



Fès, le

N° d'ordre : 12/2016

THESE DE DOCTORAT

Présentée par

Mr: Omar Cherrak

Discipline : Génie électrique
Spécialité : Génie électrique

Séparation aveugle de mélanges linéaires de sources : application à la surveillance maritime

Thèse en co-tutelle présentée et soutenue le 19 mars 2016 devant le jury composé de

Nom Prénom	Titre	Etablissement	
MARS Jérôme	PES	INP Grenoble, France	Président
BADRI Abdelmajid	PES	FST Mohammedia	Rapporteur
BOURRENANE Salah	PES	Ecole centrale de Marseille, France	Rapporteur
ABDI Farid	PES	FST Fès	Rapporteur
GHENNIOUI Hicham	PA	FST Fès	Invité
ABARKAN El Hossain	PES	FST Fès	Directeur de thèse
THIRION-MOREAU Nadège	PES	Université de Toulon	Co-directrice de thèse

Laboratoires d'accueil : -- Signaux, Systèmes et Composantes (LSSC)
-- Sciences de l'Information et des Systèmes (LSIS)



Etablissements : -- Faculté des Sciences et Techniques (FST)
-- Université de Toulon (UTLN)



Remerciements

Réussir à exprimer toute ma gratitude envers ces personnes avec les mots exacts me semble à ce moment là bien plus compliqué que de rédiger une thèse. Toutefois, ce travail n'aurait pas pu aboutir sans l'encouragement direct ou indirect de certaines personnes à qui je rends hommage par le présent manuscrit.

Tout d'abord, mes profonds et sincères remerciements s'adressent à mon directeur de thèse, le Professeur El Hossain Abarkan pour m'avoir contacté et accueilli dans son bureau pour me proposer d'effectuer une thèse dans son équipe, pour ses encouragements permanents, sa patience, son écoute et ses conseils. Je lui suis également très reconnaissant pour les nombreuses remarques, suggestions et l'intérêt qu'il a porté à ce travail. Un grand merci pour tout ce que l'on a pu échanger aussi bien en paternité qu'en termes de discussions scientifiques.

J'adresse mes chaleureux remerciements à ma co-Directrice de thèse, Madame Nadège Thirion-Moreau, pour son encadrement maternel, ses nombreux conseils avisés qui m'ont permis de recevoir une formation très riche, sa gentillesse, sa patience, son accueil qui m'a permis d'achever ma thèse dans les meilleures conditions et surtout l'hospitalité qu'elle m'a offerte lors de mes séjours en France. Un grand merci pour la grande attention qu'elle a portée à plusieurs manuscrits incluant ce document de thèse.

Je tiens à exprimer toute ma gratitude à mon co-encadrant, le Professeur Hicham Ghennioui pour ses nombreux conseils, sa bonne humeur, sa disponibilité, son encadrement très amical et surtout pour m'avoir fait bénéficier ses idées éclairées qui m'ont souvent permis de résoudre les problèmes auxquels j'étais confronté. Je le remercie profondément de m'avoir initié au domaine de la séparation aveugle de sources. Je lui témoigne toute ma reconnaissance pour la confiance qu'il m'a accordée depuis le début. Mille fois merci pour son apport marquant à la correction de plusieurs manuscrits dont ce manuscrit de thèse.

Je ne vous remercierai jamais assez, tous les trois, de m'avoir initié à la recherche. Vos qualités humaines et scientifiques ainsi que votre expérience dans le domaine de la recherche ont été de précieux atouts durant mes années de thèse. J'espère pouvoir me montrer dans le futur, à la hauteur de la confiance et du soutien sincère que vous m'avez adressé tant sur le plan scientifique qu'humain.

Je voudrais remercier le Professeur Lahbabi M'hammed, directeur du Laboratoire Signaux Systèmes et Composants de l'Université Sidi-Mohamed Ben Abdellah (USMBA) de Fès et Monsieur Mustapha Ouladsine, directeur du Laboratoire des Sciences de l'Information et des Systèmes d'Aix Marseille Université (AMU) et de l'Université de Toulon (UTLN) pour leur accueil au sein du LSSC et du LSIS et de m'avoir permis aussi d'achever ma thèse dans les meilleures conditions possibles.

Merci à Monsieur Salah Bourennane, Monsieur Abdelmajid Badri et Monsieur Farid Abdi pour avoir accepté d'être rapporteurs. Je leur suis reconnaissant pour les nombreuses remarques pertinentes qu'ils ont apportées dans l'élaboration de la version finale de ce manuscrit.

Je remercie également Monsieur Jérôme Mars, Professeur à l'INP Grenoble de m'avoir fait l'honneur d'être président du jury de ma thèse et d'avoir porté ce rôle avec autant d'intérêt.

De nombreuses pensées envers toutes mes collègues extraordinaires de Fès et de Toulon avec lesquelles j'ai passé de très bons moments de convivialité et ai pu nouer des contacts et des amitiés et avec qui je garderai de très bons souvenirs.

Clin d'oeil à tous les frères et soeurs que j'ai choisis (sans oublier ZedKa bien sûr) qui m'ont aidé et encouragé dans les moments de difficulté et que je n'ai pas oublié mais, qui feraient multiplier sans aucun doute la taille de ce document si je devais tous les nommer (je vois d'ici leur réaction).

Enfin, il me serait impossible de terminer mes remerciements sans adresser d'énormes et chaleureuses pensées à ma petite famille : ma mère, mon père, ma soeur et son époux, pour leur soutien inépuisable dans les moments difficiles et de doute. Je leur exprime toute ma reconnaissance et leurs dédie ce manuscrit sans oublier ma petite nièce Rita. Je vous remercie éternellement pour votre confiance, tendresse, soutien et amour qui me portent et me guident tous les jours. Merci mes parents pour avoir fait de moi ce que je suis aujourd'hui. Je vous aime.

Résumé

Dans cette thèse, nous nous intéressons au système d'identification automatique spatial lequel est dédié à la surveillance maritime par satellite. Ce système couvre une zone bien plus large que le système standard à terre correspondant à plusieurs cellules traditionnelles ce qui peut entraîner des risques de collision des données envoyées par des navires localisés dans des cellules différentes et reçues au niveau de l'antenne du satellite. Nous présentons différentes approches afin de répondre au problème de collision considéré. Elles ne reposent pas toujours sur les mêmes hypothèses en ce qui concerne les signaux reçus, et ne s'appliquent donc pas toutes dans les mêmes contextes (nombre de capteurs utilisés, mode semi-supervisé avec utilisation de trames d'apprentissage et information a priori ou mode aveugle, problèmes liés à la synchronisation des signaux, etc...).

Dans un premier temps, nous proposons des méthodes permettant la séparation/dé-collision des messages en modèle sur-déterminé (plus de capteurs que de messages). Elles sont fondées sur des algorithmes de décompositions matricielles conjointes combinés à des détecteurs de points temps-fréquence (retard-fréquence Doppler) particuliers permettant la construction d'ensembles de matrices devant être (bloc) ou zéro (bloc) diagonalisées conjointement. En ce qui concerne les algorithmes de décompositions matricielles conjointes, nous proposons quatre nouveaux algorithmes de bloc diagonalisation conjointe (de même que leur version à pas optimal) fondés respectivement sur des algorithmes d'optimisation de type gradient conjugué, gradient conjugué pré-conditionné, Levenberg-Marquardt et Quasi-Newton. Le calcul exact du gradient matriciel complexe et des matrices Hessiennes complexes est mené. Nous introduisons également un nouveau problème dénommé zéro-bloc diagonalisation conjointe non-unitaire lequel généralise le problème désormais classique de la zéro-diagonalisation conjointe non-unitaire. Il implique le choix d'une fonction de coût adaptée et à nouveau le calcul de quantités telles que gradient matriciel complexe et les matrices Hessiennes complexes. Nous proposons ensuite trois nouveaux algorithmes à pas optimal fondés sur des algorithmes d'optimisation de type gradient conjugué, gradient conjugué pré-conditionné et Levenberg-Marquardt.

Finalement, nous terminons par des approches à base de techniques de détection multi-utilisateurs conjointe susceptibles de fonctionner en contexte sous-déterminé dans lequel nous ne disposons plus que d'un seul capteur recevant simultanément plusieurs signaux sources. Nous commençons par développer une première approche par déflation consistant à supprimer successivement les interférences. Nous proposons ensuite un deuxième mode opératoire fondé quant à lui sur l'estimateur du maximum de vraisemblance conjoint qui est une variante de l'algorithme de VITERBI.

Mots clés

Système d'identification automatique spatial, surveillance maritime, (dé)-collision, décompositions matricielles conjointes, algorithmes de bloc et zéro-bloc diagonalisation conjointe, séparation aveugle de sources, détection automatique, spectres et distributions spatiaux (spatiales) bilinéaires et quadratiques, représentation temps-fréquence, détection multi-utilisateurs conjointe, détection par suppression successive des interférences, estimateur de séquence à maximum de vraisemblance conjoint.

Abstract

This PHD thesis concerns the spatial automatic identification system dedicated to marine surveillance by satellite. This system covers a larger area than the traditional system corresponding to several satellite cells. In such a system, there are risks of collision of the messages sent by vessels located in different cells and received at the antenna of the same satellite. We present different approaches to address the considered problem. They are not always based on the same assumptions regarding the received signals and are not all applied in the same contexts (they depend on the number of used sensors, semi-supervised mode with use of training sequences and a priori information versus blind mode, problems with synchronization of signals, etc.). Firstly, we develop several approaches for the source separation/de-collision in the over-determined case (more sensors than messages) using joint matrix decomposition algorithms combined with detectors of particular time-frequency (delay-Doppler frequency) points to build matrix sets to be joint (block) or zero (block) diagonalized. Concerning joint matrix decomposition algorithms, four new joint block-diagonalization algorithms (with optimal step-size) are introduced based respectively on conjugate gradient, preconditioned conjugate gradient, Levenberg-Marquardt and Quasi-Newton optimization schemes. Secondly, a new problem called non-unitary joint zero-block diagonalization is introduced. It encompasses the classical joint zero diagonalization problem. It involves the choice of a well-chosen cost function and the calculation of quantities such as the complex gradient matrix and the complex Hessian matrices. We have therefore proposed three new algorithms (and their optimal step-size version) based respectively on conjugate gradient, preconditioned conjugate gradient and Levenberg-Marquardt optimization schemes. Finally, we suggest other approaches based on multi-user joint detection techniques in an under-determined context where we have only one sensor receiving simultaneously several signals. First, we have developed an approach by deflation based on a successive interferences cancelation technique. Then, we have proposed a second method based on the joint maximum likelihood sequence estimator which is a variant of the VITERBI algorithm.

Key words

spatial automatic identification system, marine surveillance, collision avoidance (de-collision), joint matrix decompositions, joint block and zero-bloc diagonalization algorithms, blind source separation, automatic detection, spatial bilinear and quadratic time-frequency spectra and distributions, time-frequency representations, joint multi-user detection, successive interference cancelation detection, joint maximum likelihood sequence estimator.

Table des matières

Remerciements	iii
Résumé	v
Abstract	vii
Table des figures	xiii
Liste des tableaux	xvii
Liste des acronymes	xix
Glossaire	xxiii
Introduction générale	1
1 Séparation aveugle de sources et surveillance maritime par satellite	5
1.1 Éléments classiques de la séparation aveugle de sources	5
1.1.1 Modélisation mathématique du problème	6
1.1.2 Modélisation du système de mélange linéaire	6
1.1.3 Hypothèses classiques de la séparation de sources	9
1.1.4 Blanchiment d'ordre deux	10
1.1.5 Indéterminations inhérentes au problème de la séparation aveugle de sources	11
1.1.6 Indice d'erreur	13
1.2 Méthodes de séparation de sources : un bref historique	14
1.2.1 Mélange instantané	14
1.2.2 Mélange convolutif	21
1.3 Quelques applications de la séparation aveugle de sources	22
1.4 Le contexte applicatif de la surveillance maritime : le système d'identification automatique	23

1.4.1	Introduction	23
1.4.2	Principe de fonctionnement du système AIS	24
1.4.3	Structure des messages AIS	27
1.4.4	Codage différentiel en ligne NRZI	28
1.4.5	Modulation GMSK	28
1.5	Réception des messages au niveau d'un récepteur satellitaire	29
1.5.1	Introduction	29
1.5.2	Problèmes liés à une réception spatiale des messages	30
1.5.3	Modélisation du problème de collision des messages AIS	32
1.6	Conclusion	34
2	Problèmes de la bloc-diagonalisation et de la zéro bloc-diagonalisation conjointe et nouveaux algorithmes	37
2.1	Bloc-diagonalisation conjointe : problème et hypothèses	39
2.1.1	Formulation du problème de la bloc-diagonalisation conjointe	40
2.1.2	Principe de la bloc-diagonalisation conjointe	40
2.1.3	Lien entre BDC et DC	42
2.1.4	Quelques rappels sur l'optimisation et le préconditionnement	42
2.1.5	Quatre nouveaux algorithmes de bloc-diagonalisation conjointe non unitaire	44
2.1.6	Schéma d'optimisation déterministes itératifs classiques	46
2.1.7	Algorithmes de BDC_{NU} en résumé	50
2.1.8	Complexité algorithmique	50
2.2	Zéro bloc-diagonalisation conjointe : problème et hypothèses	52
2.2.1	Formulation du problème de la zéro bloc-diagonalisation conjointe et ses hypothèses	53
2.2.2	Principe de la zéro-bloc diagonalisation conjointe	53
2.2.3	Lien entre ZBDC et ZDC	54
2.2.4	Trois nouveaux algorithmes de zéro-bloc diagonalisation conjointe non unitaire	55
2.2.5	Gradient matriciel complexe	55
2.2.6	Algorithmes de ZBDC_{NU} en résumé	58
2.2.7	Complexité algorithmique	58
2.3	Simulations informatiques et comparaisons	59
2.3.1	Cas général de la BDC d'un ensemble de matrices complexes	60
2.3.2	Cas particulier : la DC d'un ensemble de matrices complexes	68
2.3.3	Cas général de la ZBDC d'un ensemble de matrices complexes	68

2.3.4	Cas particulier : la ZDC d'un ensemble de matrices complexes	78
2.4	Conclusion	81
3	Nouvelles méthodes de séparation en surveillance maritime par satellite : cas sur-déterminé	83
3.1	Approches basées sur les algorithmes de décomposition matricielle conjointe .	84
3.2	Approches exploitant la diversité temps-fréquence des signaux $S - AIS$	85
3.2.1	Approches basées sur l'utilisation combinée de spectres temps-fréquence spatiaux quadratiques et d'algorithmes de diagonalisation conjointe .	85
3.2.2	Énoncé du problème	86
3.2.3	Principe des approches $DC_{LM_{EH}-TF}$, $DC_{GCP_{EH}-TF}$ et DC_{GC-TF}	86
3.2.4	Détecteur de "bons" points temps-fréquences	87
3.2.5	Résumé des approches proposées	87
3.2.6	Simulations informatiques	87
3.3	Approches exploitant la diversité Doppler-Retard des sources	90
3.3.1	Approches exploitant le caractère cyclo-stationnaire à l'ordre deux des sources $S - AIS$	90
3.3.2	Les outils de ce domaine : fonction d'ambiguïté et fonction d'ambiguïté moyenne spatiale	91
3.3.3	Sous l'effet d'un mélange linéaire instantané	93
3.3.4	Approches de séparation basées sur les fonctions d'ambiguïté moyennes spatiales combinées à des algorithmes de zéro-diagonalisation conjointe non unitaire	94
3.3.5	Construction de l'ensemble $\mathcal{X}$ de matrices à partir des points Doppler-Retard sélectionnés	95
3.3.6	Un récapitulatif du principe des approches suggérées	95
3.3.7	Simulations informatiques	95
3.4	Conclusion	98
4	Nouvelles méthodes de séparation en surveillance maritime par satellite : cas sous-déterminé	101
4.1	Technique de dé-collision fondée sur la méthode de suppression successive des interférences SSI	102
4.1.1	Principe de l'approche $DMUS_{SSI}$	102
4.1.2	Approche $DMUS_{SSI}$ en résumé	103
4.1.3	Simulations informatiques	103
4.2	Technique de dé-collision à base de l'algorithme de $JMLSE$	105

4.2.1	Principe de l'approche $\text{DMUC}_{\text{JMLSE}}$	105
4.2.2	Approche $\text{DMUC}_{\text{JMLSE}}$ en résumé	107
4.2.3	Simulations informatiques	107
4.3	Conclusion	110
Conclusions et perspectives		113
Bibliographie		115
Annexe 1 : Quelques propriétés algébriques importantes		133
Annexe 2 : Fonction de coût $\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)$ et bloc diagonalisation conjointe		135
1	Calcul exact des matrices Hessiennes complexes de la fonction de coût $\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)$	135
2	Calcul du pas optimal	137
Annexe 3 : Fonction de coût $\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)$ et zéro-bloc diagonalisation conjointe		139
1	Calcul exact des gradients matriciels de la fonction de coût $\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)$. .	139
2	Calcul exact des matrices Hessiennes complexes de la fonction de coût $\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)$	141
3	Calcul du pas optimal	143
Annexe 4 : Bref état de l'art sur la détection automatique de points (α, β) et sur les méthodes de SAS en contexte cyclo-stationnaire		145
1	Bref état de l'art en matière de détection automatique de points (α, β) . . .	145
1.1	Détecteurs de points (α_i, β_i) particuliers	146
1.2	En contexte instantané non blanchi	148
1.3	En contexte convolutif blanchi	150
1.4	En contexte convolutif non blanchi	151
2	Résumé de l'état de l'art sur les méthodes de SAS en contexte cyclo-stationnaire	151
Documents annexes		153

