de source avec la minimisation de la pseudo-norme l_p paramétrée est un algorithme prometteur qui ouvre la voie à l'étude de différentes configurations possibles de ce type de fonctions de coût paramétrée. En effet, plusieurs types de contraintes de parcimonie peuvent être étudiés en changeant la forme de la fonction logistique et la valeur minimale à atteindre du paramètre p . Une étude théorique de la convergence de ce genre d'algorithme est aussi nécessaire.



(a) Les écarts types des SIR, SDR et SAR (b) Les écarts types des OPS, TPS, IPS et APS

FIGURE 9.7 – Les écarts types des résultats de séparations des algorithmes BSS- l_p -param (barres grises) et BSS- l_1 (barres violettes) : évaluation sur la base de données Theo-RI-studio.

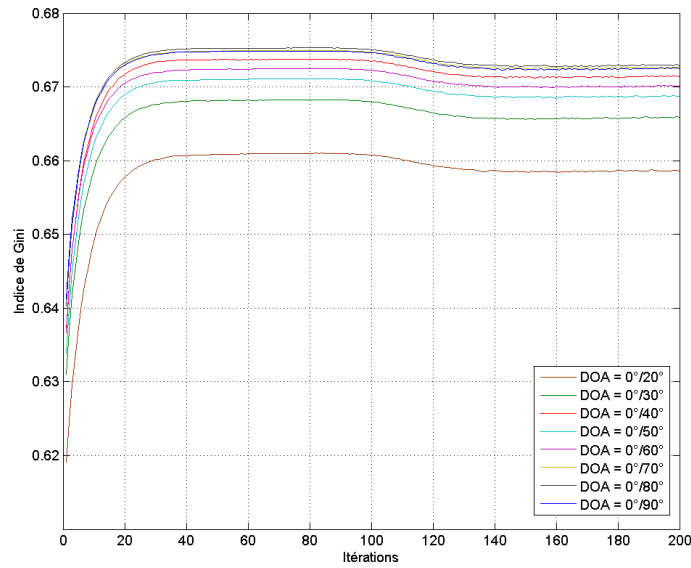


FIGURE 9.8 – L'indice de Gini moyen de BSS- l_p -param au cours des itérations sur la base de données Theo-IR-Studio.

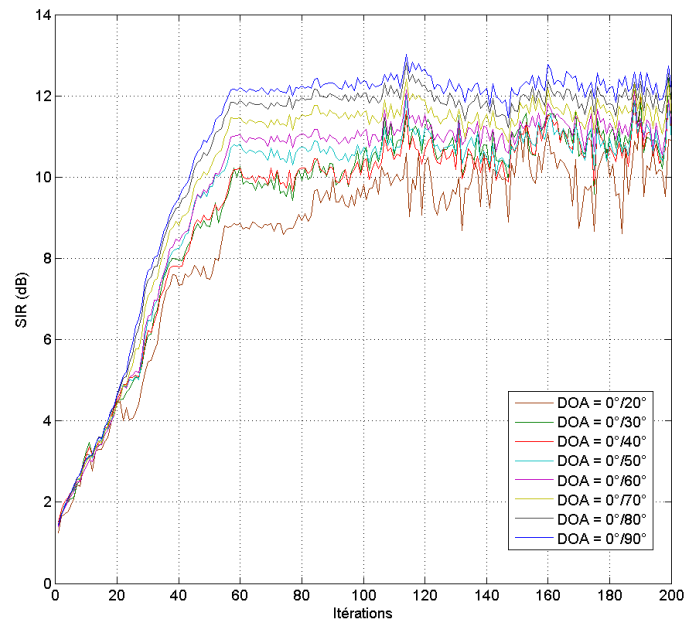


FIGURE 9.9 – Evaluation du rapport source-à-interférences SIR de l'algorithme BSS- l_p -param au cours des itérations sur la base de données Theo-RI-Studio pour les différentes directions d'arrivées DOA considérées.

9.3 Évaluation de l'apport du prétraitement par formation de voies des algorithmes de séparation à deux étapes

Nous étudions dans cette sous-section l'apport du prétraitement par formation de voies en termes de performance de séparation. Une partie de ces résultats a été publiée dans EURASIP Journal on Advances in Signal Processing (*cf.* [3, 4]). Nous considérons les configurations suivantes des algorithmes de séparation à deux étapes avec prétraitement de formation de voies :

- BF_DOA+BSS- l_1 : formation de N voies vers les directions d'arrivées suivi de BSS- l_1 , l'algorithme de séparation de source avec la minimisation de la norme l_1 ;
- BF_fixed[30°]+BSS- l_1 , BF_fixed[30°]+ICA : formation de 7 voies de -90° à 90° avec un angle inter-lobes égal à 30°, suivi respectivement de BSS- l_1 ou ICA ;
- BF_fixed[15°]+BSS- l_1 : formation de 13 voies de -90° à 90° avec un angle inter-lobes égal à 15° suivi de BSS- l_1 ;
- BF_fixed[10°]+BSS- l_1 : formation de 19 voies de -90° à 90° avec un angle inter-lobes égal à 10° suivi de BSS- l_1 ;
- BF_fixed[5°]+BSS- l_1 : formation de 37 voies de -90° à 90° avec un angle inter-lobes égal à 5° suivi de BSS- l_1 ;
- BF_fixed[30°]_BS+BSS- l_1 : formation de 7 voies de -90° à 90° avec un angle inter-lobes égal à 30° suivi d'une sélection de N lobes et BSS- l_1 ;
- BF_fixed[5°]_BS+BSS- l_1 : formation de 37 voies de -90° à 90° avec un angle inter-lobes égal à 5° suivi d'une sélection de N lobes et BSS- l_1 .

9.3.1 Influence du prétraitement par formation de voies

Les figures 9.10 et 9.11 montrent que le rapport source-à-interférences SIR et le rapport source-à-distorsion SDR, obtenus sur la base de données Theo-RI-studio avec l'algorithme BF_fixed[5°]+BSS- l_1 sont supérieurs à ceux obtenus sans prétraitement par formation de voies avec l'algorithme BSS- l_1 et bien supérieur à ceux obtenus avec seulement l'étape de formation de voies BF_fixed[5°]. Nous prenons comme référence le SIR et le SDR des signaux reçus aux capteurs 1 et 2 de Theo que nous avons nommé `signaux_capteurs` dans les figures. Cependant, cette amélioration des performances de séparation apportée par le prétraitement par formation de voies fixe est limitée et n'atteint pas les performances apportées par le prétraitement par

formation de voies vers les directions d'arrivées $\text{BF_DOA+BSS-}l_1$ comme le montre les figures 9.10 et 9.11. Mais nous pouvons dépasser cette limitation par la sélection des lobes comme nous le verrons dans la suite. Les rapports sources-à-artéfacts SAR de l'algorithme de séparation par la minimisation de la norme l_1 , de la formation de voies seule $\text{BF_fixed}[5^\circ]$, de l'algorithme de séparation par la minimisation de la norme l_1 avec formation de voies fixe $\text{BF_fixed}[5^\circ]+\text{BSS-}l_1$ ou vers les directions d'arrivée des sources $\text{BF_DOA+BSS-}l_1$ sont proches.

La figure 9.14 montre que le prétraitement par formation de voies augmente le score perceptuel relatif à la source cible TPS, c'est à dire que les sources séparées sont mieux préservées avec le prétraitement par formation de voies. Cependant, cette amélioration des performances de séparation apportée par le prétraitement par formation de voies fixe ne concerne pas la suppression de la source cible. En effet, les scores perceptuels relatifs aux interférences IPS (*cf.* figure 9.15) des algorithmes $\text{BSS-}l_1$ et $\text{BF_fixed}[5^\circ]+\text{BSS-}l_1$ sont proches alors que celui de l'algorithme avec la formation de voies vers les directions d'arrivée $\text{BF_DOA+BSS-}l_1$ est bien meilleur. D'après la figure 9.16, nous constatons que la formation de voies fixe avec un angle inter-lobes de 5° $\text{BF_fixed}[5^\circ]+\text{BSS-}l_1$ introduit des artéfacts, ce qui se reflète aussi dans son score perceptuel global OPS (*cf.* figure 9.13). Ce résultat n'a pas été mis en évidence par l'analyse objective en termes de SIR, SDR et SAR.

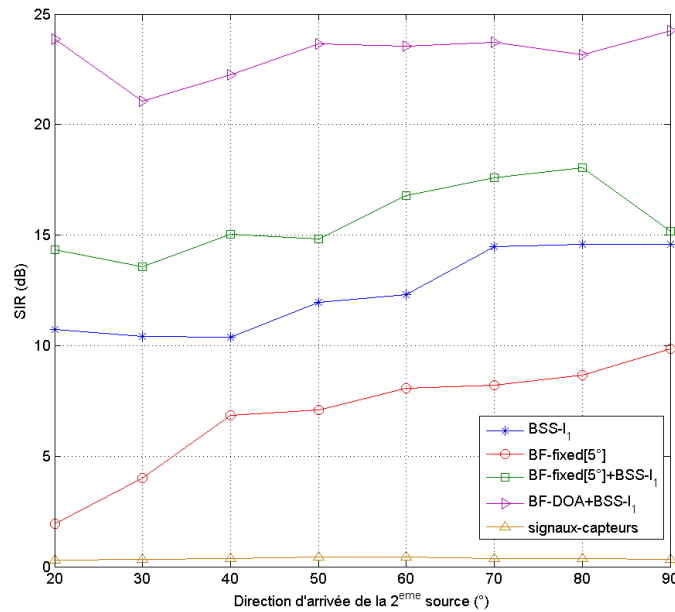


FIGURE 9.10 – Influence du prétraitement par formation de voies en termes de rapport source-à-interférences SIR : évaluation sur la base de données Theo-RI-studio.