Table des figures

1.1	Schéma général du modèle de mélange convolutif de sources.	7
1.2	Principe du protocole SO – TDMA utilisé pour le système AIS.	27
1.3	Structure de la trame AIS.	27
1.4	Dessin synoptique résumant les différentes étapes de la modulation GMSK.	29
1.5	Scénario de la détection des signaux AIS par un satellite de type LEO qui traite les signaux AIS et les envoie aux stations côtières à proximité.	30
1.6	Scénario de signaux AIS en provenance de cellules SO – TDMA différentes mais reçus par le satellite au même instant (collision).	32
2.1	Comparaison des algorithmes $BDC_{GCP_{BFGS-HA}}$, $BDC_{GCP_{DFP-HIA}}$, $BDC_{QN_{BFGS-HA}}$, $BDC_{QN_{DFP-HA}}$, $BDC_{GCP_{HE}}$, BDC_{GC} , $BDC_{QN_{HE}}$, $BDC_{LM_{HE}}$, $BDC_{LM_{HA}}$, BDC_{GradR} , $BDC_{App-Grad}$, $BDC_{GradAbs}$, BDC_{Alg} et $BDC_{GC_{Nion}}$ dans un contexte quasi non-bruité ($RSB = 100$ dB) pour le cas d’une matrice $\mathbf{A}$ carrée. Gauche : $l_{Conv}(\mathbf{G})$ en dB en fonction du nombre d’itérations pour un ensemble $\mathcal{M}$ de taille $N_m = 100$ matrices. Droite : $l_{Conv}(\mathbf{G})$ en dB en fonction du nombre d’itérations pour un ensemble $\mathcal{M}$ de taille $N_m = 5$ matrices.	61
2.2	Comparaison des algorithmes $BDC_{GCP_{BFGS-HA}}$, $BDC_{GCP_{DFP-HIA}}$, $BDC_{QN_{BFGS-HA}}$, $BDC_{QN_{DFP-HA}}$, $BDC_{GCP_{HE}}$, BDC_{GC} , $BDC_{QN_{HE}}$, $BDC_{LM_{HE}}$, $BDC_{LM_{HA}}$, BDC_{GradR} , $BDC_{App-Grad}$, $BDC_{GradAbs}$, BDC_{Alg} et $BDC_{GC_{Nion}}$ dans un contexte bruité ($RSB = 30$ dB) pour le cas d’une matrice $\mathbf{A}$ carrée. Gauche : $l_{Conv}(\mathbf{G})$ en dB en fonction du nombre d’itérations pour un ensemble $\mathcal{M}$ de taille $N_m = 100$ matrices. Droite : $l_{Conv}(\mathbf{G})$ en dB en fonction du nombre d’itérations pour un ensemble $\mathcal{M}$ de taille $N_m = 5$ matrices.	61
2.3	Initialisation aléatoire : comparaison des algorithmes $BDC_{GCP_{BFGS-HA}}$, $BDC_{GCP_{DFP-HIA}}$, $BDC_{QN_{BFGS-HA}}$, $BDC_{QN_{DFP-HA}}$, $BDC_{GCP_{HE}}$, BDC_{GC} , $BDC_{QN_{HE}}$, $BDC_{LM_{HE}}$, $BDC_{LM_{HA}}$, BDC_{GradR} , $BDC_{App-Grad}$, $BDC_{GradAbs}$, BDC_{Alg} et $BDC_{GC_{Nion}}$ dans un contexte quasi non-bruité ($RSB = 100$ dB) pour le cas d’une matrice $\mathbf{A}$ carrée. Gauche : $l_{Conv}(\mathbf{G})$ en dB en fonction du nombre d’itérations pour un ensemble $\mathcal{M}$ de taille $N_m = 100$ matrices. Droite : $N_m = 10$ matrices.	63
2.4	Initialisation aléatoire : comparaison des algorithmes $BDC_{GCP_{BFGS-HA}}$, $BDC_{GCP_{DFP-HIA}}$, $BDC_{QN_{BFGS-HA}}$, $BDC_{QN_{DFP-HA}}$, $BDC_{GCP_{HE}}$, BDC_{GC} , $BDC_{QN_{HE}}$, $BDC_{LM_{HE}}$, $BDC_{LM_{HA}}$, BDC_{GradR} , $BDC_{App-Grad}$, $BDC_{GradAbs}$, BDC_{Alg} et $BDC_{GC_{Nion}}$ dans un contexte bruité ($RSB = 30$ dB) pour le cas d’une matrice $\mathbf{A}$ carrée. Gauche : $l_{Conv}(\mathbf{G})$ en dB en fonction du nombre d’itérations pour un ensemble $\mathcal{M}$ de taille $N_m = 100$ matrices. Droite : $N_m = 10$ matrices.	63

- 2.5 Comparaison des algorithmes $BDC_{GCP_{BFGS-HA}}$, $BDC_{GCP_{DFP-HIA}}$, $BDC_{QN_{BFGS-HA}}$, $BDC_{QN_{DFP-HA}}$, $BDC_{GCP_{HE}}$, BDC_{GC} , $BDC_{QN_{HE}}$, $BDC_{LM_{HE}}$, $BDC_{LM_{HA}}$, BDC_{GradR} , $BDC_{App-Grad}$, $BDC_{GradAbs}$ et $BDC_{GC_{Nion}}$ dans un contexte non-bruité dans le cas d'une matrice $\mathbf{A}$ rectangulaire. Gauche : $l_{Conv}(\mathbf{G})$ en dB en fonction du nombre d'itérations pour un ensemble $\mathcal{M}$ de taille $N_m = 100$ matrices. Droite : $N_m = 5$ matrices. 64
- 2.6 Comparaison des algorithmes $BDC_{GCP_{BFGS-HA}}$, $BDC_{GCP_{DFP-HIA}}$, $BDC_{QN_{BFGS-HA}}$, $BDC_{QN_{DFP-HA}}$, $BDC_{GCP_{HE}}$, BDC_{GC} , $BDC_{QN_{HE}}$, $BDC_{LM_{HE}}$, $BDC_{LM_{HA}}$, BDC_{GradR} , $BDC_{App-Grad}$, $BDC_{GradAbs}$ et $BDC_{GC_{Nion}}$ dans un contexte bruité ($RSB = 30$ dB) et pour une matrice $\mathbf{A}$ rectangulaire. Gauche : $l_{Conv}(\mathbf{G})$ en dB en fonction du nombre d'itérations pour un ensemble $\mathcal{M}$ de taille $N_m = 100$ matrices. Droite : $N_m = 5$ matrices. 64
- 2.7 Comparaison des algorithmes ($BDC_{GCP_{BFGS-HA}}$, $BDC_{GCP_{DFP-HIA}}$, $BDC_{QN_{BFGS-HA}}$, $BDC_{QN_{DFP-HA}}$, $BDC_{GCP_{HE}}$, BDC_{GC} , $BDC_{QN_{HE}}$, $BDC_{LM_{HE}}$, $BDC_{LM_{HA}}$, BDC_{GradR} , $BDC_{App-Grad}$, $BDC_{GradAbs}$, BDC_{Alg} et $BDC_{GC_{Nion}}$). À gauche : $RSB = 25$ dB. À droite : $RSB = 100$ dB. 66
- 2.8 Taille d'un bloc $r = 4$: performance en fonction du rapport signal à bruit des algorithmes $BDC_{GCP_{BFGS-HA}}$, $BDC_{GCP_{DFP-HIA}}$, $BDC_{QN_{BFGS-HA}}$, $BDC_{QN_{DFP-HA}}$, $BDC_{GCP_{HE}}$, BDC_{GC} , $BDC_{QN_{HE}}$, $BDC_{LM_{HE}}$, $BDC_{LM_{HA}}$, BDC_{GradR} , $BDC_{App-Grad}$, $BDC_{GradAbs}$, BDC_{Alg} et $BDC_{GC_{Nion}}$. Gauche : $N_m = 5$ matrices. Droite : $N_m = 100$ matrices. 66
- 2.9 Pour une taille de bloc fixée à $r = 2$, performances en fonction du nombre de matrices N_m pour les différents algorithmes ($BDC_{GCP_{BFGS-HA}}$, $BDC_{GCP_{DFP-HIA}}$, $BDC_{QN_{BFGS-HA}}$, $BDC_{QN_{DFP-HA}}$, $BDC_{GCP_{HE}}$, BDC_{GC} , $BDC_{QN_{HE}}$, $BDC_{LM_{HE}}$, $BDC_{LM_{HA}}$, BDC_{GradR} , $BDC_{App-Grad}$, $BDC_{GradAbs}$, BDC_{Alg} et $BDC_{GC_{Nion}}$). À gauche : $RSB = 30$ dB. À droite : $RSB = 100$ dB. 67
- 2.10 Pour une taille de bloc fixée à $r = 2$, performances en fonction du rapport signal à bruit pour les algorithmes $BDC_{GCP_{BFGS-HA}}$, $BDC_{GCP_{DFP-HIA}}$, $BDC_{QN_{BFGS-HA}}$, $BDC_{QN_{DFP-HA}}$, $BDC_{GCP_{HE}}$, BDC_{GC} , $BDC_{QN_{HE}}$, $BDC_{LM_{HE}}$, $BDC_{LM_{HA}}$, BDC_{GradR} , $BDC_{App-Grad}$, $BDC_{GradAbs}$, BDC_{Alg} et $BDC_{GC_{Nion}}$. À gauche : $N_m = 5$ matrices. À droite : $N_m = 100$ matrices. 67
- 2.11 Comparaison des algorithmes $BDC_{GCP_{BFGS-HA}}$, $BDC_{GCP_{DFP-HIA}}$, $BDC_{QN_{BFGS-HA}}$, $BDC_{QN_{DFP-HA}}$, $BDC_{GCP_{HE}}$, BDC_{GC} , $BDC_{QN_{HE}}$, $BDC_{LM_{HE}}$, $BDC_{LM_{HA}}$, BDC_{GradR} , $BDC_{App-Grad}$, $BDC_{GradAbs}$, BDC_{Alg} et $BDC_{GC_{Nion}}$ dans un contexte quasi non-bruité. Gauche : $l_{Conv}(\mathbf{G})$ en dB en fonction de la taille d'un bloc r pour un ensemble $\mathcal{M}$ de taille $N_m = 100$ matrices. Droite : $N_m = 5$ matrices. 69
- 2.12 Comparaison des algorithmes $BDC_{GCP_{BFGS-HA}}$, $BDC_{GCP_{DFP-HIA}}$, $BDC_{QN_{BFGS-HA}}$, $BDC_{QN_{DFP-HA}}$, $BDC_{GCP_{HE}}$, BDC_{GC} , $BDC_{QN_{HE}}$, $BDC_{LM_{HE}}$, $BDC_{LM_{HA}}$, BDC_{GradR} , $BDC_{App-Grad}$, $BDC_{GradAbs}$, BDC_{Alg} et $BDC_{GC_{Nion}}$ dans un contexte bruité ($RSB = 30$ dB). Gauche : $l_{Conv}(\mathbf{G})$ en dB en fonction de la taille d'un bloc r pour un ensemble $\mathcal{M}$ de taille $N_m = 100$ matrices. Droite : $N_m = 5$ matrices. 69
- 2.13 Diagonalisation conjointe ($r = 1$) : performances en fonction du nombre de matrices N_m pour les algorithmes $DC_{GCP_{BFGS-HA}}$, $DC_{GCP_{DFP-HIA}}$, $DC_{QN_{BFGS-HA}}$, $DC_{QN_{DFP-HA}}$, $DC_{GCP_{HE}}$, DC_{GC} , $DC_{QN_{HE}}$, $DC_{LM_{HE}}$, $DC_{LM_{HA}}$, DC_{GradR} , $DC_{App-Grad}$, $DC_{GradAbs}$, DC_{Alg} et $DC_{GC_{Nion}}$. À gauche : $RSB = 30$ dB. À droite : $RSB = 100$ dB. 70
- 2.14 Diagonalisation conjointe ($r = 1$) : performances en fonction du rapport signal à bruit pour les algorithmes $DC_{GCP_{BFGS-HA}}$, $DC_{GCP_{DFP-HIA}}$, $DC_{QN_{BFGS-HA}}$, $DC_{QN_{DFP-HA}}$, $DC_{GCP_{HE}}$, DC_{GC} , $DC_{QN_{HE}}$, $DC_{LM_{HE}}$, $DC_{LM_{HA}}$, DC_{GradR} , $DC_{App-Grad}$, $DC_{GradAbs}$, DC_{Alg} et $DC_{GC_{Nion}}$. À gauche : $N_m = 5$ matrices. À droite : $N_m = 100$ matrices. 70

2.15	Comparaison des algorithmes ZBDC_{GC} , $\text{ZBDC}_{\text{LM}_{\text{HE}}}$, $\text{ZBDC}_{\text{GCP}_{\text{HE}}}$ dans un contexte quasi non-bruité ($\text{RSB} = 100$ dB) pour le cas d'une matrice $\mathbf{A}$ carrée. Gauche : $\text{I}_{\text{Conv}}(\mathbf{G})$ en dB en fonction du nombre d'itérations pour un ensemble $\mathcal{N}$ de taille $N_m = 100$ matrices. Droite : $\text{I}_{\text{Conv}}(\mathbf{G})$ en dB en fonction du nombre d'itérations pour un ensemble $\mathcal{N}$ de taille $N_m = 5$ matrices à droite.	72
2.16	Comparaison des algorithmes ZBDC_{GC} , $\text{ZBDC}_{\text{LM}_{\text{HE}}}$, $\text{ZBDC}_{\text{GCP}_{\text{HE}}}$ dans un contexte bruité ($\text{RSB} = 30$ dB) pour le cas d'une matrice $\mathbf{A}$ carrée. Gauche : $\text{I}_{\text{Conv}}(\mathbf{G})$ en dB en fonction du nombre d'itérations pour un ensemble $\mathcal{N}$ de taille $N_m = 100$ matrices. Droite : $\text{I}_{\text{Conv}}(\mathbf{G})$ en dB en fonction du nombre d'itérations pour un ensemble $\mathcal{N}$ de taille $N_m = 5$ matrices.	72
2.17	Initialisation aléatoire : comparaison des algorithmes ZBDC_{GC} , $\text{ZBDC}_{\text{LM}_{\text{HE}}}$, $\text{ZBDC}_{\text{GCP}_{\text{HE}}}$ dans un contexte non-bruité ($\text{RSB} = 100$ dB) pour le cas d'une matrice $\mathbf{A}$ carrée. Gauche : $\text{I}_{\text{Conv}}(\mathbf{G})$ en dB en fonction du nombre d'itérations pour un ensemble $\mathcal{N}$ de taille $N_m = 100$ matrices . Droite : $N_m = 20$ matrices.	73
2.18	Initialisation aléatoire : comparaison des algorithmes ZBDC_{GC} , $\text{ZBDC}_{\text{LM}_{\text{HE}}}$, $\text{ZBDC}_{\text{GCP}_{\text{HE}}}$ dans un contexte bruité ($\text{RSB} = 30$ dB) pour le cas d'une matrice $\mathbf{A}$ carrée. Gauche : $\text{I}_{\text{Conv}}(\mathbf{G})$ en dB en fonction du nombre d'itérations pour un ensemble $\mathcal{M}$ de taille $N_m = 100$ matrices. Droite : $N_m = 20$ matrices.	73
2.19	Comparaison des algorithmes ZBDC_{GC} , $\text{ZBDC}_{\text{LM}_{\text{HE}}}$ et $\text{ZBDC}_{\text{GCP}_{\text{HE}}}$ dans un contexte non-bruité dans le cas d'une matrice $\mathbf{A}$ rectangulaire. Gauche : $\text{I}_{\text{Conv}}(\mathbf{G})$ en dB en fonction du nombre d'itérations pour un ensemble $\mathcal{M}$ de taille $N_m = 100$ matrices. Droite : $N_m = 10$ matrices.	74
2.20	Comparaison des algorithmes ZBDC_{GC} , $\text{ZBDC}_{\text{LM}_{\text{HE}}}$ et $\text{ZBDC}_{\text{GCP}_{\text{HE}}}$ dans un contexte bruité ($\text{RSB} = 25$ dB) et pour une matrice $\mathbf{A}$ rectangulaire. Gauche : $\text{I}_{\text{Conv}}(\mathbf{G})$ en dB en fonction du nombre d'itérations pour un ensemble $\mathcal{M}$ de taille $N_m = 100$ matrices. Droite : $N_m = 10$ matrices.	74
2.21	Comparaison des algorithmes (ZBDC_{GC} , $\text{ZBDC}_{\text{LM}_{\text{HE}}}$, $\text{ZBDC}_{\text{GCP}_{\text{HE}}}$). À gauche : $\text{RSB} = 20$ dB. À droite : $\text{RSB} = 100$ dB.	76
2.22	Taille d'un bloc $r = 4$: performances en fonction du rapport signal à bruit des algorithmes ZBDC_{GC} , $\text{ZBDC}_{\text{LM}_{\text{HE}}}$, $\text{ZBDC}_{\text{GCP}_{\text{HE}}}$. Gauche : $N_m = 5$ matrices. Droite : $N_m = 100$ matrices.	76
2.23	Pour une taille de bloc fixée à $r = 2$, performances en fonction du nombre de matrices N_m pour les différents algorithmes (ZBDC_{GC} , $\text{ZBDC}_{\text{LM}_{\text{HE}}}$, $\text{ZBDC}_{\text{GCP}_{\text{HE}}}$). À gauche : $\text{RSB} = 30$ dB. À droite : $\text{RSB} = 100$ dB.	77
2.24	Pour une taille de bloc fixée à $r = 2$, performances en fonction du rapport signal à bruit pour les algorithmes ZBDC_{GC} , $\text{ZBDC}_{\text{LM}_{\text{HE}}}$, $\text{ZBDC}_{\text{GCP}_{\text{HE}}}$. À gauche : $N_m = 5$ matrices. À droite : $N_m = 100$ matrices.	77
2.25	Comparaison des algorithmes $\text{ZBDC}_{\text{GCP}_{\text{HE}}}$, ZBDC_{GC} et $\text{ZBDC}_{\text{LM}_{\text{HE}}}$ dans un contexte quasi non-bruité. Gauche : $\text{I}_{\text{Conv}}(\mathbf{G})$ en dB en fonction de la taille d'un bloc r pour un ensemble $\mathcal{N}$ de taille $N_m = 100$ matrices (resp. $N_m = 5$ matrices à droite). . .	79
2.26	Comparaison des algorithmes $\text{ZBDC}_{\text{GCP}_{\text{HE}}}$, ZBDC_{GC} et $\text{ZBDC}_{\text{LM}_{\text{HE}}}$ dans un contexte bruité ($\text{RSB} = 30$ dB). Gauche : $\text{I}_{\text{Conv}}(\mathbf{G})$ en dB en fonction de la taille d'un bloc r pour un ensemble $\mathcal{N}$ de taille $N_m = 100$ matrices (resp. $N_m = 5$ matrices à droite). . .	79
2.27	Zéro-diagonalisation conjointe ($r = 1$) : performances en fonction du nombre de matrices N_m pour les algorithmes $\text{ZDC}_{\text{GCP}_{\text{HE}}}$, ZDC_{GC} , $\text{ZDC}_{\text{LM}_{\text{HE}}}$. À gauche : $\text{RSB} = 30$ dB. À droite : $\text{RSB} = 100$ dB.	80

2.28	Zéro-diagonalisation conjointe ($r = 1$) : performances en fonction du rapport signal à bruit pour les algorithmes $ZDC_{GCP_{HE}}$, ZDC_{GC} , $ZDC_{LM_{HE}}$. À gauche : $N_m = 5$ matrices. À droite : $N_m = 100$ matrices.	80
3.1	Partie réelle des sources $S - AIS$ après codage NRZI et GMSK. Haut : source 1 avec retard de 5 ms et Doppler de 10 Hz. Bas : source 2 avec retard de 0 ms et Doppler de 0 Hz.	88
3.2	À gauche : partie réelle de la distribution de Pseudo Wigner-Ville des sources $S - AIS$. À droite : partie imaginaire de la distribution de Pseudo Wigner-Ville des sources $S - AIS$	89
3.3	Partie réelle des trois signaux supposés reçus au niveau des capteurs de l'antenne.	90
3.4	À gauche : partie réelle de la distribution de Pseudo Wigner-Ville des observations. À droite : partie imaginaire de la distribution de Pseudo Wigner-Ville des observations.	91
3.5	Points temps-fréquence sélectionnés à l'aide du détecteur C_{G_1} . À gauche : $RSB = 100$ dB. À droite : $RSB = 20$ dB.	92
3.6	I en dB en fonction du RSB . Comparaison des résultats obtenus avec les différentes méthodes proposées $DC_{LM_{HE}-TF}$, $DC_{GCP_{HE}-TF}$ et $DC_{GC}-TF$ à $DC_{GC_{NION}-TF}$	92
3.7	Partie réelle des sources (trames $S - AIS$ générées après codage NRZI et GMSK). Haut : première source avec retard de 10 ms et Doppler de 4 KHz. Bas : deuxième source avec retard de 0 ms et Doppler de 0 Hz.	96
3.8	À gauche : partie réelle de la FAMSG des sources $S - AIS$. À droite : partie imaginaire de la FAMSG des sources $S - AIS$	97
3.9	À gauche : partie réelle de la FAMSG des observations. À droite : partie imaginaire de la FAMSG des observations.	98
3.10	Points Doppler-retard sélectionnés à l'aide du détecteur C_{F-ZD} . À gauche : $RSB = 100$ dB. À droite : $RSB = 15$ dB.	99
3.11	I en dB en fonction du RSB . Comparaison des algorithmes proposés ZDC_{GC-DR} , $ZDC_{LM_{HE}-DR}$, $ZDC_{GCP_{HE}-DR}$ à $ZDC_{Chabriel-DR}$	99
4.1	Séquenceur SSI pour soustraction, démodulation et estimation successive des sources $S - AIS$	103
4.2	Comportement de l'algorithme $DMUS_{SSI}$: taux d'erreur binaire (TEB) en fonction du RSB . Trame 1 avec retard de 0 ms et Doppler de 0 kHz et trame 2 avec retard de 0.0260 ms et Doppler de 100 Hz.	104
4.3	Comportement de l'algorithme $DMUS_{SSI}$: TEB en fonction du RSB : la source 1 avec retard de 0 ms et Doppler de 0 kHz, la source 2 avec retard de 0.0260 ms et un Doppler de 1 kHz et la source 3 avec un retard de 0.0521 ms et un Doppler de 3 kHz.	104
4.4	Les trois séquences émises de la cadre de l'exemple didactique.	108
4.5	Approche $DMUC_{JMLSE}$: treillis construit de l'exemple explicatif.	108
4.6	Comportement de l'algorithme $DMUC_{JMLSE}$: EMQ en fonction du RSB . À droite : 3 sources $S - AIS$ en bande de base en haut et 10 sources $S - AIS$ en bande de base en bas. À gauche : 2 sources $S - AIS$ en bande de base en haut et 5 sources $S - AIS$ en bande de base en bas.	109
4.7	Comportement de l'algorithme $DMUC_{JMLSE}$: EQM en fonction du RSB . À droite : 4 sources $S - AIS$. À gauche : 2 sources $S - AIS$	110

Liste des tableaux

1.1	Quelques RTFQ de la classe Cohen et noyaux correspondants.	20
1.2	Intervalles de temps occupés par les diverses informations du navire.	25
1.3	Intervalles de temps occupés pour l'émission des informations dynamiques (rapports) AIS de Classe A en fonction du type du navire.	25
1.4	Rapport de position AIS pour les navires de classe A (messages 1, 2 et 3).	26
2.1	Résumé des algorithmes de BDC_{NU}	50
2.2	Complexité totale pour l'ensemble des algorithmes de BDC_{NU} étudiés.	52
2.3	Résumé des algorithmes de $ZBDC_{NU}$ proposés.	58
2.4	Complexité totale des algorithmes de $ZBDC_{NU}$ proposés.	59
2.5	Cas de la zéro-diagonalisation conjointe ($r = 1$) : comparaison des performances des algorithmes itératifs $ZDC_{GCP_{HE}}$, ZDC_{GC} , $ZDC_{LM_{HE}}$ à celles d'une méthode directe $ZDC_{Chabriel}$ pour un $RSB = 20$ dB et deux tailles d'ensembles $N_m = 5$ et 100.	78
2.6	Zéro-diagonalisation conjointe ($r = 1$) : comparaison des performances des algorithmes itératifs $ZDC_{GCP_{HE}}$, ZDC_{GC} , $ZDC_{LM_{HE}}$ à celles d'une méthode directe $ZDC_{Chabriel}$ pour un $RSB = 100$ dB et deux tailles d'ensembles $N_m = 5$ et 100 . . .	81
4.1	Estimation des paramètres d'entrée : effet Doppler Δf_j , retard τ_j et amplitude A_j	102
4.2	Métriques de branches calculées dans l'exemple didactique utilisant l'approche $DMUC_{JMLSE}$	108

Liste des acronymes

ACI	Analyse en Composantes Indépendantes
FAMSG	Fonction d'Ambiguïté Moyenne Spatiale Généralisée
AMUSE	Algorithm for Multiple Unknown Signals Extraction
AIS	Automatic Identification System
BDC	Bloc Diagonalisation Conjointe
(B)DC _{GC}	(Bloc-) Diagonalisation Conjointe/Gradient Conjugué
(B)DC _{LM_{HE}}	(Bloc-) Diagonalisation Conjointe/Levenberg Marquardt
(B)DC _{GCP_{HE}}	(Bloc-) Diagonalisation Conjointe/Gradient Conjugué Préconditionné
(B)DC _{QN_{HE}}	(Bloc-) Diagonalisation Conjointe/Quasi-Newton utilisant les matrices Hessiennes Exactes
BDC _{NU}	Bloc Diagonalisation Conjointe Non Unitaire
BFGS	Broyden-Fletcher-Goldfarb-Shanno
BGML	Block-Gaussian Maximum Likelihood
BIRTH	Blind Identification of mixtures of sources using Redundancies in the daTa Hexacovariance matrix
BJ	Born-Jordan
<i>c.à.d.</i>	c'est à dire
COG	Course Over Ground
COM2	COntrast Maximisation 2
CRC	Cyclic Redundancy Code
CYCLOBLISS	CYCLOstationary BLInd Source Separation
CYMOD	CYclic Multiple Output Deconvolution
DC	Diagonalisation Conjointe
DFP	Davidon-Fletcher-Powell
DMUC	Détection Multi-Utilisateurs Conjointe
DMUS	Détection Multi-Utilisateurs Séquentielle
DOA	Direction-Of-Arrival
DVP	Décomposition en Valeur Propre
DVPG	Décomposition en Valeur Propre Généralisée
DVS	Décomposition en Valeur Singulière
ECG	ElectroCardioGramme
EEG	ElectroEncéphaloGramme
EF	End Flag
EMG	ElectroMyoGramme
EMQ	Erreur Moyenne Quadratique
FCS	Frame Control Sequence
FAS	Fonction d'Ambiguïté Spatiale

FFI	Forsvarets ForskningsInstitutt
FM	Frequency Modulation
FOBIUM	Fourth Order Blind Identification of Under-determined Mixtures
FSE	Facteur de Sur-Échantionnage
FSK	Frequency Shift Keying
GMSK	Gaussian Minimum Shift Keying
GSM	Global System for Mobile
HE	Hessienne Exacte
H(I)A	Hessienne (Inverse) Approximée
ICAR	Independent Component Analysis using Redundancies
IMO	International Maritime Organization
INFOMAX	INFORmation MAXimization
ITU	International Telecommunication Union
JADE	Joint Approximate Diagonalization of Eigenmatrices
JMLSE	Joint Maximum Likelihood Sequence Estimation
LEO	Low Earth Orbit
LMA	Levenberg-Marquardt
LMS	Least Mean Squares
MEMS	Multi-Entrées Multi-Sorties
MEUS	Multi-Entrées Unique-Sortie
MIMO	Multiple Input Multiple Output
MISO	Multiple Input Single Output
MMSI	Maritime Mobile Service Identity
MSE	Mean Square Error
NRZI	Non Return to Zero Inverse
NMF	Nonnegative Matrix Factorization
OF	Ordre du Filtre gaussien
PJOAD	Partial Approximate JOint Diagonalization
PW	Pseudo Wigner
PWL	Pseudo Wigner lissée
PWV	Pseudo Wigner-Ville
PWVL	Pseudo Wigner-Ville lissée
QN	Quasi-Newton
RAIM	Receiver Autonomous Integrity Monitoring
RD	Retard-Doppler
RIF	Réponse Impulsionnelle Finie
ROT	Rate of Turn
RSB	Rapport Signal à Bruit
RSOBI	Robust Second Order Blind Identification
RTF	Représentation Temps-Fréquence
RTFSQ	Représentation Temps-Fréquence Spatiale Quadratique
S – AIS	Satellite-Automatic Identification System
SAS	Séparation Aveugle de Sources
SDMA	Spatial Division Multiple Access
SEONS	Second-Order Non-stationary source Separation
SF	Start Flag
SIA	Système d'identification automatique
SIC	Successive Interference Cancellation

SOD	Statistiques d'Ordre Deux
SO – TDMA	Self Organized Time Division Multiple Access
SOG	Speed Over Ground
SOBI	Second Order Blind Identification
SOLAS	Safety Of Life At Sea
SOS	Statistiques d'Ordre Supérieur
SSI	Suppression Successive des Interférences
STOTD	Simultaneous Thir Order Tensor Diagonalization
STFB	Spectre Temps-Fréquence Bilinéaire
STFQ	Spectre Temps-Fréquence Quadratique
STFSB	Spectre Temps-Fréquence Spatial Bilinéaire
STFSQ	Spectre Temps-Fréquence Spatiale Quadratique
$t - f$	Temps-Fréquence
TDSEP	Temporal Decorrelation Source SEParation
TEB	Taux d'Erreur Binaire
TF	Transformée de Fourier
TFBSS	Time-Frequency Blind Source Separation
TFCT	Transformée de Fourier à Court Terme
TTFB	Transformée Temps-Fréquence Bilinéaire
TTFQ	Transformée Temps-Fréquence Quadratique
TTFSB	Transformée Temps-Fréquence Spatiale Bilinéaire
TTFSQ	Transformée Temps-Fréquence Spatiale Quadratique
VHF	Very High Frequency
ZBDC	Zéro Bloc-Diagonalisation Conjointe
$Z(B)DC_{GC}$	Zéro-(Bloc) Diagonalisation Conjointe/Gradient Conjugué
$Z(B)DC_{LM_{HE}}$	Zéro-(Bloc) Diagonalisation Conjointe/Levenberg Marquardt
$Z(B)DC_{GCP_{HE}}$	Zéro-(Bloc) Diagonalisation Conjointe/Gradient Conjugué Préconditionné
$ZBDC_{NU}$	Zéro Bloc-Diagonalisation Conjointe Non Unitaire
ZDC	Zéro Diagonalisation Conjointe

Glossaire

Dans tout le manuscrit les notations suivantes seront utilisées : les scalaires seront représentés par des lettres minuscules en caractères non gras, les vecteurs par des lettres minuscules en caractères gras et les matrices par des lettres majuscules en caractères gras.

$\mathbb{R}$	corps des réels
$\mathbb{C}$	corps des complexes
$s_j(t)$	j -ème source complexe pour tout j entier
$x_i(t)$	i -ème observation complexe pour tout i entier
$n_i(t)$	i -ème bruit complexe pour tout i entier
n	nombre de sources
m	nombre de capteurs
$(\cdot)^H$	opérateur de transposition conjugaison
$(\cdot)^T$	opérateur de transposition
$\mathbf{s}(t)$	vecteur des sources de dimension $n \times 1$
$\mathbf{x}(t)$	vecteur des observations de dimension $m \times 1$
$\mathbf{x}_b(t)$	vecteur des observations blanchies de dimension $m \times 1$
$\mathbf{n}(t)$	vecteur des bruits de dimension $m \times 1$
$\mathcal{A}\{\cdot\}$	application définie de $\mathbb{C}^n$ vers $\mathbb{C}^m$
$\mathbf{A}(t)$	système de mélange à MEMS RIF
L	ordre d'un filtre RIF
$*$	produit de convolution
$\odot$	produit de Hadamard
$\otimes$	produit de Kronecker
$a_{ij}(t)$	réponse impulsionnelle entre la j -ème source et le i -ème capteur
$\mathbf{A}$	matrice d'un système de mélange instantané de scalaires a_{ij}
L'	nombre d'observations retardées
$\Re\{\cdot\}$	partie réelle d'un nombre complexe
$\mathbf{H}$	matrice bloc-Toeplitz de dimension $M \times N$
$\mathbf{S}_j(t)$	vecteur contenant la j -ème source et ses versions retardées de dimension $Q \times 1$ ($Q = L + L'$)
$\mathbf{X}_i(t)$	vecteur contenant la i -ème observation et ses versions retardées de dimension $L' \times 1$
$\mathbf{N}_i(t)$	vecteur contenant le i -ème bruit et ses versions retardées de dimension $L' \times 1$
$\mathbf{H}_{ij}$	matrice de structure Toeplitz de dimension $L' \times (L + L')$ contenant les coefficients $a_{ij}(l)$, pour tout $l = 1, \dots, L$
$\mathbf{S}(t)$	vecteur contenant les sources et ses versions retardées de dimension $N \times 1$ ($N = nQ$)

$\mathbf{X}_b(t)$	vecteur contenant les observations blanchies et ses versions retardées de dimension $M \times 1$
$\mathbf{W}$	matrice de blanchiment de dimension $N \times M$
$\mathbf{N}(t)$	vecteur contenant les bruits et ses versions retardées de dimension $M \times 1$
$\mathbf{X}(t)$	vecteur contenant les observations et ses versions retardées de dimension $M \times 1$ ($M = mL'$)
$\mathbf{U}, \mathbf{V}$	matrice unitaire de dimension $N \times N$, matrice unitaire de dimension $M \times M$
$\Delta^{\frac{1}{2}}$	matrice diagonale de taille $M \times N$
Δ'	matrice diagonale carrée de dimension $M \times M$
$\hat{\mathbf{U}}$	matrice estimée de la matrice unitaire $\mathbf{U}$ de dimension $N \times N$
$\text{rg}\{\mathbf{H}\}$	rang de la matrice $\mathbf{H}$
t, ν, τ	temps, fréquence et retard
$\mathbf{E}\{\cdot\}$	opérateur d'espérance mathématique
$(\cdot)^{-1}$	opérateur d'inversion
$\mathbf{R}_{\mathbf{X}}(t, \tau)$	matrice de corrélation d'un vecteur $\mathbf{X}(t)$ de signaux aléatoires
$\mathbf{R}_{\mathbf{X}}(\tau)$	matrice de corrélation d'un vecteur $\mathbf{X}(t)$ de signaux aléatoires supposé stationnaire
$C_R\{\cdot\}$	cumulant d'ordre R
$\text{cum}\{\cdot\}$	l'opérateur cumulant
$\hat{\mathbf{S}}(t)$	vecteur contenant les sources estimées et ses versions retardées de dimension $N \times 1$
$\mathbf{P}$	matrice de permutation
$\mathbf{I}_m$	matrice identité de dimension $m \times m$
$\mathbf{D}$	matrice bloc-diagonale ou diagonale
$\mathbf{B}$	matrice de séparation de dimension $N \times M$
$\mathbf{K}^\#$	pseudo-inverse de la matrice $\mathbf{K}$ sachant que $\mathbf{K}^\# = (\mathbf{K}^T \mathbf{K})^{-1} \mathbf{K}^T$
$\ \mathbf{K}\ _F$	norme de Frobenius de la matrice $\mathbf{K}$
$\hat{\mathbf{B}}$	matrice estimée de la matrice de séparation $\mathbf{B}$
$\mathbf{G} = \hat{\mathbf{B}}\mathbf{H}$	matrice globale de dimension $N \times N$
$\delta(\cdot)$	distribution de Dirac
$\ \mathbf{M}\ _F$	norme de Frobenius de la matrice $\mathbf{M}$
$\text{Diag}\{\cdot\}$	opérateur permettant d'extraire la matrice diagonale de la matrice passée en argument
$\text{OffDiag}\{\cdot\}$	opérateur permettant d'extraire la matrice zéro diagonale de la matrice passée en argument
$\text{Bdiag}_{(n)}\{\cdot\}$	opérateur permettant d'extraire une matrice bloc-diagonale de la matrice passée en argument. Les r matrices bloc sont de dimension $n_j \times n_j$
$\text{OffBdiag}_{(n)}\{\cdot\}$	opérateur permettant d'extraire une matrice zéro bloc-diagonale de la matrice passée en argument. Les r matrices bloc sont de dimension $n_j \times n_j$
$\text{vec}(\cdot)$	opérateur de vectorisation : permet de concaténer les colonnes d'une matrice en un vecteur colonne
$\text{unvec}(\cdot)$	opérateur réalisant l'opération inverse de $\text{vec}(\cdot)$: remet un vecteur colonne sous forme matricielle
$\mathbf{T}_{\text{Diag}}$	matrice de transformation définie comme $\mathbf{T}_{\text{Diag}} = \text{diag}\{\text{vec}(\text{Bdiag}_{(n_j)}\{\mathbf{1}_N\})\}$
$\mathbf{T}_{\text{Boff}}$	matrice de transformation définie comme $\mathbf{T}_{\text{Boff}} = \mathbf{I}_{N^2} - \mathbf{T}_{\text{Diag}}$
$\mathbf{K}_{N,M}$	matrice carrée de commutation de taille $NM \times NM$

$\max\{.\}$	L'opérateur du maximum
-------------	------------------------

Introduction générale

Cadre de la thèse et ses objectifs

Le cadre applicatif considéré dans cette thèse est celui de la surveillance maritime par satellite. Nous nous intéressons plus spécifiquement au système d'identification automatique spatial (SIA – S¹). Le système SIA est un système d'échange automatique de données fonctionnant par radio VHF. Il est standard, en ce sens qu'il est utilisé par tous les navires afin de communiquer aux autres navires (situés à proximité) ou aux stations de contrôle côtières (les bases à terre) un certain nombre d'informations les concernant notamment leur identité, leur statut, leur position et leur route. Toutefois, au cours des dernières années, le trafic maritime a connu une évolution très importante à l'échelle internationale, si bien que le système SIA classique a commencé à atteindre ses limites. Le besoin s'est alors fait ressentir de pouvoir disposer de systèmes de surveillance à couverture beaucoup plus étendue tels qu'une veille Radar. C'est la raison pour laquelle le ministère de la défense norvégien (FFI²) a lancé une vaste étude de faisabilité quant à la détection des messages SIA par des récepteurs embarqués dans des satellites. De cette étude, il est ressorti qu'il était possible de contrôler un très grand volume de trafic maritime le long des côtes en utilisant un satellite de type Low Earth Orbit (LEO situé à une orbite ≤ 1000 km) équipé d'un récepteur AIS (ou SIA). Le système résultant fut alors baptisé Satellite-AIS (S – AIS). Il requiert toutefois des aménagements/modifications par rapport au système AIS traditionnel dont il s'inspire fortement. Dans le cas de l'AIS standard, les données navire et les données terre sont numérisées et transmises par paquet, à raison d'un découpage de chaque minute en 2 250 créneaux sur les canaux VHF 87B et 88B. Le protocole d'accès multiple à répartition auto-organisé dans le temps SO – TDMA³ est alors utilisé : il permet d'affecter des créneaux (ou slots temporels) à l'ensemble des navires et des stations à terre d'une zone (cellule AIS) de sorte qu'à un instant donné, jamais plus d'un seul émetteur ne soit en fonction. Toutefois, dans le contexte d'une surveillance par satellite, une zone beaucoup plus large est couverte, correspondant à plusieurs cellules AIS traditionnelles ce qui peut entraîner des risques de collision des données envoyées par des navires localisés dans des cellules AIS différentes, mais reçues au niveau de l'antenne d'un même satellite. A cela s'ajoute également le fait que les temps de propagation des signaux transmis des navires au satellite peuvent fluctuer en fonction de la position des navires et de la zone de couverture maximale de l'antenne du satellite. Ce problème affecte lui aussi le mécanisme d'organisation des signaux S – AIS.