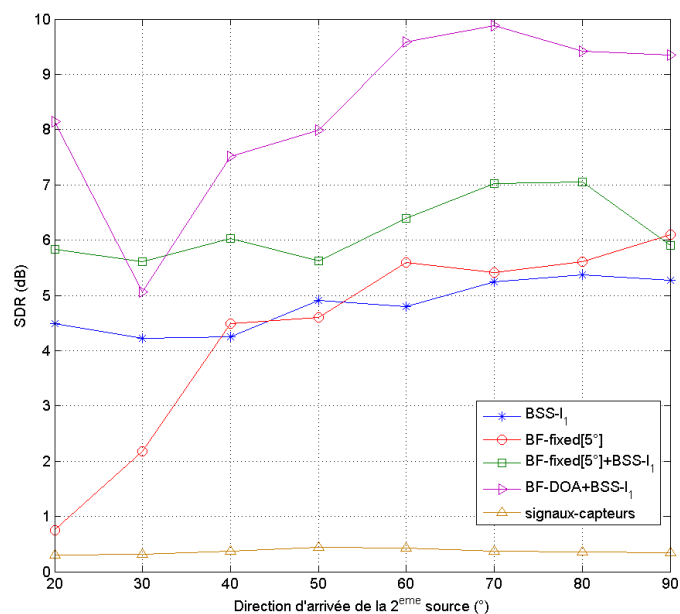


FIGURE 9.11 – Influence du prétraitement par formation de voies en termes de rapport source-à-distorsion SDR : évaluation sur la base de données Theo-RI-studio.

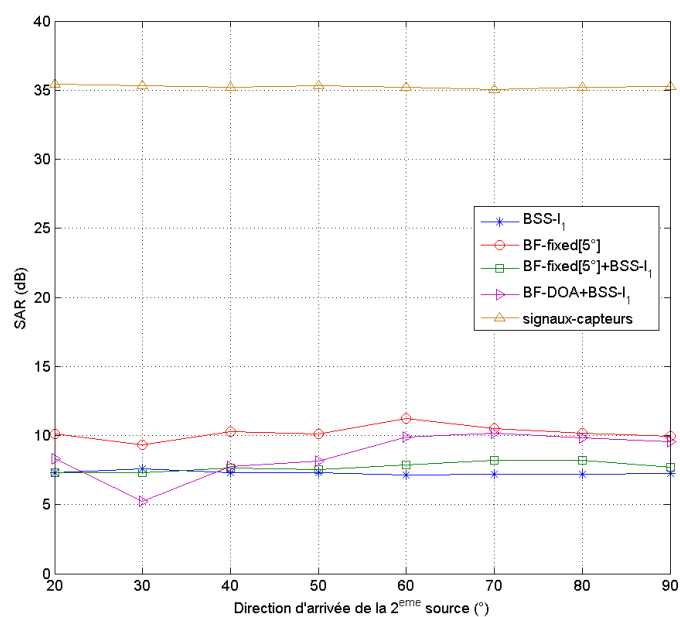


FIGURE 9.12 – Influence du prétraitement par formation de voies en termes de sources-à-artéfacts SAR : évaluation sur la base de données Theo-RI-studio.

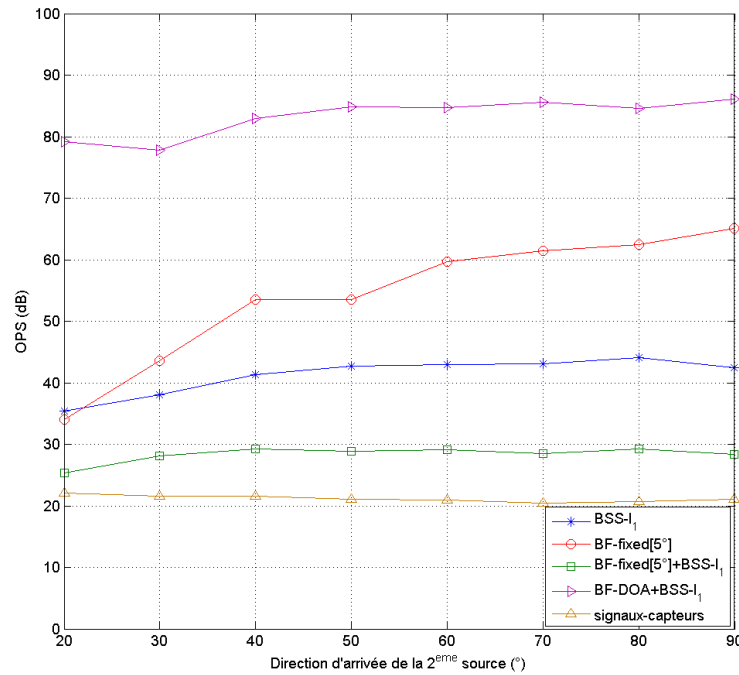


FIGURE 9.13 – Influence du prétraitement par formation de voies en termes de score perceptuel global OPS : évaluation sur la base de données Theo-RI-studio.

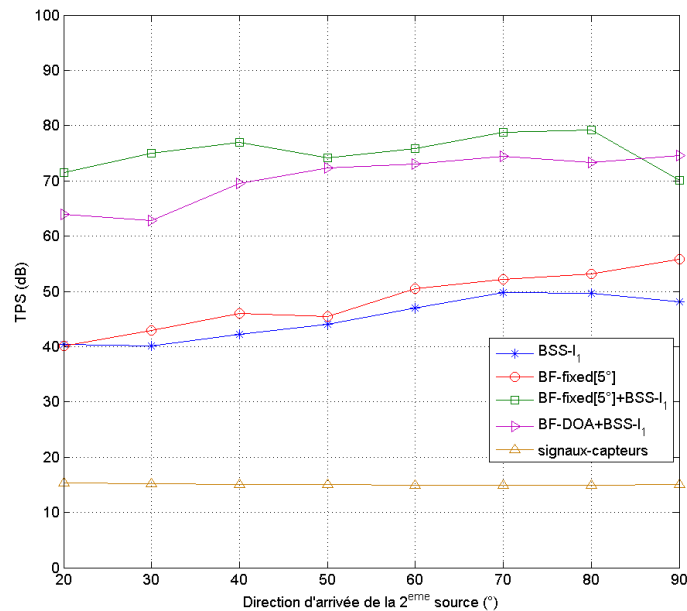


FIGURE 9.14 – Influence du prétraitement par formation de voies en termes de score perceptuel relatif à la source cible TPS : évaluation sur la base de données Theo-RI-studio.

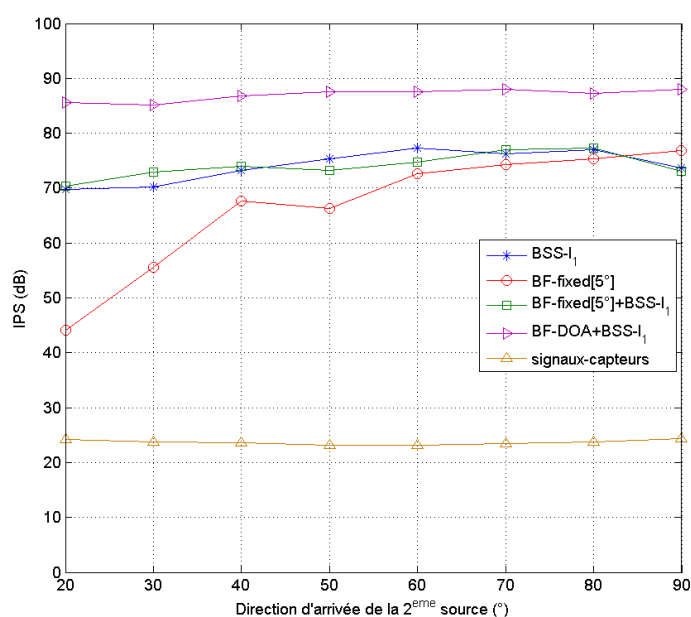


FIGURE 9.15 – Influence du prétraitement par formation de voies en termes de score perceptuel relatif aux interférences IPS : évaluation sur la base de données Theo-RI-studio.

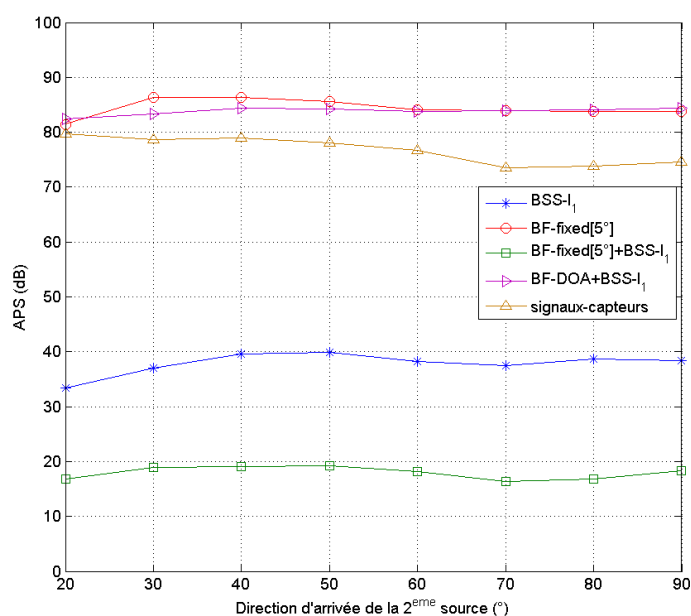


FIGURE 9.16 – Influence du prétraitement par formation de voies en termes de score perceptuel relatifs aux artéfacts APS : évaluation sur la base de données Theo-RI-studio.

9.3.2 Influence de la séparation angulaire entre les lobes :

Les figures 9.17, 9.18 et 9.19 montrent le rapport source-à-interférences SIR, le rapport source-à-distorsion SDR et le rapport sources-à-artéfacts SAR obtenus avec différents angles séparant les lobes de la formation de voies, ces configurations ont été présentées dans l'introduction de cette section. Les résultats montrent que quand nous augmentons le nombre de lobes dans l'espace de formation de voies qui nous intéresse, dans notre cas de -90° à 90° , le SIR et spécialement le SDR augmentent. En effet, en augmentant le nombre de lobes de la formation de voies, nous augmentons la chance de l'algorithme de tomber sur les bonnes directions d'arrivée des sources. Notons que dans les algorithmes $\text{BF_fixed}[15^\circ]+\text{BSS-}l_1$, $\text{BF_fixed}[10^\circ]+\text{BSS-}l_1$ et $\text{BF_fixed}[5^\circ]+\text{BSS-}l_1$, le prétraitement par formation de voies augmente le SDR par rapport à $\text{BSS-}l_1$. Le SIR des algorithmes avec prétraitement par formation de voies est meilleur que celui de l'algorithme de séparation seul $\text{BSS-}l_1$ et ceci dans toutes les configurations testées de la formation de voies fixes.