Cette thèse est donc consacrée au problème de la collision des signaux S – AIS. Notre objectif est de proposer des méthodes permettant la séparation (dé-collision) de ces signaux. Le problème traité pourrait s'écrire comme un problème de traitement d'antenne classique voire de traitement d'antenne paramétrique en tenant compte i) de l'effet Doppler engendré par la vitesse de dépla-

1. En anglais S – AIS pour Satellite Automatic Identification System.

2. Forsvarets ForskningsInstitutt in Norwegian.

3. En anglais Self Organized Time Division Multiple Access.

cement du satellite (lequel induit des décalages en fréquence au niveau des signaux $S - AIS$) de même que ii) du délai ou retard lié à la propagation des signaux depuis les navires jusqu'au satellite. Toutefois, en raison de leur plus grande robustesse par rapport à d'éventuels problèmes de calibration d'antenne et pour leur indépendance vis-à-vis de la configuration de l'antenne utilisée, nous choisissons de nous focaliser plutôt sur les méthodes de séparation aveugle de sources (SAS) en reformulant le problème de la collision des signaux $S - AIS$ comme un problème de SAS . Les signaux sources considérés dans le contexte de surveillance maritime qui est le nôtre possèdent des propriétés bien particulières : ils sont complexes, cyclo-stationnaires avec des fréquences cycliques proches. C'est la raison pour laquelle, nous montrerons qu'une approche possible consiste à utiliser des méthodes fondées sur l'utilisation de distributions spatiales quadratiques et plus précisément celles du domaine retard en temps - fréquence Doppler (fonction d'ambiguïté spatiale moyenne), l'utilisation d'une diversité supplémentaire ayant pour but de rendre possible la séparation des signaux reçus au niveau du récepteur multi-capteurs. Les approches de ce type reposent toutes sur plusieurs étapes de traitement : i) construction de la distribution spatiale quadratique (dans notre cas la fonction d'ambiguïté moyenne des observations le long de l'antenne), ii) dans le plan utilisé, sélection de points particuliers qui seront utilisés afin de construire des ensembles de matrices aux propriétés particulières iii) estimation de la matrice de séparation (ou de mélange) afin de pouvoir restituer les messages $S-AIS$ transmis. Là encore, une manière classique de procéder est de recourir à des algorithmes de décompositions matricielles conjointes. Une partie de ce manuscrit sera donc consacrée à l'élaboration de nouveaux algorithmes de décompositions matricielles conjointes non unitaires dans le but de les appliquer à notre problème de dé-collision des messages AIS (bloc-diagonalisation conjointe qui généralise la diagonalisation conjointe et zéro bloc-diagonalisation conjointe qui généralise la zéro-diagonalisation conjointe). Dans un dernier temps, nous nous intéresserons également au cas sous-déterminé : le nombre de signaux AIS reçus sera alors supposé être strictement supérieur au nombre de capteurs utilisés.

Organisation du document

Ce manuscrit est organisé de la manière suivante. Après cette introduction très générale sur les objectifs de la thèse et son contexte applicatif, le Chapitre 1 est consacré à des rappels : tout d'abord, sur le problème de la séparation aveugle de sources pour des mélanges linéaires de sources (modélisation et indéterminations qui en découlent, hypothèses classiquement utilisées, principe de blanchiment). Nous effectuons ensuite un rapide état de l'art sur les principales méthodes de SAS utilisant des algorithmes de décompositions matricielles conjointes. Nous passons ensuite aux représentations temps-fréquence spatiales quadratiques ($RTFSQ$) et à leurs principales propriétés mathématiques puis nous donnons quelques exemples concrets d'applications de la SAS . Enfin, nous détaillons le principe de l'application en surveillance maritime que nous traitons dans ce manuscrit. Cela nous amène à décrire le fonctionnement du système $S - AIS$ et plus précisément les phases de transmission et de réception des messages de même que les problèmes spécifiques posés par une réception depuis un satellite.

Le Chapitre 2 est entièrement consacré aux problèmes de décompositions matricielles conjointes. Nous commençons par y rappeler le principe général de la bloc-diagonalisation conjointe, de ses hypothèses et de son lien avec la diagonalisation conjointe. Après un rapide état de l'art sur les méthodes d'optimisation itératives (du premier ordre (descente, gradient, gradient conjugué) et du second ordre (préconditionnement)) et l'introduction de la fonction de coût sur laquelle nous travaillons, nous passons au calcul exact des matrices Hessiennes dans le cas complexe (matrices des

dérivées secondes de la fonction de coût utilisée). Cela nous permet alors de proposer et d'étudier quatre nouveaux algorithmes de bloc diagonalisation conjointe. Ils sont fondés respectivement sur des algorithmes d'optimisation de type gradient conjugué, gradient conjugué préconditionné, Levenberg-Marquardt et Quasi-Newton. Nous en proposons des variantes à pas optimal. Les coûts algorithmiques sont également évalués. Nous introduisons ensuite un problème nouveau que nous choisissons d'appeler "problème de zéro-bloc diagonalisation conjointe non-unitaire" dans la mesure où il généralise le problème désormais classique de la zéro-diagonalisation conjointe non-unitaire. Nous effectuons le calcul exact du gradient matriciel complexe puis des matrices Hessiennes complexes. Cela nous permet de proposer de nouveaux algorithmes à base respectivement de gradient conjugué, gradient conjugué préconditionné et Levenberg-Marquardt. Les variantes à pas optimal sont également étudiées et la complexité algorithmique des différents algorithmes est calculée. Nous concluons ce chapitre sur des simulations numériques sur des données synthétiques afin d'illustrer le bon comportement de l'ensemble des algorithmes proposés. Nous les testons également en mode "dégradé" c'est-à-dire dans le cas particulier de la diagonalisation conjointe et de la zéro diagonalisation conjointe et les comparons à d'autres algorithmes de la littérature.

Le Chapitre 3 est consacré au développement de nouvelles méthodes de SAS exploitant les nouveaux algorithmes de décompositions matricielles conjointes développés au chapitre précédent. Les méthodes proposées exploitent le fait que les sources considérées sont des signaux stochastiques complexes et cyclo-stationnaires. C'est pourquoi nous suggérons d'utiliser des approches à base de distributions spatiales quadratiques et plus précisément celles du domaine retard en temps - fréquence Doppler (c'est-à-dire utilisant la fonction d'ambiguïté spatiale quadratique moyenne). L'utilisation d'une diversité supplémentaire a pour but de rendre possible la séparation des signaux reçus au niveau du récepteur multi-capteurs. Nous passons ensuite à l'application en surveillance maritime par satellite et au problème de la dé-collision des messages S – AIS. Nous terminons ce chapitre par des simulations informatiques illustrant le comportement des approches proposées puis les comparons à d'autres approches de la littérature.

Dans le dernier chapitre, nous traitons le problème de collision des trames S – AIS de manière analytique dans le cas où nous ne disposons plus que d'un seul capteur recevant simultanément plusieurs signaux ou trames. Nous commençons par développer une première approche séquentielle qui consiste à supprimer successivement les interférences par déflation. Puis nous proposons également un deuxième mode opératoire fondé quant à lui sur l'estimateur du maximum de vraisemblance conjoint. Nous terminons ce chapitre sur des simulations numériques afin d'illustrer le comportement de ces approches dans le contexte de la surveillance maritime par satellite.

Enfin, nous dressons une conclusion sur toute cette étude et ouvrons des perspectives à ces travaux.

Chapitre 1

Séparation aveugle de sources et surveillance maritime par satellite

Avant de présenter les trois chapitres suivants nos contributions méthodologiques, nous consacrons ce chapitre à des rappels sur le problème de la Séparation Aveugle de Sources (**SAS**). Nous commençons par la manière dont ce problème se modélise mathématiquement que l'on considère le cas de mélanges instantanés ou le cas de mélanges convolutifs de sources (le mélange instantané n'étant qu'un cas particulier du mélange convolutif). Nous remémorons ensuite les hypothèses qui sont classiquement considérées sur les sources, le système de mélange étudié (instantané ou convolutif) de même que sur les bruits. Nous évoquons les indéterminations inhérentes au problème de la séparation de sources puis introduisons les indices d'erreur utilisés afin d'évaluer les performances des méthodes présentées et de disposer d'un outil de comparaison des méthodes entre elles. Nous passons ensuite à un panorama non exhaustif des méthodes de **SAS** fondées sur l'utilisation de décompositions matricielles conjointes (le Chapitre 2 étant consacré au développement de nouveaux algorithmes de décompositions matricielles conjointes non unitaires), et ayant, des liens avec les approches que nous développerons aux Chapitres 3 et 4, dans l'optique de les appliquer à un nouveau domaine *i.e.* celui de la surveillance maritime par satellite. Puis, nous évoquons quelques applications de la séparation aveugle de sources, en l'occurrence, l'audio, les radio-communications, le biomédical et la prospection sismique.

Enfin, nous abordons le contexte applicatif de la surveillance maritime en détaillant le fonctionnement du système **AIS**. Nous décrivons ensuite le principe d'une surveillance par satellite et les "nouveaux" problèmes auxquels il faut alors faire face. Une première partie est consacrée à la description de la phase de transmission (structure des trames, principe du codage et de la modulation utilisés par ce système). La seconde partie porte quant à elle sur les différentes perturbations induites par une réception au niveau d'un satellite (et non plus à terre ou en mer) dont la possible entrée en collision des trames émises à partir de cellules **AIS** distinctes. C'est à ce problème de la séparation des messages **AIS** qu'est consacré le reste du manuscrit.

1.1 Éléments classiques de la séparation aveugle de sources

Le problème de la **SAS** consiste à restituer n signaux sources à partir de la seule connaissance de m signaux d'observation se référant à des mélanges de sources reçues au niveau d'une antenne de capteurs. Le mélange des sources (qui sont inobservables) se produit au cours de leur propagation jusqu'aux capteurs est inconnu d'où la terminologie "aveugle".

Le problème de la SAS n'est soluble qu'à l'aide d'hypothèses portant i) sur les sources ou via l'utilisation de modalités/diversités supplémentaires (domaine temps-fréquence-capteur par exemple au lieu du domaine temps-capteur ou fréquence-capteur) qui permettent alors la levée de l'hypothèse d'indépendance et ii) sur le modèle de mélange. Dans une approche statisticienne de ce problème, l'hypothèse la plus communément utilisée est celle de l'indépendance mutuelle des sources. La formalisation mathématique dans le cas le plus simple (mélange linéaire instantané) a été effectuée par P. Comon [68] ce qui a donné lieu au concept d'Analyse en Composantes Indépendantes (ACI).

En ce qui concerne le modèle de mélange, plusieurs types de mélange ont été considérés :

- le mélange peut être supposé linéaire statique sans mémoire on parle alors du modèle de mélange instantané, *c.à.d.* que les signaux reçus au niveau des capteurs ne dépendent que des combinaisons linéaires des sources aux mêmes instants,
- le mélange peut être linéaire dynamique avec mémoire plus connu sous la terminologie de mélange convolutif, *c.à.d.* les signaux reçus au niveau des capteurs dépendent à la fois des sources et de leurs versions retardées.
- enfin le mélange peut être non-linéaire, ce qui est plus général mais aussi plus complexe à traiter. La difficulté inhérente au traitement de ce type de mélanges en limite l'étude à des configurations bien particulières comme par exemple les mélanges post non-linéaires [214][7][8], *c.à.d.* le mélange est d'abord linéaire avant d'arriver sur les capteurs qui eux vont introduire une non-linéarité. Ce problème de séparation de mélanges non-linéaires reste encore relativement peu étudié, en faisant un axe relativement ouvert de recherche.

Tout au long de ce manuscrit nous ne nous intéresserons qu'aux modèles de mélange linéaire instantané et linéaire convolutif, le cas des mélanges non linéaires est en dehors du champ d'étude de ce manuscrit.

1.1.1 Modélisation mathématique du problème

Dans un contexte général, le problème de la SAS se formule de la manière simple suivante :

$$\mathbf{x}(t) = \mathcal{A}\{\mathbf{s}(t)\} + \mathbf{n}(t), \quad (1.1)$$

avec :

- $\mathbf{s}(t) = [s_1(t), \dots, s_n(t)]^T \in \mathbb{C}^n$ représente le vecteur colonne des sources émises.
- $\mathbf{x}(t) = [x_1(t), \dots, x_m(t)]^T \in \mathbb{C}^m$ est le vecteur colonne des observations reçues au niveau des capteurs.
- $\mathcal{A}\{\cdot\}$ est le système liant les observations aux sources (de $\mathbb{C}^n$ vers $\mathbb{C}^m$).
- $\mathbf{n}(t) = [n_1(t), \dots, n_m(t)]^T \in \mathbb{C}^m$ est le vecteur colonne des bruits perturbateurs additifs.
- n (resp. m) désigne le nombre de sources émises (resp. le nombre de capteurs).
- $(\cdot)^T$ désigne l'opérateur de transposition.

1.1.2 Modélisation du système de mélange linéaire

Après avoir modélisé le problème de la SAS, il reste à modéliser le système de mélange linéaire, mais auparavant, notons que le système de mélange linéaire instantané peut être considéré comme un cas particulier du système de mélange convolutif, c'est la raison pour laquelle nous commençons par la modélisation du modèle convolutif, puis passons du modèle convolutif au modèle instantané. Il faut également noter que lorsque le modèle de mélange considère plus de sources que de capteurs ($n > m$), on dit que le système de mélange est sous-déterminé sinon il est

considéré comme (sur)-déterminé ($m \geq n$). Dans ce manuscrit, le but poursuivi est principalement le développement d'approches pour la résolution du problème de la séparation de mélanges linéaires convolutifs, réels ou complexes, sur-déterminés de sources à valeurs réelles ou complexes. Ces sources pourront être aléatoires stationnaires, cyclo-stationnaires voire non-stationnaires. Le cas de mélanges sous-déterminés sera considéré dans le dernier chapitre et l'étude analytique nous permettra de proposer deux premières classes de solutions.

1.1.2.1 Mélange convolutif

En ce qui concerne l'application $\mathcal{A}\{.\}$, le signal reçu se modélise pour un mélange linéaire convolutif sous la forme matricielle comme suit :

$$\mathbf{x}(t) = \sum_{\ell=0}^L \mathbf{A}(\ell) \mathbf{s}(t - \ell) + \mathbf{n}(t) = \{\mathbf{A} * \mathbf{s}\}(t) + \mathbf{n}(t), \quad (1.2)$$

- $\mathbf{s}(t)$ est un vecteur de dimension $n \times 1$ contenant les sources $s_j(t)$, avec $j = 1, \dots, n$.
- $\mathbf{x}(t)$ est un vecteur de dimension $m \times 1$ contenant les m signaux observés $x_i(t)$, avec $i = 1, \dots, m$.
- $\mathbf{n}(t)$ est un vecteur de dimension $m \times 1$ contenant les m signaux bruits $n_i(t)$, $i = 1$, avec $\dots, m$.
- $\mathbf{A}(\ell)$, $\ell = 1, \dots, L$, sont des matrices de dimension $m \times n$ contenant les coefficients du filtre $a_{ij}(t)$ (représente la réponse impulsionnelle entre la j -ème source et le i -ème capteur) qui caractérisent le système de mélange puisque ce dernier est considéré comme un système multi-entrées multi-sorties (MEMS⁴) à réponse impulsionnelle finie (RIF) à temps discret, que l'on note $\mathbf{A}(t)$.
- L désigne l'ordre du filtre RIF.
- $*$ désigne l'opérateur du produit de convolution.

Un schéma général de la séparation de mélanges linéaires de sources est donné au niveau de la Fig.1.1 :

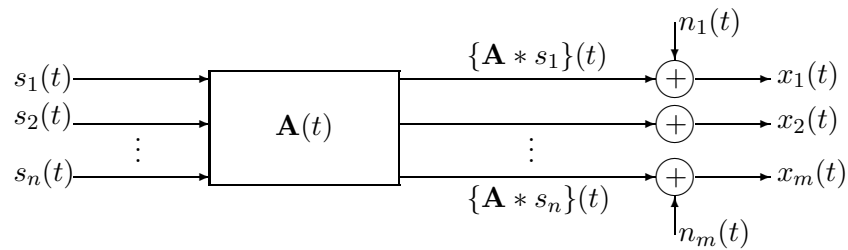


FIGURE 1.1 – Schéma général du modèle de mélange convolutif de sources.

1.1.2.2 Mélange instantané

Comme nous l'avons mentionné auparavant, le système de mélange instantané peut être considéré comme un cas particulier du système de mélange convolutif. En effet, quand les réponses impulsionnelles du système $\mathbf{A}(t)$ liant les observations aux sources sont nulles pour tous les

4. En anglais MIMO pour Multiple Input Multiple Output.

instants $t \neq 0$, le système $\mathbf{A}(t)$ équivaut alors à une matrice de scalaires a_{ij} ne dépendant plus du temps t , *c.à.d.* que :

$$\begin{cases} \mathbf{A}(0) = \mathbf{A} \text{ si } t \doteq 0, \\ \mathbf{A}(t) = 0 \text{ sinon.} \end{cases} \quad (1.3)$$

Le signal reçu se modélise alors pour un mélange linéaire instantané sous la forme matricielle suivante :

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \quad (1.4)$$

avec

- $\mathbf{s}(t)$ est un vecteur de dimension $n \times 1$ contenant les sources $s_j(t)$, avec $j = 1, \dots, n$.
- $\mathbf{x}(t)$ est un vecteur de dimension $m \times 1$ contenant les m signaux d'observation $x_i(t)$, avec $i = 1, \dots, m$.
- $\mathbf{n}(t)$ est un vecteur de dimension $m \times 1$ contenant les m signaux bruits $n_i(t)$, $i = 1$, avec $\dots, m$.
- $\mathbf{A}$ est une matrice de dimension $m \times n$ contenant les coefficients a_{ij} telle que,

$$\mathbf{A} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}. \quad (1.5)$$

1.1.2.3 Passage du modèle convolutif au modèle instantané

Dans ce paragraphe, nous rappelons de la même manière que dans [164][33][121][34][2][97][96][114][110] comment le mélange convolutif de l'Eq. (1.2) peut être reformulé au niveau du domaine temporel en un mélange linéaire instantané. Notons que pour $i = 1, \dots, m$ et $j = 1, \dots, n$, nous avons :

$$\begin{aligned} \mathbf{S}_j(t) &= [s_j(t), s_j(t-1), \dots, s_j(t-(L+L')+1)]^T, \quad \forall j = 1, \dots, n \\ \mathbf{X}_i(t) &= [x_i(t), x_i(t-1), \dots, x_i(t-L'+1)]^T, \quad \forall i = 1, \dots, m \\ \mathbf{N}_i(t) &= [n_i(t), n_i(t-1), \dots, n_i(t-L'+1)]^T, \quad \forall i = 1, \dots, m, \end{aligned} \quad (1.6)$$

avec L' est un entier représentant le nombre de retards considéré en posant $M = mL'$ et $N = n(L+L')$. Notons que le mélange est sur-déterminé si $M \geq N$, *c.à.d.* L' vérifie $mL' \geq n(L+L')$, nous introduisons donc les vecteurs suivants :

$$\begin{aligned} \mathbf{S}(t) &= [\mathbf{S}_1(t), \mathbf{S}_2(t), \dots, \mathbf{S}_n(t)]^T, \\ \mathbf{X}(t) &= [\mathbf{X}_1(t), \mathbf{X}_2(t), \dots, \mathbf{X}_m(t)]^T, \\ \mathbf{N}(t) &= [\mathbf{N}_1(t), \mathbf{N}_2(t), \dots, \mathbf{N}_m(t)]^T, \end{aligned} \quad (1.7)$$

avec $\mathbf{S}(t)$ est un vecteur de taille $(n(L+L') \times 1)$, $\mathbf{X}(t)$ et $\mathbf{N}(t)$ sont deux vecteurs de taille $(mL' \times 1)$. Pour tout $j = 1, \dots, n$ et pour tout $i = 1, \dots, m$, nous définissons la matrice $\mathbf{H}_{ij}$ dite de Toeplitz dont les coefficients sur une diagonale descendant de gauche à droite sont les mêmes :

$$\mathbf{H}_{ij} = \begin{pmatrix} a_{ij}(0) & \dots & \dots & a_{ij}(L) & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & a_{ij}(0) & \dots & \dots & a_{ij}(L) \end{pmatrix}, \quad (1.8)$$

enfin nous définissons :

$$\mathbf{X}(t) = \mathbf{H}\mathbf{S}(t) + \mathbf{N}(t), \quad (1.9)$$

avec $\mathbf{H}$ est une matrice de dimension $M \times N$ telle que :

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_{11} & \dots & \mathbf{H}_{1n} \\ \vdots & \ddots & \vdots \\ \mathbf{H}_{m1} & \dots & \mathbf{H}_{mn} \end{pmatrix}. \quad (1.10)$$

Dans le cas où $L = 0$, $L' = 1$ et $\mathbf{A} = \mathbf{H}$, nous retombons alors sur un problème de SAS en contexte linéaire instantané qui a été présenté au niveau de l'Eq. (1.4). Nous signalons que les auteurs dans [211][242][245][188][185][238] proposent de repasser au modèle instantané en travaillant au niveau du domaine fréquentiel et en exploitant la propriété de la Transformée de Fourier (TF) qui transforme un produit de convolution en un produit simple. Cependant, ce type de technique pose le problème de la reconstruction avec l'obligation de la levée des indéterminations locales propres à chaque fréquence. C'est la raison pour laquelle nous avons choisi de ne détailler et de n'utiliser que des techniques opérant au niveau du domaine temporel. Elles présentent en outre, l'avantage d'une structure algébrique bien définie : i) en convolutif : les matrices de corrélation de sources sont bloc-diagonales⁵ ii) en instantané : pour des sources mutuellement décorréliées et individuellement colorées, ces matrices de corrélation sont diagonales.