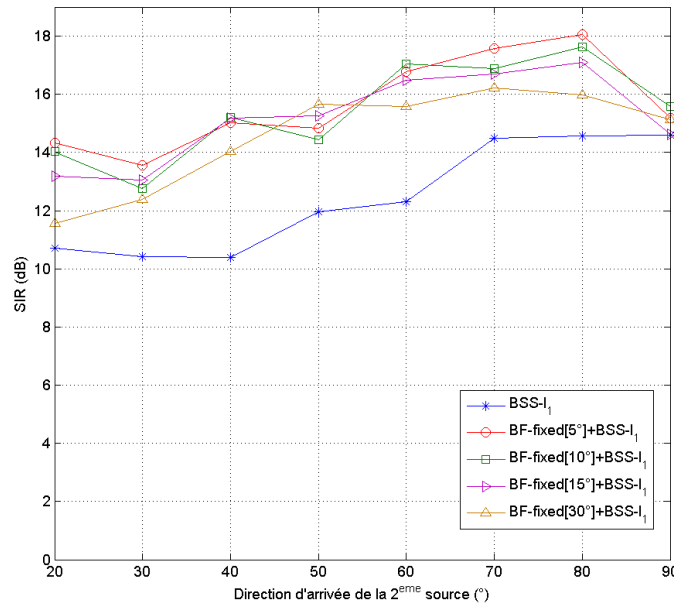


FIGURE 9.17 – Influence de l'angle inter-lobes de la formation de voies en termes de rapport source-à-interférences SIR : angle inter-lobes égal à 5° $\text{BF_fixed}[5^\circ]+\text{BSS-}l_1$, 10° $\text{BF_fixed}[10^\circ]+\text{BSS-}l_1$, 15° $\text{BF_fixed}[15^\circ]+\text{BSS-}l_1$, et 30° $\text{BF_fixed}[30^\circ]+\text{BSS-}l_1$, évaluation sur la base de données Theo-RI-studio.

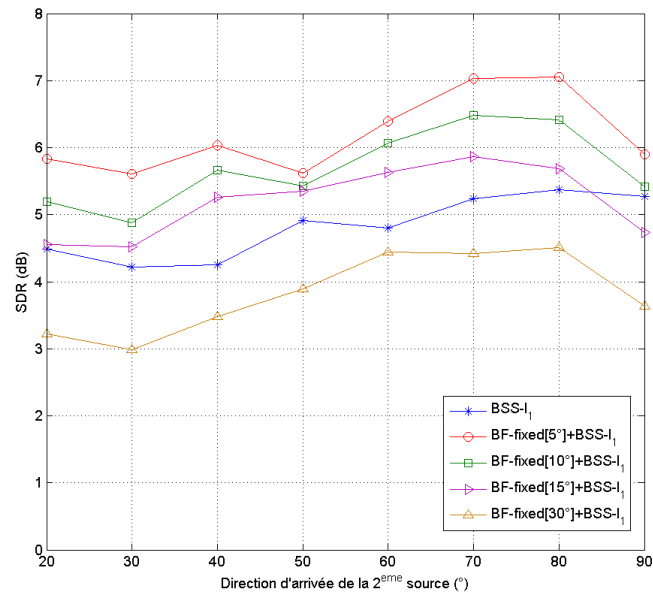


FIGURE 9.18 – Influence de l’angle inter-lobes de la formation de voies en termes de rapport source-à-distorsion SDR : angle inter-lobes égal à 5° BF_fixed[5°]+BSS- l_1 , 10° BF_fixed[10°]+BSS- l_1 , 15° BF_fixed[15°]+BSS- l_1 , et 30° BF_fixed[30°]+BSS- l_1 , évaluation sur la base de données Theo-RI-studio.

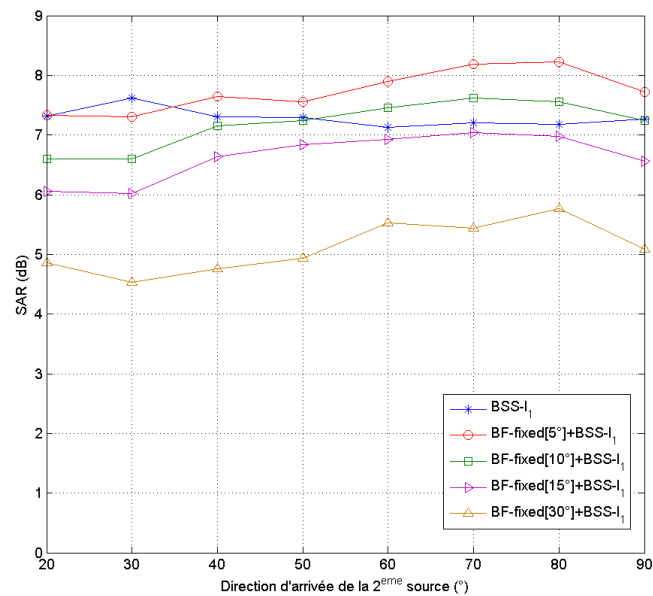


FIGURE 9.19 – Influence de l’angle inter-lobes de la formation de voies en termes de rapport sources-à-artéfacts SAR : angle inter-lobes égal à 5° BF_fixed[5°]+BSS- l_1 , 10° BF_fixed[10°]+BSS- l_1 , 15° BF_fixed[15°]+BSS- l_1 , et 30° BF_fixed[30°]+BSS- l_1 , évaluation sur la base de données Theo-RI-studio.

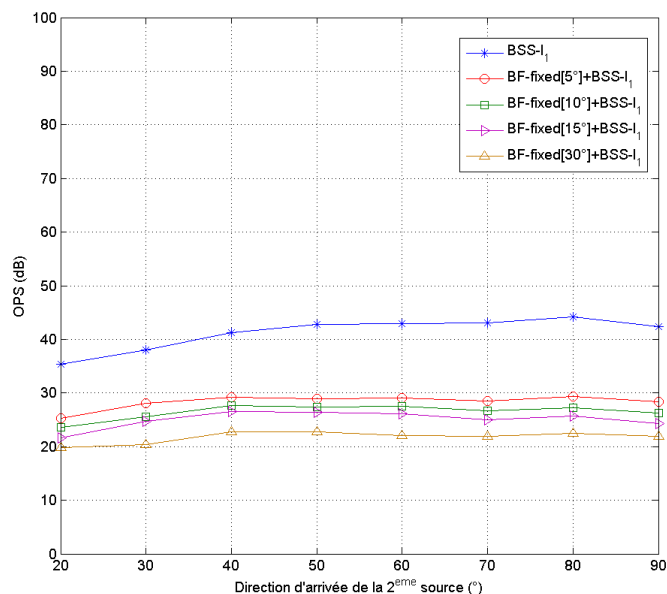


FIGURE 9.20 – Influence de l'angle inter-lobes de la formation de voies en termes de score perceptuel global OPS : angle inter-lobes égal à 5° BF_fixed[5°]+BSS- l_1 , 10° BF_fixed[10°]+BSS- l_1 , 15° BF_fixed[15°]+BSS- l_1 , et 30° BF_fixed[30°]+BSS- l_1 , évaluation sur la base de données Theo-RI-studio.

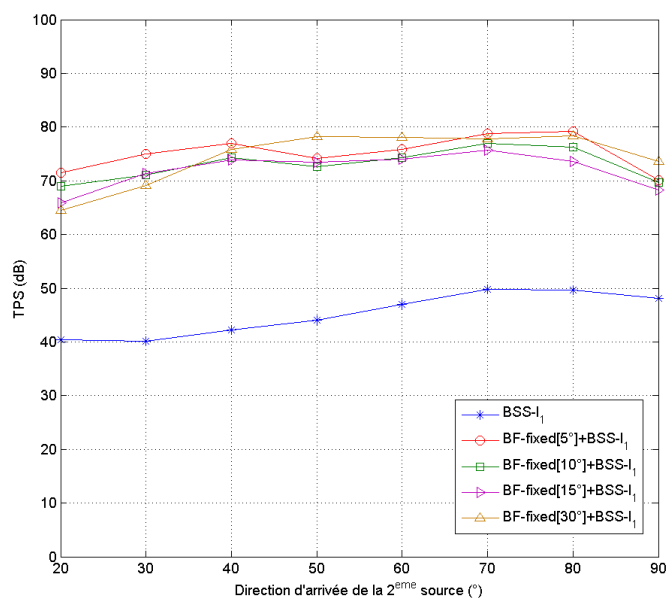


FIGURE 9.21 – Influence de l'angle inter-lobes de la formation de voies en termes de score perceptuel relatif à la cible TPS : angle inter-lobes égal à 5° BF_fixed[5°]+BSS- l_1 , 10° BF_fixed[10°]+BSS- l_1 , 15° BF_fixed[15°]+BSS- l_1 , et 30° BF_fixed[30°]+BSS- l_1 , évaluation sur la base de données Theo-RI-studio.

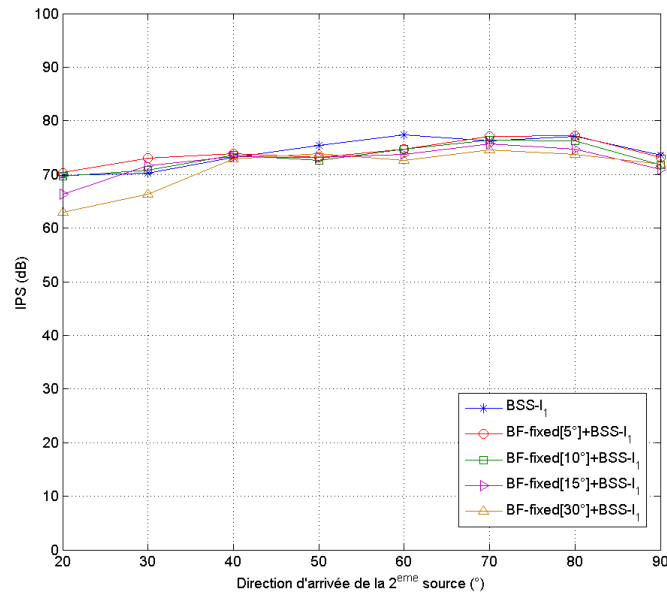


FIGURE 9.22 – Influence de l'angle inter-lobes de la formation de voies en termes de score perceptuel relatif aux interférences IPS : angle inter-lobes égal à 5° BF_fixed[5°]+BSS- l_1 , 10° BF_fixed[10°]+BSS- l_1 , 15° BF_fixed[15°]+BSS- l_1 , et 30° BF_fixed[30°]+BSS- l_1 , évaluation sur la base de données Theo-RI-studio.

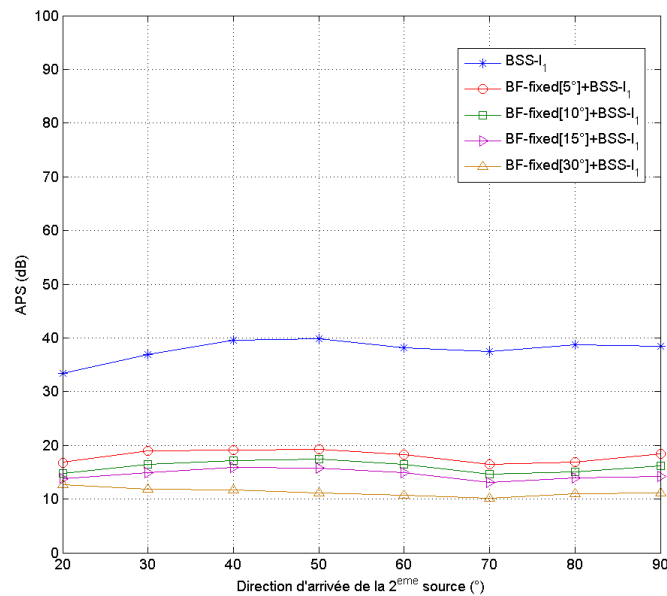


FIGURE 9.23 – Influence de l'angle inter-lobes de la formation de voies en termes de score perceptuel relatifs aux artéfacts APS : angle inter-lobes égal à 5° BF_fixed[5°]+BSS- l_1 , 10° BF_fixed[10°]+BSS- l_1 , 15° BF_fixed[15°]+BSS- l_1 , et 30° BF_fixed[30°]+BSS- l_1 , évaluation sur la base de données Theo-RI-studio.

Quant aux résultats perceptuels, ils confirment ce que nous avons constaté avec le prétraitement par formation de voies vers des directions fixes. Le prétraitement par formation de voies fixes, quel que soit l'angle entre les lobes, augmente le score perceptuel relatif à la source cible IPS par rapport à l'utilisation de $\text{BSS-}l_1$ seul (*cf.* figure 9.22). Par contre, le prétraitement par formation de voies fixe introduit des artéfacts comme le montre la figure 9.23 des scores perceptuels relatifs aux artéfacts APS, ce qui conduit à un score perceptuel global inférieur à celui de $\text{BSS-}l_1$ (*cf.* figure 9.20). Le score perceptuel relatif aux interférences est quant à lui presque le même avec ou sans prétraitement par formation de voies et pour tous les angles inter-lobes considérés.

Conclusion sur le prétraitement par formation de voies vers des directions de visée fixe :

Pour conclure sur cette partie, nous pouvons affirmer que le prétraitement par formation de voies vers des directions de visée fixes améliore les performances de séparation des sources audio par rapport à l'utilisation d'un algorithme de séparation seul et ceci pour tous les angles inter-lobes testés. Ceci est vrai pour toutes les mesures de performances à part le score perceptuel relatif aux artéfacts (APS) qui baisse lors de ce type de prétraitement. Néanmoins, avec le prétraitement par formation de voies vers des directions de visée fixes, nous n'atteignons pas encore les performances de séparations observées avec le prétraitement par formation de voies vers les directions d'arrivées.

9.3.3 Influence de la sélection de lobes

Comme nous pouvons le noter des figures 9.24, 9.25 9.26 et 9.31a, le prétraitement par formation de voies fixe avec sélection de lobes ($\text{BF_fixed}[30^\circ]_\text{BS}+\text{BSS-}l_1$ et $\text{BF_fixed}[5^\circ]_\text{BS}+\text{BSS-}l_1$) et le prétraitement par formation de voies fixe vers les directions d'arrivées des sources ($\text{BF_DOA}+\text{BSS-}l_1$) ont des performances proches en termes de rapport source-à-interférences SIR, de rapport source-à-distorsion SDR et de rapport sources-à-artéfacts SAR. Cependant, dans un milieu réverbérant où les directions d'arrivées ne peuvent pas être estimées avec précision, le prétraitement avec formation de voies fixe et sélection de lobes peut être une bonne solution pour améliorer le SIR et le SDR des sources estimées, par comparaison avec l'utilisation d'un algorithme de séparation seul ($\text{BSS-}l_1$).

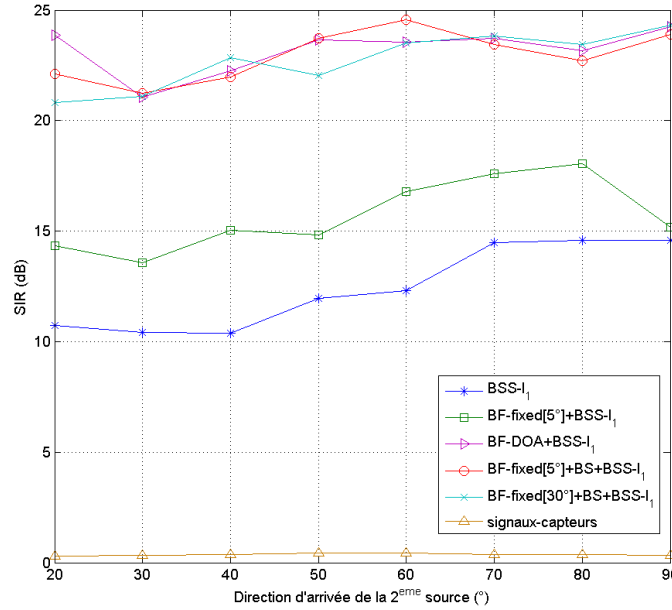


FIGURE 9.24 – Influence de la sélection de lobes en termes de rapport source-à-interférences SIR : évaluation sur la base de données Theo-RI-studio.

En comparant les algorithmes $\text{BF_fixed}[30^\circ]_\text{BS}+\text{BSS-}l_1$ et $\text{BF_fixed}[5^\circ]_\text{BS}+\text{BSS-}l_1$, nous pouvons noter que l'impact de l'angle inter-lobe est relativement faible par rapport au gain de performance et ceci grâce à la phase de sélection de lobes. Cependant, la formation de voies avec un angle inter-lobes de 5° nous permet d'estimer les directions d'arrivées des sources avec une résolution de 5° comme le montre la figure 9.32. Cette dernière représente une moyenne sur les lobes sélectionnés pour la base de données Theo-RI-studio, c'est à dire pour toutes les paires de sources à séparer et pour toutes les directions d'arrivées considérées.

L'analyse perceptuelle montre que la sélection de lobes améliore la qualité perceptuelle des sources séparées par rapport l'algorithme de séparation sans et avec prétraitement par formation de voies fixe. En effet, la figure 9.27 montre une amélioration du score perceptuel global OPS des sources séparées grâce à la phase de sélection de lobes. Plus particulièrement, le score perceptuel relatif à la source cible TPS (*cf.* figure 9.28), le score perceptuel relatif aux interférences IPS (*cf.* figure 9.29) et le score relatif aux artéfacts APS (*cf.* figure 9.30) ont augmenté. La sélection de lobes a aussi permis de réduire les écarts types des scores perceptuels TPS IPS et APS par rapport à ceux obtenus sans phase de sélection de lobes (*cf.* figure 9.31b).

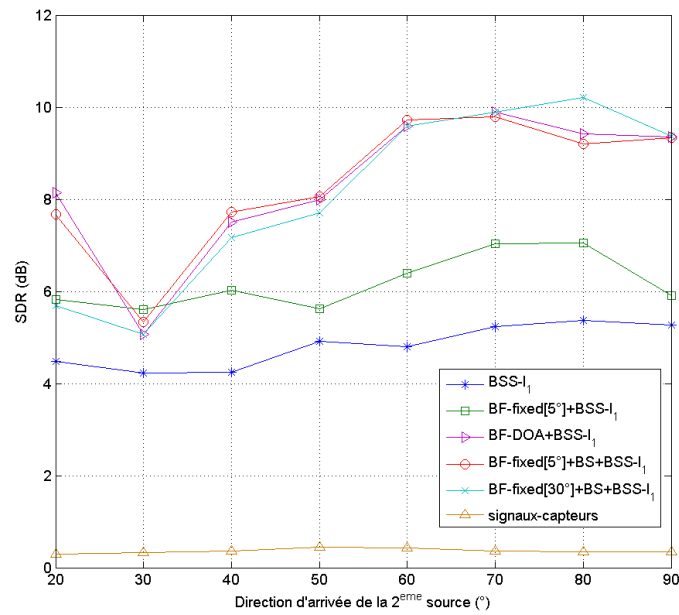


FIGURE 9.25 – Influence de la sélection de lobes en termes de rapport source-à-distorsion SDR : évaluation sur la base de données Theo-RI-studio.

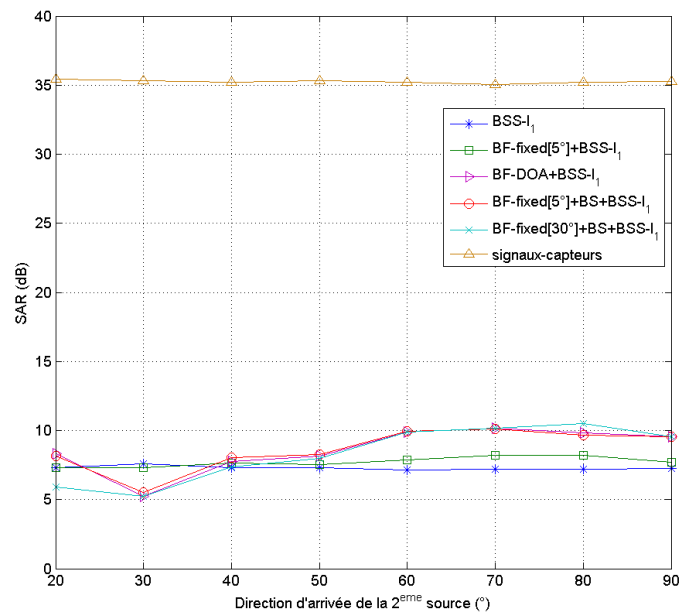


FIGURE 9.26 – Influence de la sélection de lobes en termes de rapport sources-à-artéfacts SAR : évaluation sur la base de données Theo-RI-studio.