1.1.3 Hypothèses classiques de la séparation de sources

La séparation aveugle de sources comme son nom le laisse entendre, ne s'appuie sur aucune information à priori en ce qui concerne les sources ou le système de mélange. La majorité des approches développées en séparation aveugle de sources supposent toutefois que les sources sont des processus stochastiques, et qu'elles sont mutuellement indépendantes. Mathématiquement, cela veut dire que la densité de probabilité conjointe des n sources est égale au produit de leurs densités marginales, autrement dit, $\forall \tau$ et pour $i \neq j$, la fonction d'intercorrélation entre deux sources $s_i(t)$ et $s_j(t)$ est nulle ($R_{ij}(t, \tau) = \mathbb{E}\{s_i(t)s_j^*(t+\tau)\} = 0$, avec $\mathbb{E}\{\cdot\}$ l'opérateur d'espérance mathématique). Plus récemment, d'autres approches ont été développées afin de résoudre le problème de la SAS, pouvant s'appuyer sur des hypothèses autres que celle de l'indépendance des sources (les sources pourront être des signaux déterministes voire même corrélés, elles pourront être supposées non stationnaires, parcimonieuses, etc.). Par ailleurs, il arrive fréquemment que l'on s'appuie sur un certain nombre d'hypothèses supplémentaires :

Hyp₁. la fonction d'intercorrélation entre une source et sa version retardée et/ou bien entre deux versions retardées de cette source est non-nulle *c.à.d* que les composantes du vecteur des sources $\mathbf{s}(t)$ sont colorées (cette hypothèse est notamment exploitée dans certaines approches de SAS utilisant des décompositions matricielles conjointes)

Hyp₂. de nombreuses méthodes de séparation de sources reposent sur l'utilisation de cumulants, ou de covariances, etc. Des quantités qu'il est alors nécessaire d'estimer. Une hypothèse souvent implicite est celle de signaux ergodiques. Cependant en pratique, cette hypothèse n'est souvent pas valide (observation sur des durées finies relativement courtes). Le vecteur des sources $\mathbf{s}(t)$ considéré tout au long de ce manuscrit pourrait donc être stationnaire ou cyclo-stationnaire voire même non-stationnaire.

5. Une matrice bloc-diagonale est une matrice dont les blocs diagonaux sont des matrices carrées de taille quelconque et les blocs hors diagonaux sont des matrices nulles.

Hyp₃. En général, on suppose que le nombre de capteurs m est supérieur ou égal au nombre de sources n ⁶ (on parle alors d'un modèle sur-déterminé).

Hyp₄. Afin d'assurer l'existence de la matrice de séparation dans le contexte convolutif (resp. instantané), il faut garantir l'existence d'une inverse à gauche de la matrice $\mathbf{H}$ (resp. $\mathbf{A}$) *c.à.d.* qu'elle est supposée de rang plein ou encore que $\text{rg}\{\mathbf{H}\} = N$ (resp. $\text{rg}\{\mathbf{A}\} = n$) si $\text{rg}\{\cdot\}$ désigne le rang d'une matrice.

Hyp₅. Le vecteur $\mathbf{n}(t)$ contenant les bruits est stationnaire au sens strict.

Hyp₆. Les composantes de $\mathbf{s}(t)$ sont, statistiquement, indépendantes des composantes de $\mathbf{n}(t)$.

Hyp₇. Les composantes du vecteur $\mathbf{n}(t)$ sont considérées comme étant blanches gaussiennes, mutuellement indépendantes, de moyenne nulle et de variance σ_n^2 . Les cumulants d'ordre strictement supérieur à deux étant nuls pour des signaux gaussiens, l'hypothèse de gaussianité du bruit rend les approches de SAS fondées sur des cumulants d'ordre strictement supérieur à deux très robustes car insensibles aux bruits.

D'autres hypothèses peuvent être ajoutées soit sur les signaux sources (parcimonieux, de lois connues, etc.), les bruits et/ou la matrice de mélange. Notons enfin que la majeure partie des méthodes de séparation aveugle de sources présupposent que le nombre de sources est connu, sinon il faudra une étape préalable d'estimation de ce nombre. Les hypothèses présentées ci-dessus ont permis de développer des solutions au problème de la séparation de mélanges (réels ou complexes) linéaires instantanés ou convolutifs dans le cas sur-déterminé.

1.1.4 Blanchiment d'ordre deux

Dans ce paragraphe, nous rappelons le principe du blanchiment, étape importante pour un certain nombre de méthodes de SAS de la littérature que nous citons dans §1.2. Dans la recherche d'un vecteur aux composantes indépendantes, il est fréquent de commencer par une étape dite de blanchiment des observations. Cette étape consiste à décorréler les composantes du vecteur $\mathbf{X}(t)$ tout en forçant leur puissance à être unitaire. Pour ce faire, nous devons chercher la matrice $\mathbf{W}$ dite de "blanchiment" rendant le vecteur $\mathbf{X}_b(t) = \mathbf{W}\mathbf{X}(t)$ blanc, ce qui implique que la matrice de covariance de $\mathbf{X}_b(t)$ est égale à l'identité. D'après l'hypothèse **Hyp₄** (§1.1.3), la matrice $\mathbf{A}$ est régulière, elle peut donc se décomposer en valeurs singulières de la manière suivante : $\mathbf{A} = \mathbf{V}\mathbf{\Delta}^{\frac{1}{2}}\mathbf{U}$, avec $\mathbf{V}$ et $\mathbf{U}$ deux matrices unitaires de taille respectivement $M \times M$ et $N \times N$ et $\mathbf{\Delta}^{\frac{1}{2}}$ est une matrice diagonale de taille $M \times N$ $\mathbf{\Delta} = \text{diag}\{\delta_1, \dots, \delta_N\}$ avec $\Delta_{ii} = \delta_i$ pour tout $i = 1, \dots, N$ et $\delta_1 \geq \delta_2 \geq \dots \geq \delta_N \geq 0$, ainsi

$$\mathbf{A}\mathbf{A}^H = \mathbf{V}\mathbf{\Delta}'\mathbf{V}^H, \quad (1.11)$$

avec $\mathbf{\Delta}' = \mathbf{\Delta}^{\frac{1}{2}}(\mathbf{\Delta}^{\frac{1}{2}})^H = \text{diag}\{\delta_1, \dots, \delta_N, 0, \dots, 0\}$ est une matrice diagonale carrée de dimension $M \times M$ dont les N premiers termes diagonaux sont non-nuls et les $M - N$ suivants sont nuls. La matrice de corrélation des signaux observés $\overline{\mathbf{R}}_{\mathbf{X}}(\tau)$ en contexte stationnaire, est donnée en $\tau = 0$ par l'équation suivante :

$$\overline{\mathbf{R}}_{\mathbf{X}}(t, \tau) = \overline{\mathbf{R}}_{\mathbf{X}}(\tau) = \mathbb{E}\{\overline{\mathbf{X}}(t + \frac{\tau}{2})\overline{\mathbf{X}}^H(t - \frac{\tau}{2})\}, \quad (1.12)$$

$$\begin{aligned} \overline{\mathbf{R}}_{\mathbf{X}}(0) &= \mathbf{A}\overline{\mathbf{R}}_{\mathbf{S}}(0)\mathbf{A}^H + \overline{\mathbf{R}}_{\mathbf{N}}(0) \\ &= \mathbf{A}\mathbf{A}^H + \sigma_n^2\mathbf{I}_M, \end{aligned} \quad (1.13)$$

6. Cependant, nous présentons au Chapitre 4 des approches permettant de réaliser la séparation pour des mélanges contenant plus de sources que de capteurs, on parle alors de mélanges sous-déterminés.

ce qui simplifie en $\bar{\mathbf{R}}_{\mathbf{X}}(0) = \mathbf{V}\Delta'\mathbf{V}^H + \sigma_{\mathbf{n}}^2\mathbf{I}_M$ à l'aide des hypothèses sur les signaux sources (§1.1.3). Alors, en injectant l'Eq. (1.12), nous tombons sur :

$$\begin{aligned}\bar{\mathbf{R}}_{\mathbf{X}}(0) &= \mathbf{V}\Delta'\mathbf{V}^H + \sigma_{\mathbf{n}}^2\mathbf{I}_M \\ &= \mathbf{V}(\Delta' + \sigma_{\mathbf{n}}^2\mathbf{I}_M)\mathbf{V}^H,\end{aligned}\tag{1.14}$$

avec $\Delta' + \sigma_{\mathbf{n}}^2\mathbf{I}_M = \text{diag}\{\delta_1 + \sigma_{\mathbf{n}}^2, \dots, \delta_N + \sigma_{\mathbf{n}}^2, \sigma_{\mathbf{n}}^2, \dots, \sigma_{\mathbf{n}}^2\}$ une matrice diagonale et la relation donnée par l'Eq. (1.12) représente la Décomposition en Valeurs Propres (DVP) de la matrice $\bar{\mathbf{R}}_{\mathbf{X}}(0)$.

Pour l'estimation de la matrice de blanchiment, on distingue donc deux cas :

- Si $M > N$: nous pouvons estimer la variance du bruit $\sigma_{\mathbf{n}}^2$ au moyen des $M - N$ plus petites valeurs propres de $\bar{\mathbf{R}}_{\mathbf{X}}(0)$, la matrice Δ' est donc estimée par les N plus grandes valeurs propres.
- Si $M = N$, étant donné que la DVP de $\bar{\mathbf{R}}_{\mathbf{X}}(0)$ ne permet plus l'estimation de la variance du bruit $\sigma_{\mathbf{n}}^2$, nous devons la supposer connue, afin que la matrice Δ' puisse être estimée par la DVP de la matrice $\bar{\mathbf{R}}_{\mathbf{X}}(0) - \sigma_{\mathbf{n}}^2\mathbf{I}_M$.

La matrice de blanchiment spatiale $\mathbf{W}$ de taille $N \times M$ est alors définie par :

$$\mathbf{W} = (\Delta^{\frac{1}{2}})^{\#}\mathbf{V}^H,\tag{1.15}$$

avec $(\Delta^{\frac{1}{2}})^{\#}$ est la pseudo-inverse de $\Delta^{\frac{1}{2}}$ de dimension $N \times M$. Nous pouvons, maintenant, écrire le vecteur des signaux blanchis de taille $N \times 1$ sous la forme suivante :

$$\mathbf{X}_{\mathbf{b}}(t) = \mathbf{W}\mathbf{X}(t) = \mathbf{W}\mathbf{A}\mathbf{S}(t) + \mathbf{W}\mathbf{N}(t) = \mathbf{U}\mathbf{S}(t) + \mathbf{W}\mathbf{N}(t).\tag{1.16}$$

Dans le cas non-bruité, nous avons alors simplement : $\bar{\mathbf{R}}_{\mathbf{X}_{\mathbf{b}}}(0) = \mathbf{U}\bar{\mathbf{R}}_{\mathbf{S}}(0)\mathbf{U}^H = \mathbf{I}_N$. Dans ce cas là, l'étape préalable de blanchiment des observations permet de se ramener à un mélange unitaire des sources. Il ne reste alors plus qu'à estimer la matrice unitaire $\mathbf{U}$ pour parvenir à la séparation.

Dans cette thèse, notre objectif est de développer des méthodes de SAS s'affranchissant de cette étape de blanchiment des observations qui borne les performances des méthodes. C'est la raison pour laquelle au chapitre suivant, tous les algorithmes de décompositions matricielles conjointes que nous allons proposer ne reposent sur aucune contrainte d'unitarité sur la matrice recherchée ((bloc)-diagonaliseur ou zéro-(bloc) diagonaliseur conjoint).

1.1.5 Indéterminations inhérentes au problème de la séparation aveugle de sources

La séparation aveugle de sources - à moins d'hypothèses supplémentaires sur les sources (comme par exemple : la somme des puissances est égale à un, la puissance est unitaire, etc) - ne permet pas de restituer exactement les sources de départ (elles seront restituées à un facteur d'échelle près ou à un filtre près), ni leur ordre. Ces indéterminations sont inhérentes au modèle considéré et au fait que les signaux sources et la matrice de mélange sont inconnus.

Indétermination de permutation

Considérons l'Eq. (1.9), on constate que le vecteur $\mathbf{X}(t)$ peut être engendré à partir de plusieurs vecteurs $\mathbf{S}(t)$ différents. En faisant varier l'ordre des signaux sources de même que l'ordre des lignes de la matrice $\mathbf{H}$ on retrouve en effet exactement le même vecteur $\mathbf{X}(t)$. Les signaux

sources ne peuvent donc être restitués qu'à une permutation près, liée au fait que l'ordre des sources est inconnu. Cette indétermination d'ordre est modélisée au moyen d'une matrice carrée de permutation notée $\mathbf{P}$ contenant seulement un élément égal à un par ligne et par colonne et des zéros ailleurs. En effet, en utilisant la propriété de commutativité de l'opération de multiplication et celle d'addition et pour tout $i = 1, \dots, m$, nous pouvons écrire en considérant une permutation sur les 2 premières sources :

$$\begin{aligned} \mathbf{X}_i(t) - \mathbf{N}_i(t) &= \sum_{j=1}^n \mathbf{H}_{ij} \mathbf{S}_j(t) \\ &= \mathbf{H}_{i1} \mathbf{S}_1(t) + \mathbf{H}_{i2} \mathbf{S}_2(t) + \dots + \mathbf{H}_{in} \mathbf{S}_n(t) \\ &= \mathbf{H}_{i2} \mathbf{S}_2(t) + \mathbf{H}_{i1} \mathbf{S}_1(t) + \dots + \mathbf{H}_{in} \mathbf{S}_n(t), \end{aligned} \quad (1.17)$$

Ce qui conduit finalement aussi à :

$$\mathbf{X}(t) - \mathbf{N}(t) = \tilde{\mathbf{H}} \tilde{\mathbf{S}}(t), \quad (1.18)$$

avec

$$\tilde{\mathbf{S}}(t) = [\mathbf{S}_2(t), \mathbf{S}_1(t), \mathbf{S}_3(t), \dots, \mathbf{S}_n(t)]^T, \quad (1.19)$$

et

$$\tilde{\mathbf{H}} = \begin{pmatrix} \mathbf{H}_{21} & \mathbf{H}_{22} & \dots & \mathbf{H}_{2n} \\ \mathbf{H}_{11} & \mathbf{H}_{12} & \dots & \mathbf{H}_{1n} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{H}_{m1} & \mathbf{H}_{m2} & \dots & \mathbf{H}_{mn} \end{pmatrix}. \quad (1.20)$$

Ainsi,

$$\mathbf{X}(t) - \mathbf{N}(t) = \tilde{\mathbf{H}} \tilde{\mathbf{S}}(t) = \mathbf{H} \mathbf{P} \tilde{\mathbf{S}}(t) = \mathbf{H} \mathbf{P} \mathbf{P}^{-1} \mathbf{S}(t) = \mathbf{H} \mathbf{S}(t), \quad (1.21)$$

où $\mathbf{P}$ est une matrice de permutation ($\mathbf{P} \mathbf{P}^T = \mathbf{I}$ et $\mathbf{P}^{-1} = \mathbf{P}^T$).

Indétermination d'échelle

Pour tout $i = 1, \dots, m$, pour tout $\tau = 0, \dots, (L+L')+1$ et pour tout $\alpha_j^{(\tau)} \in \mathbb{C}^*$ ($j = 1, \dots, n$), si nous multiplions chaque source $\mathbf{S}_j(t)$ par un scalaire $\alpha_j^{(\tau)}$ non nul et nous divisons chaque élément correspondant de la matrice $\mathbf{H}$ par ce même scalaire, cela n'entraîne aucune modification sur le vecteur d'observation $\mathbf{X}(t)$. En effet, l'équation 1.2 peut être définie sous la forme suivante :

$$x_i(t) = \sum_{j=1}^n \sum_{\ell=0}^L \alpha_j^{(\ell)} a_{ij}(\ell) \frac{s_j(t-\ell)}{\alpha_j^{(\ell)}} + n_i(t). \quad (1.22)$$

Ainsi,

$$\mathbf{X}(t) - \mathbf{N}(t) = \overline{\mathbf{H}} \overline{\mathbf{S}}(t) = \mathbf{H} \mathbf{D} \overline{\mathbf{S}}(t) = \mathbf{H} \mathbf{D} \mathbf{D}^{-1} \mathbf{S}(t) = \mathbf{H} \mathbf{S}(t), \quad (1.23)$$

avec $\mathbf{D}$ est une matrice bloc-diagonale inversible,

$$\overline{\mathbf{S}}(t) = \left[\frac{s_1(t)}{\alpha_1^{(0)}}, \dots, \frac{s_1(t - (L+L') + 1)}{\alpha_1^{((L+L')+1)}}, \dots, \frac{s_n(t - (L+L') + 1)}{\alpha_n^{((L+L')+1)}} \right]^T, \quad (1.24)$$

et

$$\bar{\mathbf{H}} = \begin{pmatrix} \bar{\mathbf{H}}_{11} & \dots & \bar{\mathbf{H}}_{1n} \\ \vdots & \ddots & \vdots \\ \bar{\mathbf{H}}_{m1} & \dots & \bar{\mathbf{H}}_{mn} \end{pmatrix}, \quad (1.25)$$

avec

$$\bar{\mathbf{H}}_{ij} = \begin{pmatrix} \alpha_j^{(0)} a_{ij}(0) & \dots & \dots & \alpha_j^{(L)} a_{ij}(L) & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \alpha_j^{(0)} a_{ij}(0) & \dots & \dots & \alpha_j^{(L)} a_{ij}(L) \end{pmatrix}. \quad (1.26)$$

En modèle de mélange convolutif, l'indétermination d'échelle dans la restitution des sources fait qu'elles ne peuvent être retrouvées qu'à un filtre près (en modèle de mélange instantané cela se réduit à une amplitude près, dans ce cas là, la matrice $\mathbf{D}$ est alors simplement une matrice diagonale inversible).

Au final, le problème de la **SAS** consiste donc à estimer une matrice $\mathbf{B}$ dite de séparation telle que les signaux sources reconstruits $\hat{\mathbf{S}}(t)$ s'écrivent de la façon suivante :

$$\hat{\mathbf{S}}(t) = \mathbf{B}\mathbf{X}(t) = \mathbf{B}\mathbf{H}\mathbf{S}(t) = \mathbf{P}\mathbf{D}\mathbf{S}(t), \quad (1.27)$$

avec $\mathbf{P}$ matrice de permutation et $\mathbf{D}$ est une matrice bloc-diagonale inversible. Cela conduit encore à :

$$\mathbf{B}\mathbf{H} = \mathbf{P}\mathbf{D}. \quad (1.28)$$

La séparation est donc atteinte dès lors que la matrice de séparation $\mathbf{B}$ obtenue est de la forme :

$$\mathbf{B} = \mathbf{P}\mathbf{D}\mathbf{H}^\#, \quad (1.29)$$

où $(.)^\#$ désigne l'opérateur de pseudo-inverse (ou inverse de Moore-Penrose généralisée).

Ces indéterminations n'ont aucune incidence sur la propriété (d'indépendance par exemple...) sur laquelle on s'appuie pour réaliser la séparation des signaux que l'on cherche à estimer en minimisant un critère judicieusement choisi (tels que ceux que nous introduisons dans le chapitre suivant). Ajoutons malgré tout que plusieurs approches ont quand même été développées afin de lever ces indéterminations [211][180][189][147][137][168][85] [213]. Notons également que l'indétermination d'échelle dans le domaine complexe comprend l'indétermination de rotation de phase qui change d'une modulation à l'autre [136].

1.1.6 Indice d'erreur

Dans le but d'évaluer la qualité de la séparation dans nos simulations informatiques et de pouvoir comparer les algorithmes entre eux, nous utilisons l'indice d'erreur proposé en convolutif par H. Ghennioui dans [108][114] généralisant celui introduit dans [171][149] pour l'instantané :

$$l_{\text{Conv}}(\mathbf{G}) = \frac{1}{Q(Q-1)} \left[\sum_{i=1}^Q \left(\sum_{j=1}^Q \frac{\|(\mathbf{G})_{i,j}\|_F^2}{\max_{\ell} \|(\mathbf{G})_{i,\ell}\|_F^2} - 1 \right) + \sum_{j=1}^Q \left(\sum_{i=1}^Q \frac{\|(\mathbf{G})_{i,j}\|_F^2}{\max_{\ell} \|(\mathbf{G})_{\ell,j}\|_F^2} - 1 \right) \right], \quad (1.30)$$

où $Q = L + L'$, $\|\cdot\|_F$ désigne la norme de Frobenius, $(\mathbf{G})_{i,j} \forall i, j \in \{1, \dots, Q\}$ est la (i, j) -ème

matrice bloc (carrée) de $\mathbf{G} = \hat{\mathbf{B}}\mathbf{H}$ ⁷. Une séparation parfaite correspond à un indice d'erreur tendant vers $-\infty$ en échelle logarithmique (0 en échelle linéaire). Cet indice mesure la proximité de $\hat{\mathbf{B}}\mathbf{H}$ par rapport au produit $\mathbf{P}\mathbf{D}$ d'une matrice de permutation par une matrice bloc-diagonale. Une bonne séparation est obtenue lorsque la matrice $\mathbf{G}$ est égale à une matrice contenant un seul bloc non-nul par bloc-ligne et par bloc-colonne⁸ (resp. en instantané, cela signifie q'un seul terme non-nul est obtenu par ligne et par colonne). Nous rappelons que lorsque $L = 0$, $L' = 1$ et $\mathbf{A} = \mathbf{H}$, nous retrouvons alors bien l'indice d'erreur introduit dans [171][149] pour mesurer la qualité de séparation dans le cadre des systèmes de mélanges instantanés.

1.2 Méthodes de séparation de sources : un bref historique

Après avoir rappelé les points les plus importants du problème de la séparation aveugle de sources, nous effectuons, dans cette partie, un panorama (non exhaustif) d'un certain nombre de méthodes de SAS. Nous nous concentrons principalement sur les méthodes fondées sur l'utilisation de décompositions matricielles conjointes étant donné que ces dernières font l'objet du chapitre suivant (on pourra consulter [71] ou [51] pour un panorama relativement exhaustif des méthodes existantes). D'autre part, elles sont également utilisées dans la résolution du problème de SAS que nous considérons dans le contexte applicatif de la surveillance maritime. Nous avons fait le choix ici de scinder ces méthodes en deux catégories. Nous traitons d'abord des méthodes dévolues au cas des mélanges instantanés, puis nous passons ensuite au cas des mélanges convolutifs.

1.2.1 Mélange instantané

1.2.1.1 Introduction

L'analyse en composantes indépendantes (ACI), introduite en 1980, est un cas particulier de la SAS pour des mélanges instantanés sur-déterminés et des sources aléatoires. Elle a pour but la décomposition d'un signal en sous composantes additives indépendantes (dont au plus une composante peut être gaussienne). Une manière de procéder consiste à construire des fonctions de contraste utilisant les statistiques d'ordre supérieur et exploitant certaines propriétés des cumulants d'ordre supérieur. Les premières recherches ont été initiées par Jutten et al. en 1985 [15][142][143]. Ils se sont intéressés en particulier au problème de décodage du mouvement en neurobiologie [124][126][15]. Ils ont alors proposé la première approche de séparation de sources basée sur des réseaux neuromimétiques [144]. Par la suite, P. Comon a formalisé clairement le concept d'ACI en étendant la notion de contraste exploitée dans les systèmes de déconvolution aveugle de processus aléatoires scalaires [86][103] au cas des vecteurs aléatoires [68] : l'estimation de la matrice de mélange $\mathbf{A}$ (ou de sa pseudo-inverse : la matrice de séparation $\mathbf{B}$) se fait à travers l'optimisation d'un critère connu sous le nom de "fonction de contraste" ou plus simplement de "contraste".

Première définition d'une fonction de contraste proposée par P. Comon

7. $\hat{\mathbf{B}}$ représentant l'estimée de la matrice de séparation $\mathbf{B}$.

8. Par exemple : $\mathbf{G} = \begin{pmatrix} \mathbf{0} & \mathbf{G}_{11} & \cdots & \mathbf{0} \\ \mathbf{G}_{21} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{G}_{mn} \end{pmatrix}$, où $\mathbf{0}$ désigne la matrice nulle de dimension $Q \times Q$ et $\mathbf{G}_{ij}, i \in \{1, \dots, n\}, j \in \{1, \dots, n\}$ sont les matrices carrées non-nulles de dimension $Q \times Q$.

Soit un ensemble $\mathcal{Y}$ de vecteurs aléatoires, un contraste est une fonction multivariable $\mathcal{I}(\cdot)$ de $\mathcal{Y}$ dans $\mathbb{R}$, ne dépendant que de la loi de probabilité d'un vecteur aléatoire $\mathbf{y}$ de l'ensemble $\mathcal{Y}$, devant satisfaire les propriétés suivantes [68] :

P₀. $\forall \mathbf{y} \in \mathcal{Y}, \forall \mathbf{D}$, matrice diagonale inversible, $\mathcal{I}(\mathbf{D}\mathbf{y}) = \mathcal{I}(\mathbf{y})$.

P₁. $\forall \mathbf{y} \in \mathcal{Y}, \forall \mathbf{P}$, matrice de permutation inversible, $\mathcal{I}(\mathbf{P}\mathbf{y}) = \mathcal{I}(\mathbf{y})$.

P₂. $\forall \mathbf{s} \in \mathcal{Y}$ de composantes indépendantes et $\forall \mathbf{M}$, tel que $\mathbf{M}\mathbf{s} \in \mathcal{Y}, \mathcal{I}(\mathbf{M}\mathbf{s}) \leq \mathcal{I}(\mathbf{s})$.

P₃. $\mathcal{I}(\mathbf{M}\mathbf{s}) = \mathcal{I}(\mathbf{s}) \Leftrightarrow \mathbf{M} = \mathbf{D}\mathbf{P}$.

Il résulte des propriétés énoncées ci-dessus que les contrastes sont des fonctions invariantes par changement d'échelle (**P₀**) et par permutation (**P₁**) qui doivent être maximisées (propriété **P₂**) pour résoudre le problème de la séparation (**P₃**).

Exemples de contraste : l'une des premières méthodes proposée par P. Comon dans [68] nommée COM2 (CONtrast Maximisation 2) consiste à maximiser une fonction de contraste notée $\mathcal{I}_R^2(\mathbf{y})$ laquelle implique les modules au carré des cumulants d'ordre $R \geq 3$ des sources estimées $\mathbf{y}(t)$:

$$\mathcal{I}_R^2(\mathbf{y}) = \sum_{i=1}^n |\mathbf{C}_R\{y_i\}|^2, \quad R \geq 3, \quad (1.31)$$

où n est le nombre de sources et $\mathbf{C}_R\{\cdot\}$ désigne le cumulant d'ordre R d'un processus stationnaire. La maximisation de la fonction $\mathcal{I}_R^2(\mathbf{y})$ conduit à l'algorithme ACI [68], cette méthode revenant également à la DC d'un tenseur d'ordre R . Notons que ce contraste ne fait appel à aucun cumulant croisé. C'est J.-F. Cardoso et A. Souloumiac qui ont eu l'idée de proposer dans [43] l'utilisation de cumulants d'ordre quatre exploitant à la fois des auto-cumulants et des cumulants croisés. Il s'agit de la fonction de contraste utilisée dans l'algorithme JADE (Joint Approximate Diagonalization of Eigenmatrices) que l'on note $\mathcal{J}(\mathbf{y})$:

$$\mathcal{J}(\mathbf{y}) = \sum_{i_1, i_2, i_3=1}^n |\text{Cum}\{y_{i_1}, y_{i_1}, y_{i_2}, y_{i_3}\}|^2, \quad (1.32)$$

où l'opérateur $\text{Cum}\{\cdot\}$ désigne un cumulant d'ordre quatre. Le but est alors de maximiser la fonction JADE ce qui revient encore à diagonaliser conjointement un ensemble de n^2 matrices.

Un peu plus tard, L. De Lathauwer a proposé dans [80] l'algorithme STOTD (Simultaneous Third Order Tensor Diagonalization) qui suit la même logique que l'algorithme JADE, à ceci près que cet algorithme consiste à diagonaliser conjointement un ensemble de tenseurs d'ordre trois au carré et non plus un ensemble de matrices. Il repose sur la maximisation de la fonction contraste suivante notée $\mathcal{S}(\mathbf{y})$:

$$\mathcal{S}(\mathbf{y}) = \sum_{i_1, i_2}^n |\text{Cum}\{y_{i_1}, y_{i_1}, y_{i_1}, y_{i_2}\}|^2. \quad (1.33)$$

Généralisation des fonctions de contraste selon E. Moreau : la fonction de contraste suivante - généralisant les contrastes présentés précédemment - est alors proposée par E. Moreau dans [171] :

$$\mathcal{G}(\mathbf{y})_R^{R_1, R_2} = \sum_{i_1, i_2, j_1, \dots, j_{R_3}=1}^n |\mathbf{C}_R^{R_1, R_2}\{\mathbf{y}\}|^2, \quad \text{tel que} \quad \begin{cases} R \geq 3, 2 \leq R_1 \leq R, \\ 0 \leq R_2 \leq R - R_1, \\ R_3 = R - R_1 - R_2, \end{cases} \quad (1.34)$$

avec $R, R_1, R_2, R_3 \in \mathbb{N}$ et

$$\mathcal{C}_R^{R_1, R_2} \{\mathbf{y}\} = \text{Cum} \left\{ \underbrace{y_{i_1}, \dots, y_{i_1}}_{R_1 \times}, \underbrace{y_{i_2}, \dots, y_{i_2}}_{R_2 \times}, \underbrace{y_{i_3}, \dots, y_{i_{R_3+2}}}_{R_3 \text{ termes}} \right\}.$$

Pour retrouver les contrastes précédents $\mathcal{I}_R^2(\mathbf{y})$, $\mathcal{J}(\mathbf{y})$ et $\mathcal{S}(\mathbf{y})$, il suffit de considérer :

- $R_2 = R_3 = 0 \Rightarrow \mathcal{G}(\mathbf{y})_R^{R,0} = \mathcal{I}_R^2(\mathbf{y})$, suggéré par P. Comon Eq. (1.31),
- $R_2 = 0$ et $R_3 = 2 \Rightarrow \mathcal{G}(\mathbf{y})_4^{2,0} = \mathcal{J}(\mathbf{y})$, suggéré par J.F. Cardoso et A. Souloumiac (Eq. 1.32),
- $R_1 = 3 \Rightarrow \mathcal{G}(\mathbf{y})_4^{3,0} = \mathcal{S}(\mathbf{y})$, suggéré par L. Delathauwer (Eq. 1.33).

1.2.1.2 Autres méthodes exploitant la technique de diagonalisation conjointe

Méthode AMUSE : parmi les nombreuses méthodes de SAS exploitant la technique de la DC, nous trouvons la méthode AMUSE introduite par Tong et al. [224] [223] qui n'exploite que les statistiques du second ordre des sources et qui réalise la diagonalisation simultanée de deux matrices de covariance des observations, l'une associée à un retard nul et la seconde à un retard non nul (l'une de ces matrices devant nécessairement être une matrice définie positive). Cette méthode suppose que :

Hyp₀. Les sources sont mutuellement décorrélées ($\mathbb{E}\{s_i(t)s_j^*(t+\tau)\} = 0$) avec $\tau > 0$.

Hyp₁. Les statistiques des sources sont distinctes : $\mathbb{E}\{s_i(t)s_i^*(t+\tau)\} \neq \mathbb{E}\{s_j(t)s_j^*(t+\tau)\}$ pour $i \neq j$.

L'hypothèse **Hyp₀** permet aux matrices de covariance des sources $\overline{\mathbf{R}}_s(\tau)$, (où $\overline{\mathbf{R}}_s(\tau) = \langle \mathbf{R}_s(t, \tau) \rangle$) d'être forcément diagonales avec $\langle \cdot \rangle$ l'opérateur de moyennage temporel défini pour un vecteur $\mathbf{z}(t)$ comme :

$$\langle \mathbf{z}(t) \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \mathbf{z}(t) dt. \quad (1.35)$$

L'hypothèse **Hyp₁** assure le fait que ces deux matrices possèdent des valeurs propres différentes. La DC permet d'estimer la matrice unitaire $\mathbf{U}$ qui diagonalise simultanément deux matrices (en l'occurrence $\mathbf{M}_0$ qui n'est autre que la matrice de corrélation à l'origine et $\mathbf{M}_1$) :

$$\begin{aligned} \mathbf{U}\mathbf{M}_0\mathbf{U}^H &= \mathbf{I}, \\ \mathbf{U}\mathbf{M}_1\mathbf{U}^H &= \mathbf{D}_1, \end{aligned}$$

où $\mathbf{D}_1$ est une matrice diagonale. Les matrices $\mathbf{M}_0$ et $\mathbf{M}_1$ sont deux matrices symétriques sachant que $\mathbf{M}_0$ est de surcroît définie positive. La méthode AMUSE opère à partir d'observations blanches, ce qui permet de passer d'un mélange $\mathbf{A}$ quelconque à un mélange $\mathbf{U}$ unitaire. Pour ce faire, on construit : $\mathbf{x}_b(t) = \mathbf{W}\mathbf{x}(t) = \mathbf{U}\mathbf{s}(t) + \mathbf{W}\mathbf{n}(t)$, avec la matrice $\mathbf{W}$ définie précédemment au niveau de l'Eq. (1.15). Cette matrice de blanchiment est identifiée au moyen de la DVP de la matrice $\mathbf{M}_0$. La matrice $\mathbf{U}$ est ensuite obtenue sous la forme $\mathbf{U} = \mathbf{V}\mathbf{W}$ à partir de la décomposition spectrale $\mathbf{W}\mathbf{M}_1\mathbf{W}^T = \mathbf{V}^T\mathbf{D}_1\mathbf{V}$, avec $\mathbf{V}$ la matrice estimée par la DVP de la matrice $\mathbf{M}_1$ qui est la matrice de corrélation des observations blanches à un retard $\tau \neq 0$. La matrice de mélange est ensuite estimée par $\hat{\mathbf{A}} = \mathbf{W}^\# \mathbf{V}$ de même que les sources : $\hat{\mathbf{s}}(t) = \mathbf{V}^H \mathbf{x}_b(t) = \mathbf{V}^H \mathbf{W} \mathbf{x}(t)$. Toutefois, en pratique, des contraintes sont nécessaires car l'estimation des matrices de corrélation peut s'avérer difficile (elles peuvent ne plus être ni symétriques ni hermitiennes). Pour remédier

à ce problème, les matrices de corrélation peuvent être forcées à être symétriques au moyen d'une procédure du type de celle donnée ci-dessous :

$$\overline{\mathbf{R}}_{\mathbf{x}_b}(\tau) = \frac{\left(\overline{\mathbf{R}}_{\mathbf{x}_b}(\tau) + \overline{\mathbf{R}}_{\mathbf{x}_b}^H(\tau)\right)}{2}. \quad (1.36)$$

L'une d'elle peut également ne plus être définie positive. Pour résoudre ce problème, on peut améliorer le conditionnement de la matrice au moyen d'une procédure du type :

$$\overline{\mathbf{R}}_{\mathbf{x}_b}(\tau) = \overline{\mathbf{R}}_{\mathbf{x}_b}(\tau) + \delta \mathbf{I}_m, \quad (1.37)$$

où δ est une constante réelle de petite valeur et $\mathbf{I}_m$ est la matrice identité de taille $m \times m$. Notons que dans le cas où $N_m = 2$, une solution analytique a été proposée dans [190] en exploitant la diagonalisation conjointe de deux matrices définies positives.

Méthode SOBI : dans le but d'améliorer la robustesse et les performances de l'algorithme AMUSE, Belouchrani et al. ont introduit dans [22] l'algorithme SOBI (Second Order Blind Identification) qui n'est rien d'autre que l'extension de l'algorithme AMUSE : au lieu de se restreindre à la diagonalisation de deux matrices ($N_m = 2$), ils en diagonalisent simultanément tout un ensemble. Cet algorithme suit alors la même logique que l'algorithme AMUSE. L'ensemble, noté $\overline{\mathcal{M}}$, sur lequel ils travaillent est construit à partir des matrices de corrélation des observations blanchies à différents retards $i \in \{1, \dots, N_m\}$ avec N_m le nombre de matrices et donc de retards considérés. La DC des matrices de l'ensemble $\overline{\mathcal{M}}$ est ensuite réalisée au moyen d'un algorithme de type Jacobi [44]. C'est là que se situe la différence majeure entre l'algorithme AMUSE et l'algorithme SOBI : la DVP de la matrice $\mathbf{M}_1$ à un retard $\tau \neq 0$ est remplacée dans SOBI par la DC d'un ensemble de matrices à différents retards.