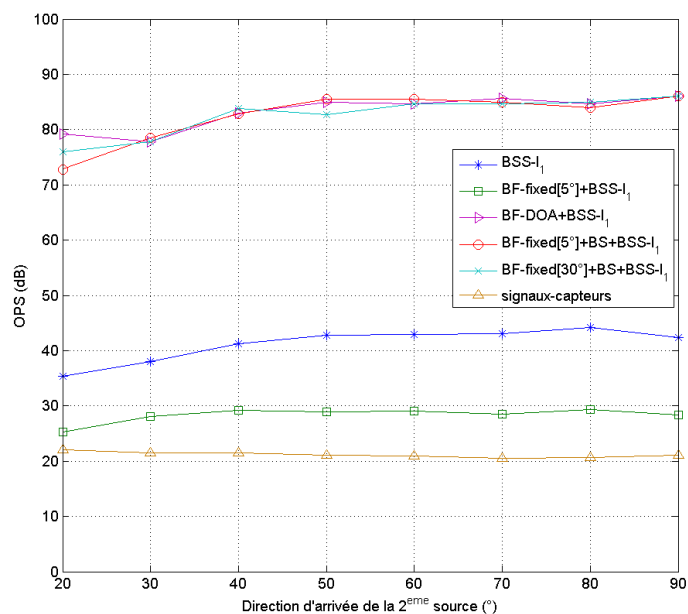


FIGURE 9.27 – Influence de la sélection de lobes en termes de score perceptuel global OPS : évaluation sur la base de données Theo-RI-studio.

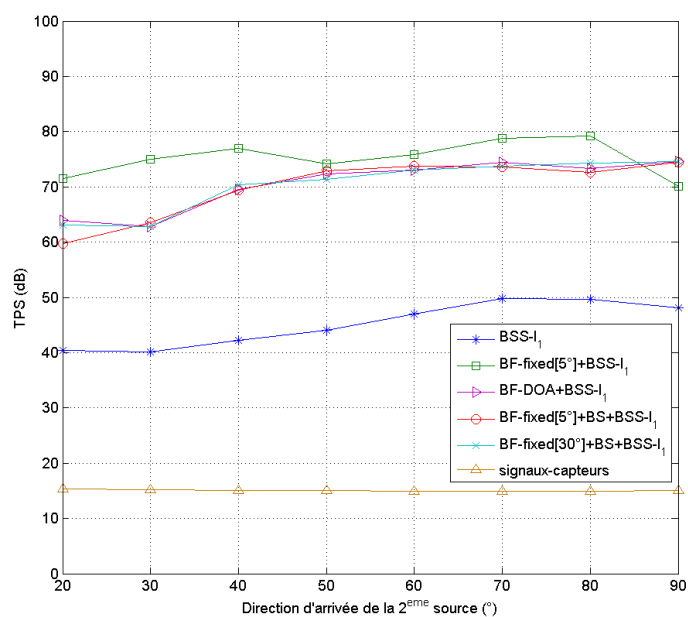


FIGURE 9.28 – Influence de la sélection de lobes en termes de score perceptuel relatif à la cible TPS : évaluation sur la base de données Theo-RI-studio.

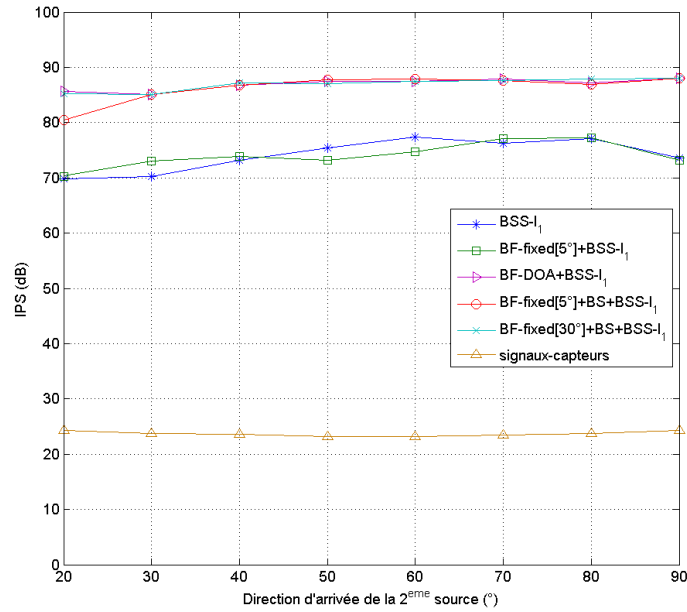


FIGURE 9.29 – Influence de la sélection de lobes en termes de score perceptuel relatif aux interférences IPS : évaluation sur la base de données Theo-RI-studio.

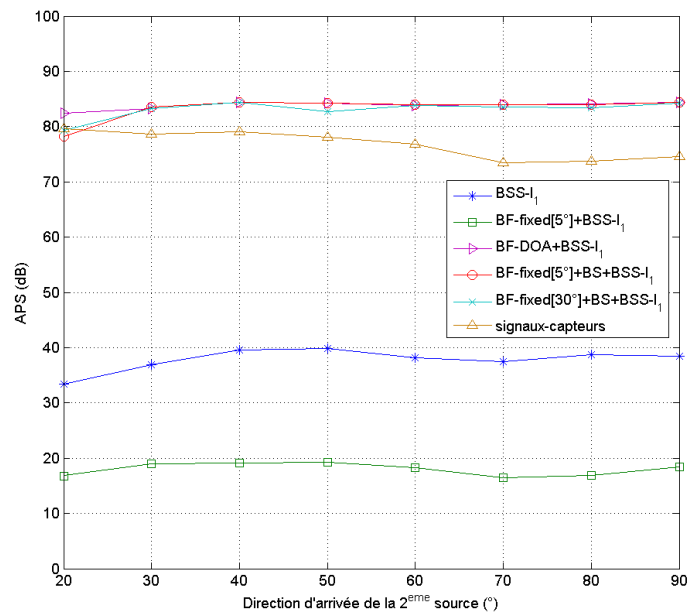


FIGURE 9.30 – Influence de la sélection de lobes en termes de score perceptuel relatifs aux artefacts APS : évaluation sur la base de données Theo-RI-studio.

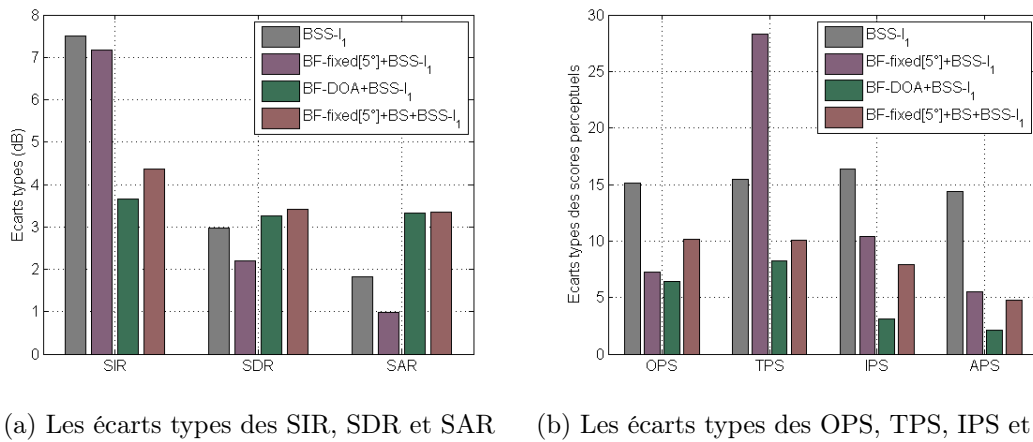


FIGURE 9.31 – Les écarts types des résultats de séparations des algorithmes $BSS-l_1$ (barres grises), $BF_fixed[5^\circ]+BSS-l_1$ (barres violettes), $BF_DOA+BSS-l_1$ (barres vertes) et $BF_fixed[5^\circ]_{BS}+BSS-l_1$ (barres rouges) : évaluation sur la base de données Theo-RI-studio.

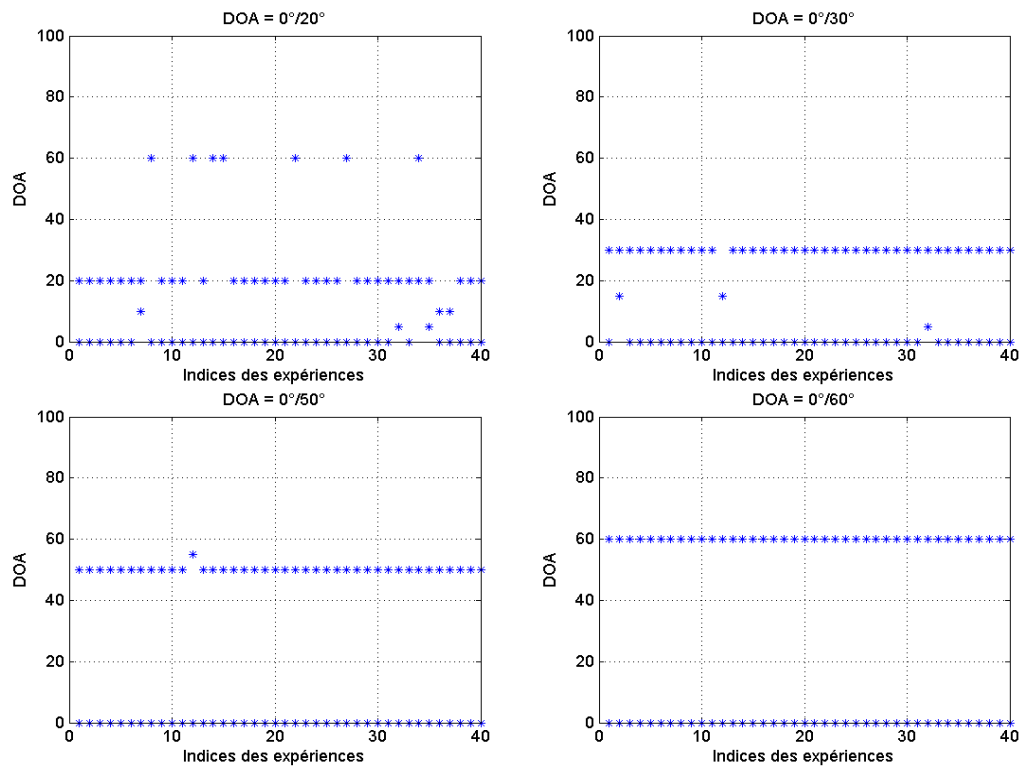


FIGURE 9.32 – Estimation des directions d'arrivées en utilisant l'algorithme de formation de voies $BF_fixed[5^\circ]_{BS}$ pour les 40 cas de séparation de source de la base de données Theo-RI-studio.

Les performances perceptuelles des algorithmes de séparation de sources avec prétraitement par formation de voies fixe et sélection de lobes (BF_fixed[30°]_BS+BSS- l_1 et BF_fixed[5°]_BS+BSS- l_1) restent proche de celles obtenues par l'algorithme avec prétraitement par formation de voies vers les directions d'arrivées des sources BF_DOA+BSS- l_1 .

9.3.4 Analyse de la convergence

Nous avons procédé à l'analyse de la convergence de l'algorithme de séparation de sources avec prétraitement de formation de voies fixe et sélection de lobes BF_fixed[5°]_BS+BSS- l_1 en observant sa vitesse de convergence à travers les itérations pour les directions d'arrivées considérées (*cf.* figure 9.33). Chaque courbe représente la moyenne sur toutes les fréquences de la fonction de coût normalisée $\psi(f) = \frac{\sum_{i=1}^N \sum_{k=1}^{N_T} |Y_i(f,k)|}{\max(\sum_{i=1}^N \sum_{k=1}^{N_T} |Y_i(f,k)|)}$ (*cf.* section 4.3).

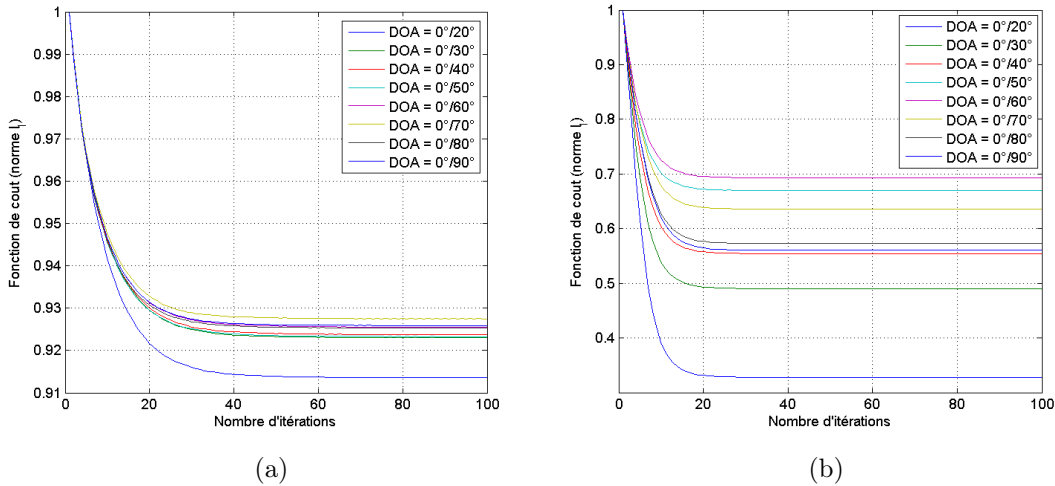


FIGURE 9.33 – Étude de la vitesse de convergence : la valeur de la fonction de coût de (a) BSS- l_1 et (b) BF_fixed[5°]_BS+BSS- l_1 à travers les itérations et pour différentes directions d'arrivées.

Comme nous pouvons le voir dans la figure 9.33b, notre algorithme itératif à deux étapes converge assez rapidement (typiquement entre 10 et 20 itérations) vers un état stable. Nous notons aussi que la vitesse de convergence de la méthode avec prétraitement par formation de voies et sélection de lobes est meilleure que celle de BSS- l_1 seul (*cf.* figure 9.33a). En effet, dans ce contexte, l'algorithme de séparation

$\text{BSS-}l_1$ converge à son état stable après 30 à 40 itérations. De plus, la fonction de coût de l'algorithme à deux étapes atteint des valeurs plus basses que celles atteintes par $\text{BSS-}l_1$ et par conséquent, le traitement par formation de voies conduit à une meilleure convergence.

9.4 Variation des performances de séparation avec le nombre de capteurs

Nous nous intéressons ici à l'étude des performances de la séparation de sources audio par rapport au nombre de capteurs [6]. Cette étude a été effectuée avec notre algorithme de séparation à deux étapes $\text{BF_fixed}[5^\circ]_\text{BS}+\text{BSS-}l_1$ en comparaison avec l'algorithme de la minimisation de la norme l_1 seul $\text{BSS-}l_1$. Par rapport à la méthode de séparation, notre but est d'étudier l'effet du nombre de capteurs sur la qualité de la séparation de sources. Plus spécifiquement, nous tentons de trouver le nombre optimal de capteurs qui doit être utilisé pour l'audition des robots, étant donnée une géométrie du réseau de capteurs. Nous montrons que dans notre cas, l'utilisation d'un réseau de capteurs augmente significativement les performances de séparation en comparaison avec le cas binaural (deux capteurs), cette augmentation se stabilise à partir d'un certain nombre de microphones que nous dévoilerons dans la suite.

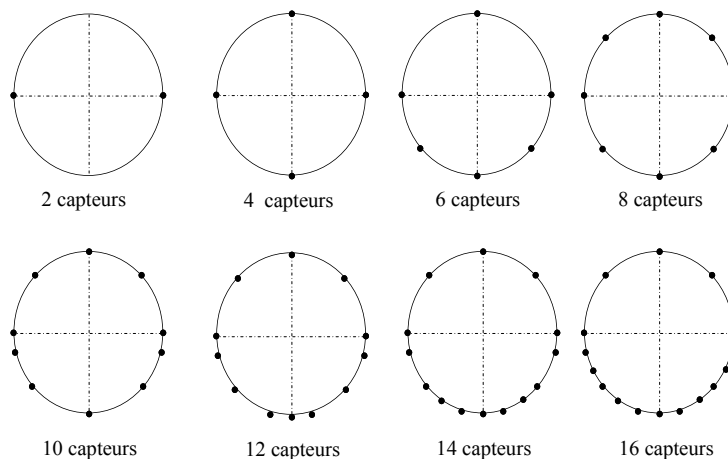


FIGURE 9.34 – Une vue de dessus des différentes configuration du réseau de capteurs : le nombre de microphones varie de 2 à 16.

Nous proposons d'évaluer le rapport source-à-interférences SIR, le rapport source-

à-distorsion SDR et le rapport sources-à-artéfacts SAR de la base de données **Theo-RI-studio** en variant le nombre de capteurs impliqué dans l'acquisition des mélanges de 2 à 16 et en considérant les configurations du réseau de microphones de la figure 9.34.

La figure 9.35 montre que, pour n'importe quel nombre de capteurs, la séparation aveugle de source avec un prétraitement par formation de voies et sélection de lobes **BF_fixed[5°]_BS+BSS- l_1** a de meilleures performances que l'algorithme de séparation seul **BSS- l_1** . Les figures 9.35 et 9.36 montrent la variation des performances objectives et perceptuelles de l'algorithme **BF_fixed[5°]_BS+BSS- l_1** avec le nombre de capteurs impliqué dans la séparation. Ces courbes représentent la moyenne des performances des cas de séparation de la base de données **Theo-RI-studio**. D'après ces courbes, les performances de séparation augmentent avec le nombre de capteurs.

La figure 9.37 montre la variation de la moyenne des SIR par rapport au nombre de capteurs. Nous pouvons noter une amélioration significative des performances quand le nombre de capteurs augmente. Quand le nombre de microphones M est supérieur ou égal à 8, il n'y a plus de gain significatif observé et ceci pour toutes les directions d'arrivées. Le même résultat est observé pour le SDR (*cf.* figure 9.38), SAR (*cf.* figure 9.39), et les scores perceptuels OPS (*cf.* figure 9.40), TPS (*cf.* figure 9.41), IPS (*cf.* figure 9.42) et APS (*cf.* figure 9.43). Le même phénomène de saturation des performances est observé pour la séparation de 3 et de 4 sources.

Ceci attire notre attention et peut être expliqué par : d'abord la réverbération modérée de la chambre dans laquelle les mesures ont été faites, nous nous attendons à un effet plus significatif de l'augmentation du nombre de capteurs dans un milieu plus réverbéré, ensuite il y a l'effet de la géométrie des lobes de la formation de voies. Nous notons aussi que plus la différence entre les directions d'arrivées est grande, plus la convergence par rapport au nombre de capteurs vers un état stable est rapide.

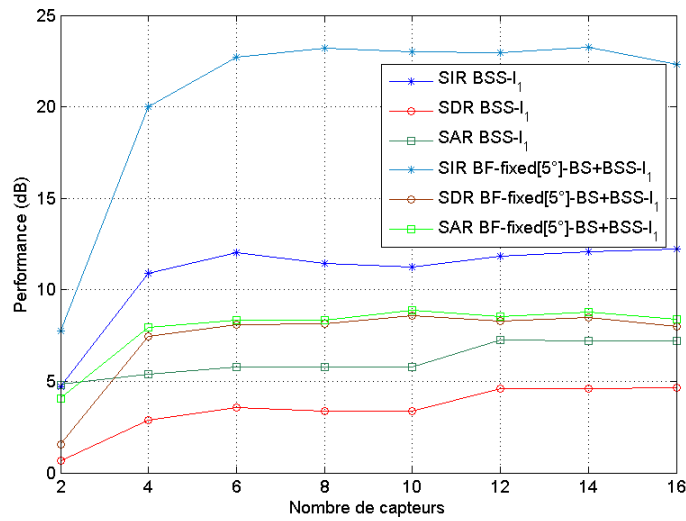


FIGURE 9.35 – Variation du rapport source-à-interférences SIR (courbes en bleu), le rapport source-à-distorsion SDR (courbes en rouge) et le rapport sources-à-artéfacts SAR (courbes en vert) avec le nombre de capteurs pour la séparation de deux sources à partir de différentes directions d'arrivées : évaluation de BF_fixed[5°]_BS+BSS- l_1 et BSS- l_1 : évaluation sur la base de données Theo-RI-studio.