Méthodes dérivées de SOBI et JADE : dans l'esprit des méthodes SOBI et JADE, de nombreuses autres méthodes ont vu le jour. Ainsi, dans [93], les auteurs ont proposé une extension de la méthode JADE afin de proposer une nouvelle famille de méthodes de SAS utilisant les statistiques d'ordre supérieur et exploitant la propriété de cyclo-stationnarité des sources. Elle a été dénommée JADE_{cyclique}. Il existe une autre méthode très proche de SOBI baptisée TDSEP (Temporal Decorrelation Source SEParation) proposée par Ziehe et Müller dans [264]. A. Belouchrani [26] a également proposé d'améliorer l'algorithme SOBI et plus précisément l'étape de blanchiment ce qui a donné naissance à une méthode plus robuste en présence d'un bruit additif décorrélié dans le temps mais pouvant être corrélé spatialement d'où son nom de RSOBI. Dans la même optique, nous trouvons plusieurs méthodes comme BIRTH proposée par Albera et al. dans [12], ICAR [9][10][11] et FOBIUM proposée par Ferréol et al. dans [95] qui ne sont rien d'autre que des extensions de la méthode SOBI à l'ordre quatre et à l'ordre six. Ces méthodes s'affranchissent toutefois de l'étape préalable de blanchiment. Il faut également noter que le principal avantage des méthodes FOBIUM et BIRTH réside dans le fait qu'elles permettent également de considérer le cas de mélanges sous-déterminés. Enfin, une autre idée a également été développée : celle de modéliser des sources non-stationnaires au moyen de gaussiennes stationnaires par morceaux d'où la possibilité d'exploiter l'approche dite du Maximum de Vraisemblance. De cette idée, sont nés les algorithmes BGML (Block-Gaussian Maximum Likelihood) [192] et SEONS (Second-Order Non-stationary source Separation) [63].

1.2.1.3 Méthodes basées sur l'utilisation de transformées temps-fréquence spatiales quadratiques

Parce que les hypothèses de sources aléatoires stationnaires et indépendantes sont des hypothèses relativement fortes sur les sources qui ne sont pas toujours satisfaites par ces dernières, plusieurs travaux ont porté sur la recherche d'autres types d'approches permettant la relaxe de ces hypothèses au bénéfice d'hypothèses moins contraignantes sur les sources (pour un panorama complet, on pourra consulter [1]). En effet, il n'est pas rare dans les problèmes de séparation/démixage de signaux (ou d'images), que les signaux à séparer soient des signaux déterministes ou encore des signaux cyclo-stationnaires voire non stationnaires voire même des signaux corrélés. Il s'est avéré que les méthodes fondées sur l'utilisation de transformées temps-fréquence spatiales quadratiques permettent d'attaquer ce type de problématique grâce à l'introduction d'une modalité/diversité supplémentaire par rapport aux méthodes plus classiques (extension des méthodes temps-capteurs ou fréquence-capteurs aux méthodes temps-fréquence-capteurs).

Ces méthodes, suggérées à l'origine par [25], se décomposent généralement en trois grandes étapes de traitement :

1. construction de distributions temps-fréquence spatiales quadratiques des observations (blanches ou non) le long de l'antenne de capteurs,
2. dans le plan "temps-fréquence", sélection de points temps-fréquence particuliers et constructions des ensembles de matrices qui leurs sont associés
3. estimation de la matrice de séparation (ou de mélange) à partir de ces ensembles de matrices aux propriétés algébriques particulières en utilisant des décompositions matricielles conjointes (diagonalisation conjointe et/ou zéro-diagonalisation conjointe) et restitution des sources aux indéterminations de la SAS près.

Nous allons donc commencer par des rappels sur les spectres et les distributions temps-fréquence. Une méthode classique d'analyse d'un signal consiste à lui associer une transformée particulière de sa représentation initiale. Pour des besoins particuliers en traitement du signal, on a donc été amené à associer à un signal des représentations dépendant non plus d'une seule mais de deux variables (le temps et la fréquence) et possédant donc simultanément les deux caractères temporel et fréquentiel. Ces transformées sont qualifiées de représentations temps-fréquence (RTF). Il existe différents types de RTF : i) paramétriques ii) et non paramétriques [20][218]. L'ensemble non paramétrique peut lui même se répartir en plusieurs sous-ensembles : i) linéaires, ii) bilinéaires/quadratiques et iii) RTF d'ordre supérieur.

Dans ce manuscrit, nous nous intéressons plus particulièrement aux RTF bilinéaires/quadratiques énergétiques. Elles aussi sont regroupées en classes particulières en fonction des propriétés qu'elles possèdent (le lecteur pourra se reporter à la référence [67] pour un panorama relativement exhaustif des RTF quadratiques). Ainsi, on distingue :

- la classe de Cohen regroupant les RTF énergétiques possédant des propriétés d'invariance par translations temporelles et fréquentielles.
- la classe affine regroupant les RTF énergétiques possédant des propriétés d'invariance par translations temporelles et changements d'échelles.

Nous allons davantage détailler les RTF qui appartiennent à la classe de Cohen dans la mesure où nous serons amenés à développer des méthodes et à présenter des simulations qui les utilisent dans la suite de ce manuscrit. Notons quand même qu'au Chapitre 3, étant donné la nature des signaux de télécommunications numériques que nous chercherons à séparer, nous développerons également une méthode fondée sur l'utilisation de la fonction d'Ambiguïté moyenne et travaillerons dans un espace dual à savoir le domaine retard - fréquence Doppler.

Spectres et transformées temps-fréquence spatiaux (spatiales) bilinéaires et quadratiques

Spectre et transformée temps-fréquence spatial(e) bilinéaire

Soient $\mathbf{x}(t) = [x_1(t), \dots, x_n(t)]^T$, $\mathbf{y}(t) = [y_1(t), \dots, y_n(t)]^T$ deux signaux (éventuellement analytiques) aléatoires, réels ou complexes. Le Spectre Temps-Fréquence Spatial Bilineaire (STFSB) associé s'écrit sous la forme suivante :

$$(\mathbf{x}(t), \mathbf{y}(t)) \xleftrightarrow{\text{STFSB}} \mathbf{D}_{\mathbf{xy}}(t, \nu; \mathbf{K}) = \int_{\mathbb{R}^2} \mathbf{R}_{\mathbf{xy}}(\theta, \theta') \odot \mathbf{K}(\theta, \theta'; t, \nu) d\theta d\theta', \quad (1.38)$$

où la fonction $\mathbf{K}(\theta, \theta'; t, \nu)$ désigne le noyau de la transformation, $\odot$ désigne le produit de Hadamard⁹ et $\mathbf{R}_{\mathbf{xy}}(\theta, \theta')$ représente la matrice de corrélation définie de la façon suivante : $\mathbf{R}_{\mathbf{xy}}(\theta, \theta') = \mathbf{E} \{ \mathbf{x}(\theta) \mathbf{y}^H(\theta') \}$ où $\mathbf{E} \{ \cdot \}$ désigne l'opérateur d'espérance mathématique. La matrice $\mathbf{R}_{\mathbf{xy}}(\theta, \theta')$ s'écrit donc :

$$\mathbf{R}_{\mathbf{xy}}(\theta, \theta') = \begin{pmatrix} R_{x_1 y_1}(\theta, \theta') & \dots & R_{x_1 y_n}(\theta, \theta') \\ \vdots & \ddots & \vdots \\ R_{x_n y_1}(\theta, \theta') & \dots & R_{x_n y_n}(\theta, \theta') \end{pmatrix}, \quad (1.39)$$

sachant que chacune des composantes $R_{x_i y_j}(\theta, \theta') = \mathbf{E} \{ x_i(\theta) y_j^*(\theta') \}$, $\forall i, j = 1, \dots, n$, représente l'inter-corrélation entre deux signaux scalaires $x_i(t)$ et $y_j(t)$. Le STFSB $\mathbf{D}_{\mathbf{xy}}(t, \nu; \mathbf{K})$ exprimé au niveau de l'Eq. (1.38) peut lui aussi être mis sous la forme matricielle suivante :

$$\mathbf{D}_{\mathbf{xy}}(t, \nu; \mathbf{K}) = \begin{pmatrix} D_{x_1 y_1}(t, \nu; K_{11}) & \dots & D_{x_1 y_n}(t, \nu; K_{1n}) \\ \vdots & \ddots & \vdots \\ D_{x_n y_1}(t, \nu; K_{n1}) & \dots & D_{x_n y_n}(t, \nu; K_{nn}) \end{pmatrix}, \quad (1.40)$$

avec $D_{x_i y_j}(t, \nu; K_{ij})$, $\forall i, j = 1, \dots, n$, est la transformée bilinéaire associée au couple de deux signaux scalaires $(x_i(t), y_j(t))$. Le Spectre Temps-Fréquence Bilineaire (STFB) $D_{x_i y_j}(t, \nu; K_{ij})$ est paramétré par le noyau K_{ij} est défini pour $i, j = 1, \dots, n$, comme :

$$(x_i(t), y_j(t)) \xleftrightarrow{\text{STFB}} D_{x_i y_j}(t, \nu; K_{ij}) = \int_{\mathbb{R}^2} R_{x_i y_j}(\theta, \theta') K_{ij}(\theta, \theta'; t, \nu) d\theta d\theta', \quad (1.41)$$

Notons que pour des signaux déterministes, il suffit juste de retirer l'espérance mathématique au niveau de la fonction d'inter-corrélation et de la matrice de corrélation, *c.à.d.* $\mathbf{R}_{\mathbf{xy}}(\theta, \theta') = \mathbf{x}(\theta) \mathbf{y}^H(\theta')$ et $R_{x_i y_j}(\theta, \theta') = x_i(\theta) y_j^*(\theta')$, $\forall i, j = 1, \dots, n$. Dans ce cas, $D_{x_i y_j}(t, \nu; K_{ij})$ (resp. $\mathbf{D}_{\mathbf{xy}}(t, \nu; \mathbf{K})$) est qualifié de Transformée Temps-Fréquence (resp. Transformée Temps-Fréquence Spatiale) Bilineaire, notée TTFB (resp. TTFSB).

Spectre et transformée temps-fréquence (spatial) quadratique

Le Spectre Temps-Fréquence (resp. Spectre Temps-Fréquence Spatial) Quadratique STFQ (resp. STFSQ) peut être défini pour deux signaux identiques (x_i, x_i) (resp. pour deux vecteurs identiques $(\mathbf{x}(t), \mathbf{x}(t))$). Ainsi :

$$(x_i(t), x_i(t)) \xleftrightarrow{\text{STFQ}} D_{x_i x_i}(t, \nu; K_{ij}) = D_{x_i}(t, \nu; K_{ij}) = \int_{\mathbb{R}^2} R_{x_i x_i}(\theta, \theta') K_{ij}(\theta, \theta'; t, \nu) d\theta d\theta', \quad (1.42)$$

9. C'est un produit terme à terme des éléments de deux matrices de même dimension.

et

$$\begin{aligned}
(\mathbf{x}(t), \mathbf{x}(t)) &\stackrel{\text{STFSQ}}{\longleftrightarrow} \mathbf{D}_{\mathbf{xx}}(t, \nu; \mathbf{K}) = \int_{\mathbb{R}^2} \mathbf{R}_{\mathbf{xx}}(\theta, \theta') \odot \mathbf{K}(\theta, \theta'; t, \nu) d\theta d\theta' \\
&= \begin{pmatrix} D_{x_1 x_1}(t, \nu; K_{11}) & \dots & D_{x_1 x_n}(t, \nu; K_{1n}) \\ \vdots & \ddots & \vdots \\ D_{x_n x_1}(t, \nu; K_{n1}) & \dots & D_{x_n x_n}(t, \nu; K_{nn}) \end{pmatrix}. \quad (1.43)
\end{aligned}$$

La remarque précédente reste valable à savoir que dans le cas de signaux déterministes, il faut juste enlever l'espérance mathématique au niveau de la fonction d'auto-corrélation $D_{x_i x_i}(t, \nu; K_{ij})$ (resp. de la matrice de corrélation $\mathbf{D}_{\mathbf{xx}}(t, \nu; \mathbf{K})$) pour retomber sur la définition des transformées temps-fréquence quadratiques, notées TTFQ (resp. transformées temps-fréquences spatiale quadratiques, notées TTFSQ).

Les termes diagonaux de la matrice $\mathbf{D}_{\mathbf{xx}}(t, \nu; \mathbf{K})$ sont des termes quadratiques que l'on appelle les "auto-termes", ils correspondent aux STFQ (resp. TTFQ pour des signaux déterministe) de chacune des composantes du vecteur $\mathbf{x}(t)$ de signaux aléatoires (resp. déterministes). Les termes hors diagonale portent l'appellation générique d'"inter-termes". Ils correspondent quant à eux aux STFB ou TTFB associés à des couples de composantes distinctes du vecteur $\mathbf{x}(t)$ de signaux aléatoires (resp. déterministes).

En agissant sur le noyau, il est alors possible d'engendrer les différents éléments de la classe Cohen (Tab. 1.1) dont nous pouvons citer quelques éléments : i) la distribution de Pseudo Wigner (PW), ii) de Pseudo Wigner-Ville (PWV), iii) de Pseudo Wigner lissée (PWL), iv) de Pseudo Wigner-Ville lissée (PWVL) et finalement v) et de Born-Jordan (BJ) (pour plus de représentations et détails, le lecteur peut se référer à [67]).

Transformation	Noyau $\mathbf{K}(t, \nu)$ de la classe Cohen
WV	$\delta(t)$ où $\delta(\cdot)$ désigne la distribution Dirac
PWV	$\delta(t)H(\tau)$ avec $H(\tau) = p^*(\frac{\tau}{2})p(-\frac{\tau}{2})$ une fenêtre glissante
PWVL	$G(t)H(\tau)$ avec $H(\cdot)$ et $G(\cdot)$ deux fenêtres glissantes
CW	$\sqrt{\frac{\sigma}{4\pi}} \frac{1}{ \tau } e^{-\frac{\sigma t^2}{4\tau^2}}$
BJ	$\begin{cases} \frac{1}{\tau}, & t/\tau < 1/2 \\ 0, & t/\tau > 1/2 \end{cases}$

TABLE 1.1 – Quelques RTFQ de la classe Cohen et noyaux correspondants.

Les méthodes de SAS qui se fondent sur les outils qui viennent d'être présentés, sont nées en 1998 [25]. Elles exploitent la non-stationarité des sources. Schématiquement, l'hypothèse d'indépendance des sources est relâchée au bénéfice de l'existence, dans le plan temps-fréquence, d'un nombre suffisant de points pour lesquels les sources se distinguent. La première méthode à avoir été proposée dans [25], passe par une étape préalable de blanchiment des observations, puis consiste à diagonaliser conjointement un ensemble de matrices de RTFSQ associés à des points temps-fréquence d'énergie suffisamment importante et ne correspondant qu'à des auto-termes sur les sources. L'aspect le plus compliqué de ces techniques réside généralement dans la construction des ensembles de matrices de RTFSQ ou de STFSQ (auxquels on appliquera ensuite des algorithmes de décompositions matricielles conjointes) et par conséquent dans la sélection de "bons" points temps-fréquence. Plusieurs travaux ont porté sur cette question de la sélection automatique des points temps-fréquence et la construction des ensembles de matrice à diagonaliser conjointement et/ou à zéro diagonaliser conjointement parmi lesquels [132][96][98] (donnant

lieu à la méthode TFBSS (Time-Frequency Blind Source Separation)), ou encore [116][118][149] et [21][2]. Par la suite ces méthodes ont été généralisées afin de pouvoir s'affranchir de l'étape de blanchiment préalable des observations [115][178] ou pour pouvoir considérer le cas sous-déterminé [149][157].

1.2.2 Mélange convolutif

Avant d'évoquer les principales méthodes de SAS dédiées aux mélanges convolutifs, précisons que la plupart de ces approches exploitent au moins l'une des propriétés suivantes : non gaussianité des signaux, non stationarité des signaux ou encore diversité spectrale.

1.2.2.1 Approches temporelles

Dans l'optique d'étendre les approches temporelles du contexte instantané au contexte convolutif, plusieurs voies ont été explorées, parmi lesquelles :

- les approches temporelles n'exploitant que les Statistique d'Ordre Deux SOD et s'appuyant donc soit sur la non-stationarité des sources soit sur leur cohérence temporelle avec à la fois des méthodes basées sur des décompositions de type sous espace [179][165][3][5][121][166][244] et des méthodes basées sur la Bloc Diagonalisation Conjointe (BDC) unitaire [33][34][82] ou non unitaire [108][110][114][220][55] d'un ensemble de matrices. C'est d'ailleurs sur ce type de décompositions matricielles que porte le deuxième chapitre de ce manuscrit, l'objectif étant de proposer de nouveaux algorithmes de BDC. Dans ce type de décomposition, les matrices recherchées ont une structure algébrique particulières puisqu'elles sont bloc diagonales¹⁰
- les méthodes temporelles exploitant les statistiques d'ordre supérieur (SOS) ont également été généralisées en commençant par exploiter la propriété de sources indépendantes identiquement distribuées (*iid*) tout en continuant à optimiser une fonction de coût de type contraste ([69] par exemple). Moreau et al. ont ensuite proposé une généralisation de la notion de contraste en développant l'idée de contrastes non symétriques [175][176][173]. De nouvelles familles de contrastes n'utilisant que des auto-cumulants [174] ou des cumulants croisés [176][219] ont été proposées. Par la suite, le cas de sources non *iid* a été considéré dans [158][46]. Toutes ces méthodes requièrent une étape de blanchiment des observations puis la recherche d'un filtre para-unitaire. Dans [72][73], Comon et al. ont montré les liens qui existent entre fonctions de contraste et critères de DC qui ont pu être exploités dans le cas des mélanges convolutifs. Ainsi, la diagonalisation conjointe partielle approchée sous contrainte de para-unitarité d'un ensemble de matrices construit à partir des cumulants des observations blanchies a donné lieu à l'algorithme PAJOD (Partial Approximate JOint Diagonalization).

1.2.2.2 Approches fréquentielles

Une autre façon d'aborder ce problème consiste à passer dans le domaine fréquentiel en utilisant soit la TF soit la Transformée de Fourier à Court Terme (TFCT)[135][180][211][188][187][170][193][194][197][199][198] afin d'exploiter le fait que le produit de convolution se transforme en un produit simple dans le domaine de Fourier. A nouveau, la non-stationarité des sources est exploitée dans bon nombre de ces méthodes [188][187][193][194][197][199][198]. Ces approches utilisent les

10. ce sont des matrices bloc dont les termes (bloc) diagonaux sont des matrices carrées tandis que les blocs hors-diagonale sont des matrices nulles.

mêmes critères de séparation de sources que ceux utilisés en contexte instantané à ceci près que les traitements doivent être étendus à tout un ensemble de fréquences distinctes. La difficulté majeure rencontrée au niveau de ces approches reste celle de la reconstruction d’une information cohérente d’une fréquence à l’autre (ceci est dû aux indéterminations d’échelle et de permutation propre à chaque fréquence qui peuvent donc induire des fluctuations d’une fréquence aux fréquences voisines). Plusieurs techniques ont donc été proposées [75][78][76][206][204][207][203][237][177] en réponse à ce problème. Araki et al. ont également montré dans [18] qu’un autre problème subsiste au niveau de la TFCT qui ne conduit qu’à peu de points pour chaque fréquence d’où des complications en terme d’estimation statistique. L’une des premières techniques a été suggérée par C. Servière dans [205].