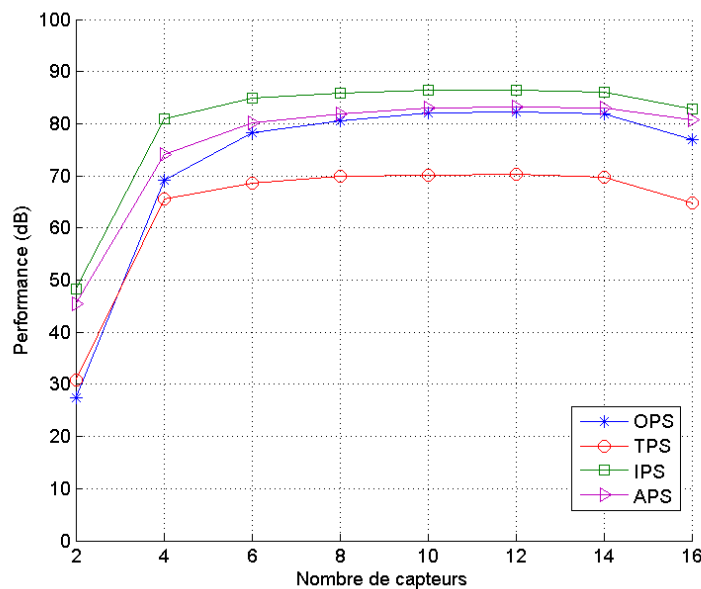


FIGURE 9.36 – Variation des mesures perceptuelles avec le nombre de capteurs pour la séparation de deux sources à partir de différentes directions d'arrivées : évaluation de BF_fixed[5°]_BS+BSS- l_1 : évaluation sur la base de données Theo-RI-studio.

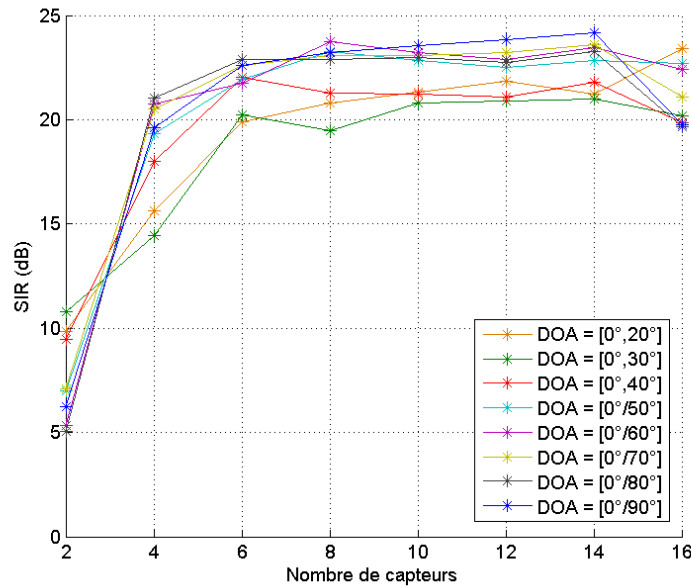


FIGURE 9.37 – Variation du rapport source-à-interférences SIR avec le nombre de capteurs et pour différentes directions d'arrivées pour l'algorithme $\text{BF_fixed}[5^\circ]_\text{BS+BSS-}l_1$: évaluation sur la base de données Theo-RI-studio.

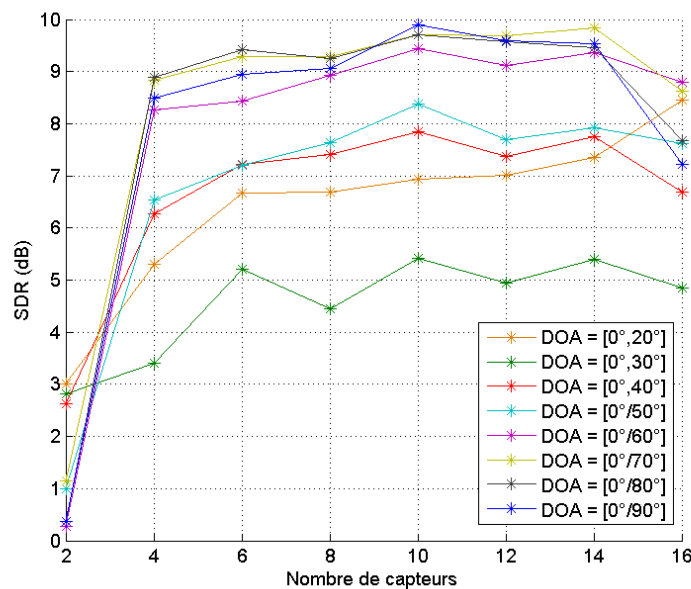


FIGURE 9.38 – Variation du rapport source-à-distorsion SDR avec le nombre de capteurs et pour différentes directions d'arrivées pour l'algorithme $\text{BF_fixed}[5^\circ]_\text{BS+BSS-}l_1$: évaluation sur la base de données Theo-RI-studio.

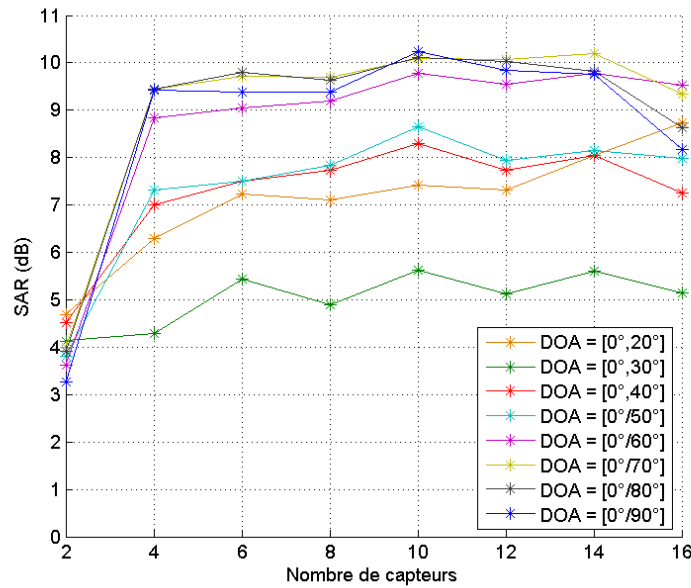


FIGURE 9.39 – Variation du rapport sources-à-artéfacts SAR avec le nombre de capteurs et pour différentes directions d'arrivées pour l'algorithme $\text{BF_fixed}[5^\circ]_\text{BS+BSS-}l_1$: évaluation sur la base de données Theo-RI-studio.

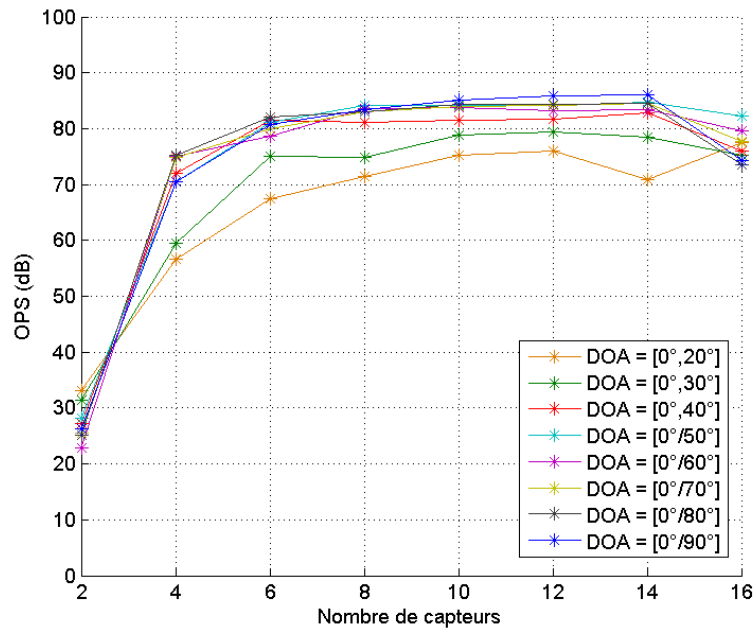


FIGURE 9.40 – Variation du score perceptuel global OPS avec le nombre de capteurs et pour différentes directions d'arrivées pour l'algorithme $\text{BF_fixed}[5^\circ]_\text{BS+BSS-}l_1$: évaluation sur la base de données Theo-RI-studio.

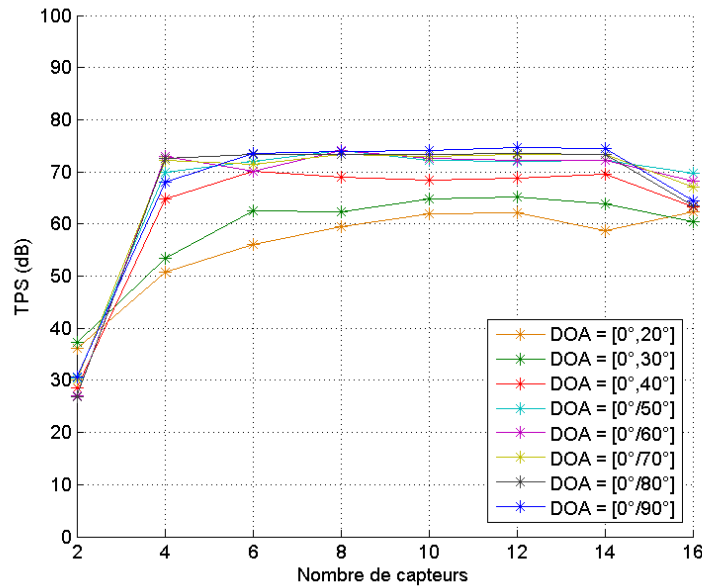


FIGURE 9.41 – Variation du score perceptuel relatif à la cible TPS avec le nombre de capteurs et pour différentes directions d'arrivées pour l'algorithme $\text{BF_fixed}[5^\circ]_{\text{BS+BSS-}l_1}$: évaluation sur la base de données Theo-RI-studio.

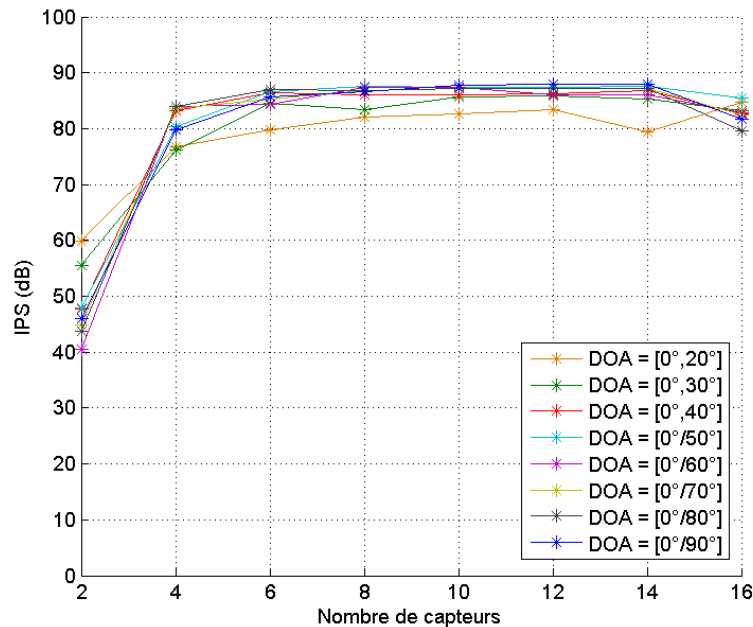


FIGURE 9.42 – Variation du score perceptuel relatif aux interférences IPS avec le nombre de capteurs et pour différentes directions d'arrivées pour l'algorithme $\text{BF_fixed}[5^\circ]_{\text{BS+BSS-}l_1}$: évaluation sur la base de données Theo-RI-studio.

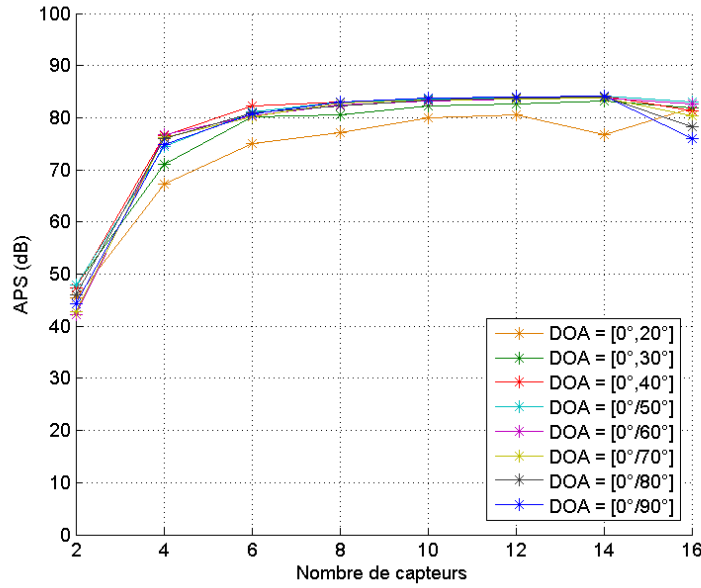


FIGURE 9.43 – Variation du score perceptuel relatif aux artéfacts APS avec le nombre de capteurs et pour différentes directions d'arrivées pour l'algorithme $\text{BF_fixed}[5^\circ]_\text{BS+BSS-}l_1$: évaluation sur la base de données Theo-RI-studio.

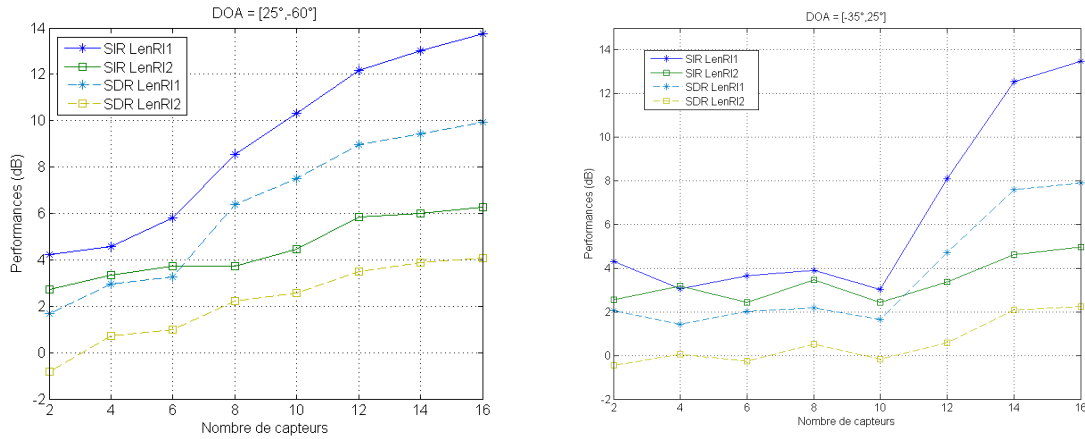


FIGURE 9.44 – Variation du rapport source-à-interférences SIR et du rapport source-à-distorsion SDR de l'algorithme $\text{BF_fixed}[5^\circ]_\text{BS+BSS-}l_1$ avec le nombre de capteurs : évaluation sur la base de données Theo-RI-IDV avec deux longueurs de réponses impulsionnelles : $\text{RI1} = 500$ échantillons et $\text{RI2} = 800$ échantillons.

La figure 9.44 montre les courbes de rapport source-à-interférences SIR et du rapport source-à-distorsion SDR de $\text{BF_fixed}[5^\circ]_\text{BS+BSS-}l_1$ en fonction de la variation du nombre de capteurs évaluées sur la base de données enregistrée à l'Institut de la Vision Theo-RI-IDV et pour deux différentes longueurs de réponses impulsionnelles :

RI1 = 500 échantillons et RI2 = 800 échantillons. La figure montre que les performances de séparation en termes de SIR et SDR ne semblent pas être stationnaire après un certain nombre de capteurs mais continuent à s'améliorer. Ce résultat ne va pas dans le sens de ce que nous avons observé pour la base de données **Theo-RI-studio** enregistrée dans le studio de Télécom ParisTech. Ceci peut s'expliquer par le fait que **Theo-RI-IDV** soit plus réverbérante que **Theo-RI-studio**. Nous pouvons supposer que si nous augmentons encore plus le nombre de capteurs (plus que 16 capteurs), nous pourrions observer le même phénomène de « saturation » des résultats de séparation observés dans le cas d'un environnement moins réverbérant.

A la figure 9.45, nous évaluons la vitesse de convergence de l'algorithme de séparation de sources avec prétraitement par formation de voies par rapport à la variation du nombre de capteurs. Nous notons que cette convergence est pratiquement indépendante de la taille du réseau de capteurs.

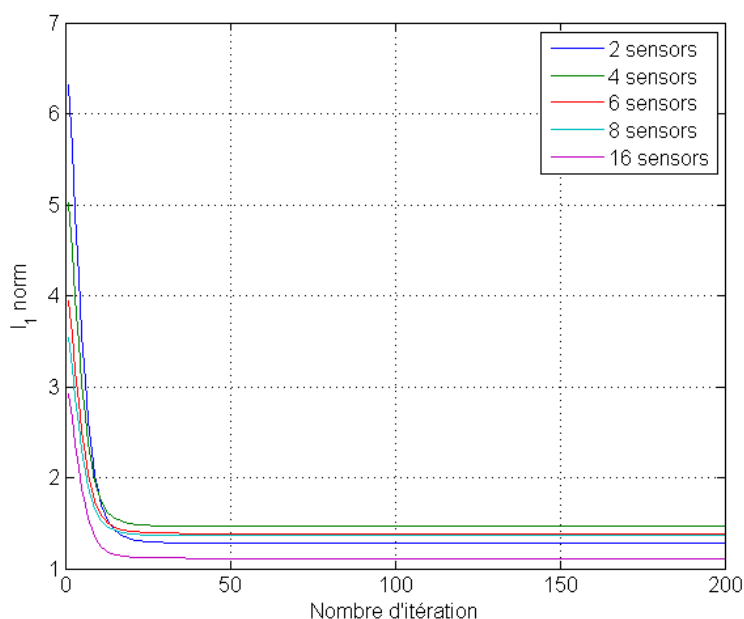


FIGURE 9.45 – La moyenne de la norme l_1 au cours des itérations de l'algorithme **BF_fixed[5°]_BS+BSS- l_1** pour différentes tailles du réseau de capteurs : évaluation sur la base de données **Theo-RI-studio**.

Conclusion

Dans ce chapitre, nous avons présenté les résultats des algorithmes itératifs de séparation de sources que nous avons développés. Les principaux résultats de ce chapitre sont :

- la séparation de sources avec un critère de parcimonie basé sur la minimisation de la norme l_1 a d'aussi bonnes performances que l'analyse en composantes indépendantes basé sur la théorie de l'information ;
 - la paramétrisation de la pseudo-norme l_p afin de rendre la contrainte de parcimonie de plus en plus dure au cours des itérations a des résultats prometteurs ;
 - le prétraitement par formation de voies fixe utilisant les HRTFs améliore les performances moyenne de séparation (de 3dB maximum) ; en particulier le prétraitement par formation de voies fixes avec sélection de lobes améliore aussi bien la qualité objective (de 10 à 15 dB) que la qualité perceptuelle de la séparation (de 40 points environ) ;
 - le nombre de capteurs nécessaire pour obtenir de bons résultats de séparation varie selon l'acoustique de l'environnement dans lequel se passe la séparation : pour la séparation d'un même mélange de sources, plus le milieu dans lequel nous effectuons cette séparation est réverbérant, plus le nombre de capteurs nécessaire pour obtenir une bonne séparation est grand.
-