1.2.2.3 Approches séquentielles

Dans ce paragraphe, nous terminons sur les approches séquentielles qui ne cherchent plus à séparer tous les signaux simultanément mais les extraient les uns après les autres au moyen d’un processus itératif moyennant certaines précautions. La plus connue de ces méthodes est la “déflation” introduite par P. Loubaton [159]. La pertinence de cette approche a été démontrée dans le cas instantané (puis dans le cas convolutif) par N. Delfosse et P. Loubaton dans [84] (en 1997 par J.-K. Tugnait dans [229][230]). Son principe se résume comme suit : ayant estimé une source, il suffit de soustraire sa contribution par une méthode de moindres carrés afin de se ramener à un mélange de $n - 1$ sources puis à recommencer de façon itérative le même processus d’extraction jusqu’à la dernière source. Plus tard, Simon et al. ont montré dans [210] que l’intérêt principal de ces approches était qu’elles restent efficaces même dans le cas non linéaire. Par ailleurs, nous notons enfin qu’il existe des méthodes dites hiérarchiques [225][226] qui consistent à maximiser une série de critères dans lesquels on introduit à chaque étape un terme de pénalisation représentant la décorrélation spatiale des sources estimées [138][45][185].

1.3 Quelques applications de la séparation aveugle de sources

Nombreuses sont les situations dans lesquelles on débouche finalement sur un problème de séparation/démixage de signaux ou d’images si bien que depuis que le problème de la SAS a été formalisé pour la première fois par des chercheurs français en traitement du signal au milieu des années 80 [125][127] jusqu’à aujourd’hui, la SAS a été omniprésente dans différents domaines d’applications parmi lesquels le génie biomédical, le traitement de la parole et des signaux audio, le traitement radar, la séparation de vibrations dans les machines tournantes, le traitement d’images et les télécommunications numériques. Précisons quand même qu’aujourd’hui pour attaquer les problèmes de séparation, on ne se limite plus forcément à des modèles du type de ceux présentés au paragraphe 1.1.1, mais que d’autres modèles mathématiques ont également été très largement étudiés dont les plus connus sont les factorisations matricielles non-négatives (Nonnegative Matrix Factorization (NMF) [151][150]) avec comme domaine d’application phare le traitement d’images (dont l’imagerie hyperspectrale) ou encore les décompositions tensorielles [70] de type décompositions CP (pour Canonical Polyadic Decomposition ou Candecomp/Parafac [128][123]) voire les décompositions tensorielles non négatives (NonNegative Tensor Factorization (NTF) [64]) avec des domaines d’applications comme la chémométrie, l’analyse de donnée environnementale, l’imagerie (donc la spectroscopie de fluorescence 3D) etc. En ce qui concerne les domaines d’application phare de la SAS, on citera en outre :

- l’application à l’audio : la plupart des signaux acoustiques résultent du mélange de plusieurs sources, si bien que le traitement du signal acoustique fut l’un des premiers domaines

dans lequel on a cherché à appliquer les techniques de SAS, le problème traité à l'origine étant celui de la lutte contre l'effet "cocktail party". En effet, lors d'une fête ou d'une prise de son audio dans une salle où se trouvent plusieurs locuteurs, il s'agit de séparer les voix des différentes personnes qui parlent en même temps, de la musique et autres bruits de fond [181][189]. À titre d'exemple, Westner et al. ont utilisé dans [242] le critère d'INFO-MAX (INFOrmation MAXimization), et prouvé qu'il était le plus adapté à la séparation des sons acoustiques lors d'expériences réelles réalisées en studios d'enregistrement. Par la suite, Mitianoudis dans [170] s'en est servi pour séparer des signaux audio au niveau du plan temps-fréquence tout en proposant une solution au problème de permutations. En rehaussement de parole, les auteurs de [30] prennent en compte le bruit blanc ou coloré additif et essaient de l'éliminer à l'aide d'algorithmes de SAS fondés sur la DC.

- l'application aux systèmes de radio-communication et plus particulièrement la téléphonie mobile SDMA (Spatial Division Multiple Access). Xavier et al. adoptent dans [246] une modélisation de type SAS pour les GSM (Global System for Mobile) (système MIMO). Il en va de même en téléphonie à main libre, le signal de parole du locuteur peut être bruité par différentes sources de bruit (musique, voitures, etc.). Dans ce cas là, la SAS est utilisée à des fins de débruitage de la source utile (en l'occurrence la "parole" du locuteur [52]).
- l'application en génie biomédical : la SAS peut être utilisée pour séparer des signaux d'ElectroCardioGramme (ECG), d'ElectroMyoGramme (EMG) ou encore d'ElectroEncéphaloGramme (EEG) [40]. Il est possible d'extraire grâce à la SAS les battements cardiaques du fœtus, à partir de plusieurs signaux ECG (ECG fœtus, ECG mère et le bruit qui peut être, par exemple, la respiration ou toute activité musculaire, etc.) reçus au niveau des capteurs placés sur le ventre de la mère [81]. Dans [256], les auteurs ont proposé une nouvelle méthode basée sur un couplage de Décomposition en Valeurs Singulières (DVS) et de SAS exploitant la propriété de cyclo-stationarité des signaux, ce qui a permis d'améliorer la précision d'extraction des signaux ECG par rapport à la technique fondée sur l'ACI proposée par [81].

Cette liste de domaines d'application de la SAS est bien entendu loin d'être exhaustive (pour un panorama plus complet on pourra consulter le livre [71] ou encore le numéro spécial d'IEEE signal processing magazine de mai 2014, tous deux consacrés à la séparation aveugle de sources). Il en existe bien d'autre comme l'électrotechnique [41][89][105], la géophysique [215], le contrôle du trafic aérien [53][228], la chimie [222], etc.

Nous allons maintenant présenter un nouveau domaine d'application celui de la surveillance maritime et plus spécifiquement le système d'identification automatique spatial, pour lequel nous allons développer de nouvelles approches inspirées de la SAS.

1.4 Le contexte applicatif de la surveillance maritime : le système d'identification automatique

1.4.1 Introduction

Le système d'identification automatique (SIA, en anglais AIS pour Automatic Identification System) inclus dans la convention internationale pour la sauvegarde de la vie Humaine en mer SOLAS¹¹ [139] est un système de transmission de données sur une longueur d'onde radio VHF¹² (ce qui permet une réception à courte et moyenne distance pour tout navire équipé d'une simple

11. Safety Of Life At Sea.

12. pour Very High Frequency.

antenne VHF). Ce système, utilisé entre navires ou de navire à stations de base, est dédié à l'échange automatique d'informations classées en trois catégories :

1. des informations statiques telles que le numéro, l'indicatif d'appel, le nom et le type du navire,
2. bien évidemment, des informations dynamiques telles que la position, la vitesse, la route et le cap réel des navires,
3. en outre, des données propres à chaque traversée des navires telles que leur cargaison, la longueur et la largeur des convois, les ports de départ et d'arrivée sont saisies manuellement par le bord et communiquées par la suite.

Cet échange se fait d'une manière régulière en mode navire-navire et en mode navire-station côtière. Par conséquent, le système AIS est un dispositif :

- de communication,
- de sûreté car il joue le rôle d'outil anti-collision pour les navires de commerce en permettant à chaque navire de "voir" les autres navires présents dans son environnement proche et ce, quelles que soient les conditions de visibilité,
- de sécurité du transport maritime international, normalisé par l'Union Internationale des Télécommunications ITU¹³ et adopté par l'Organisation Maritime Internationale IMO¹⁴. En effet, associé à une station terrestre, l'AIS permet aux autorités portuaires et aux organismes de sauvetage d'améliorer la sécurité en mer par l'assistance à l'efficacité du service de trafic des navires (VTS : Vessel Traffic Service) en réduisant les risques de collision dans leurs zones de couverture.

Enfin précisons qu'il existe deux classes de transpondeurs AIS i) ceux de « Classe A » et ii) ceux de « Classe B ». Les transpondeurs AIS de Classe A ont été rendus obligatoires en 2002 pour tous les navires de commerce (indépendamment de leur taille)

- d'un tonnage brut supérieur à 300 tonnes effectuant des voyages internationaux,
- d'un tonnage brut supérieur à 500 tonnes destinés au cabotage national,

et pour tous les navires de transport de plus de 12 passagers.

Ceux de Classe B sont destinés aux navires non-SOLAS commerciaux et de plaisance.

1.4.2 Principe de fonctionnement du système AIS

Dans ce manuscrit, nous nous intéressons aux systèmes AIS de Classe A, qui utilisent le protocole d'accès multiple à répartition auto-organisé dans le temps SO – TDMA et transmettent :

- d'une part les informations dynamiques toutes les 2 à 10 secondes dans le cas où le navire est en mouvement, (sinon ces dernières sont transmises toutes les 3 minutes), et cela, se fait à une puissance d'émission de 12,5 watts,
- d'autre part, les informations statiques et les données relatives aux traversées des navires émises toutes les 6 minutes.

Par conséquent, les différents types d'informations propres à l'AIS sont transmis à divers intervalles récapitulés dans le Tab. 1.2. Il existe deux fréquences dédiées au fonctionnement alternatif de ce système : l'AIS – 1 à 161,975 MHz et l'AIS – 2 à 162,025 MHz. Chacune des fréquences est répartie en 2250 time-slots (d'une durée de 26.67 ms chacun) réitérés toutes les 60 secondes. D'après le Tab. 1.3, les informations dynamiques du navire occupent des intervalles différents selon la mobilité du navire, le Tab. 1.4 donne en détails la gestion des informations dynamiques

13. En anglais : International Telecommunication Union.

14. En anglais : International Maritime Organization.

Informations du navire	Intervalle de temps
Statiques	– Toutes les 6 minutes. – Une donnée modifiée ou demandée.
Dynamiques	– Bateau immobile : toutes les 3 minutes. – Bateau en mouvement : entre 10 et 2 secondes. – Une donnée demandée.
Relatives au voyage	– Toutes les 6 minutes. – Une donnée modifiée ou demandée.

TABLE 1.2 – Intervalles de temps occupés par les diverses informations du navire.

à travers les intervalles occupés par un navire AIS de Classe A selon la vitesse et le changement du cap.

Conditions dynamiques du navire	Intervalle de génération du rapport
– Navire à l'ancre et se déplace à moins de 3 noeuds.	3 minutes
– Navire à l'ancre et se déplace à plus de 3 noeuds. – Navire à une vitesse de 0 – 14 noeuds.	10 secondes
– Navire à une vitesse de 0 – 14 noeuds avec changement de direction (cap).	$3\frac{1}{3}$ secondes
– Navire à une vitesse de 14 – 23 noeuds.	6 secondes
– Navire à une vitesse de 14 – 23 noeuds avec changement de direction (cap). – Navire à une vitesse supérieure à 23 noeuds. – Navire à une vitesse supérieure à 23 noeuds avec changements de direction (cap).	2 secondes

TABLE 1.3 – Intervalles de temps occupés pour l'émission des informations dynamiques (rapports) AIS de Classe A en fonction du type du navire.

Paramètre	bit	Description
ID message	6	Identifiant du message 1, 2 et 3.
Indicateur de redondance	2	Exploité par le répéteur pour indiquer le nombre de fois qu'un message a été répété, il prend la valeur de 0 (la valeur par défaut) à 3 (ne plus répéter), voir l'Annexe 2 de [161].
ID de l'utilisateur	30	le numéro de l'identité du service mobile maritime MMSI (en anglais pour Maritime Mobile Service Identity), cet unique numéro correspond à l'immatriculation de chaque navire.
L'état de navigation	4	Décrit l'état de navigation du navire à travers un numéro qui est défini entre 0 et 15. Par exemple, – 0 correspond à un navire à moteur en route. – 1 correspond à un navire ancré. – 7 correspond à un navire engagé dans la pêche. – 8 correspond à un voilier. Pour plus d'informations, voir [161].
Le taux de virage (ROT) (en anglais pour Rate of Turn)	8	Indique la vitesse à laquelle un navire est en train de tourner. Cet indicateur est exprimé en degré par minutes.
La vitesse au sol (SOG) (en anglais pour Speed Over Ground)	10	Détermine la vitesse du navire par rapport à la surface de la terre mesurée en $\frac{1}{10}$ noeuds.
La précision de la position	1	C'est un indicateur est définie comme suit : – 1 correspond à une précision élevée ($\leq 10m$). – 0 correspond à une précision faible ou la précision par défaut ($> 10m$)
La longitude	28	La longitude en 1/10000 minutes.
La latitude	27	La latitude en 1/10000 minutes.
La route sur le fond (COG) (en anglais pour Course Over Ground)	12	Il s'agit du sens (en 0,1 degré) dans lequel le navire est en mouvement.
Le cap vrai	9	Donne l'orientation du navire en degrés par rapport au nord dans le sens des aiguilles d'une montre.
L'horodatage	6	Désigne le nombre de secondes écoulées en temps universel coordonné (UTC en anglais pour Universal Time coordinated).
L'indicateur de manoeuvre spéciale	2	Indique la manoeuvre spéciale du navire, cet indicateur prend les valeurs suivants : – 0 correspond au cas indisponible ou par défaut. – 1 correspond à un navire désengagé. – 2 correspond à un navire engagé dans une manoeuvre spéciale.
Réservé	3	Ce champs non utilisé est réservé pour une utilisation dans le future. Il est donc réglé à zéro.
La vérification autonome de l'intégrité du positionnement RAIM (en anglais pour Receiver Autonomous Integrity Monitoring)	1	Il indique si l'utilisateur occupe la vérification autonome de l'intégrité du positionnement, 0 correspond à RAIM non utilisé ou par défaut, 1 correspond à RAIM utilisé.
L'état de la communication	19	Indique l'état de communication et de synchronisation. Il est utilisé dans la planification de la prochaine transmission afin d'éviter les interférences mutuelles.

TABLE 1.4 – Rapport de position AIS pour les navires de classe A (messages 1, 2 et 3).

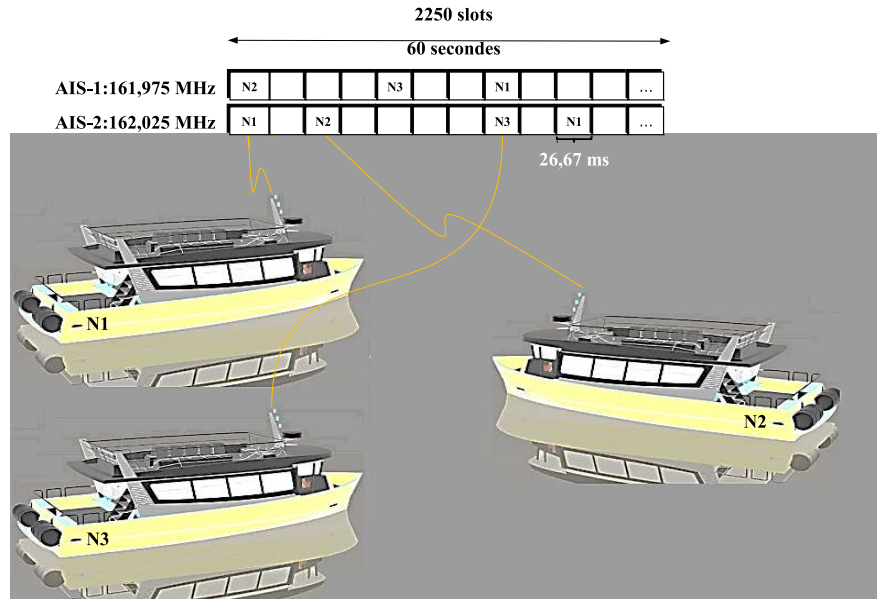


FIGURE 1.2 – Principe du protocole SO – TDMA utilisé pour le système AIS.

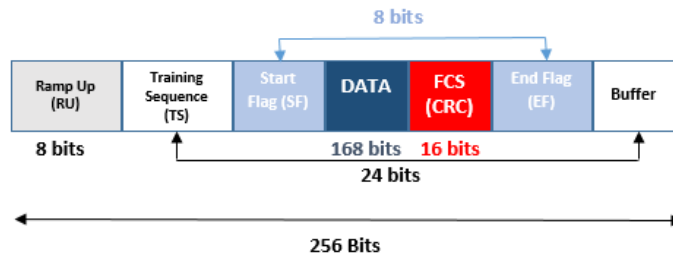


FIGURE 1.3 – Structure de la trame AIS.

1.4.3 Structure des messages AIS

Le message AIS comme illustré au niveau de la Fig. 1.3 est d'une longueur de $N = 256$ bits. Il contient une séquence d'apprentissage (TS)¹⁵ composé de 0 et 1 et qui prend 24 bits. Le fanion de début/Fin (SF/EF)¹⁶ qui prend 8 bits. Une séquence de contrôle de redondance cyclic (CRC)¹⁷ ou séquence de contrôle de trame (FCS)¹⁸ d'une longueur de 16 bits est ajoutée aux différents types d'information de 168 bits propres aux navires qui appartiennent à la zone de navigation. Un bit 0 dénommé le bit de bourrage est inséré après tous les cinq 1 successifs appartenant au bloc des données et au bloc CRC/FCS.

Les trames AIS détaillées ci-dessus subissent une chaîne de mise en forme. Premièrement, ces trames passent par un bloc fondamental dans le système AIS qui est le codage en ligne. Puis, après le codage en ligne, les trames sont modulées. Dans le paragraphe suivant, nous allons

15. En anglais pour Training Sequence.

16. En anglais pour Start/End Flag.

17. En anglais Cyclic Redundancy Code.

18. En anglais Frame Control Sequence.

présenter le codage en ligne NRZI¹⁹ et la modulation GMSK²⁰ normalisés au niveau du système AIS.

1.4.4 Codage différentiel en ligne NRZI

Le codage utilisé par le système AIS est le codage NRZI. Ce type de codage n'utilise pas directement le niveau mais le changement du niveau pour coder les valeurs logiques 0 et 1, *c.à.d* la sortie change de niveau pour représenter un bit 0 et reste constante pour un bit 1, il s'agit donc d'un codage différentiel. Les symboles s'écrivent sous la forme suivante :

$$a_n = b_n a_{n-1}, n \geq 1, \quad (1.44)$$

a_n est la séquence d'information et $b_n \in \{-1, 1\}$.

1.4.5 Modulation GMSK

La modulation à déplacement minimum gaussien (GMSK) avec son enveloppe constante et son spectre compact est une technique de modulation dédiée pour les communications par satellite et les application de la radio mobile [182]. Cette modulation est dérivée du procédé ordinaire de la modulation par déplacement de fréquence FSK²¹, qui est une version numérique de la modulation de fréquence FM²². Comme toute modulation numérique, la GMSK est décomposée en trois étapes principales :

- La première étape est le codage binaire à symbole (ou mapping en anglais). Cela implique qu'à chaque élément a_n appelé symbole est associé un ou plusieurs bits issus de la source d'informations, d'où le terme "débit symbole" signifiant le nombre de symboles transmis par unité de temps. On peut le mesurer en bauds et le déterminer par le débit binaire divisé par le nombre de bits représenté par chaque symbole.
- La deuxième étape est celle de mise en forme permettant de transformer la suite de symboles en un signal qui sera adapté au canal de transmission. Cette mise en forme est en général réalisée par une opération de filtrage linéaire. Le signal s'écrit alors de la façon suivante :

$$s_g(t) = \sum_{n=0}^{+\infty} a_n g(t - nT_s), \quad (1.45)$$

avec a_n les symboles émis, T_s la période symbole et $g(t)$ le filtre de mise en forme. Dans le cas de la modulation GMSK, $g(t)$ n'est rien d'autre qu'une Gaussienne dont l'expression est donnée par :

$$g(t) = \sqrt{\frac{2\pi}{\log 2}} B \exp\left(-\frac{2}{\log 2}(\pi B t)^2\right), \quad (1.46)$$

avec B la bande passante du filtre Gaussien (le produit temps bande spécifié dans le système AIS est $BT_s = 0.4$ avec $T_s = \frac{1}{9600} s$).

- La dernière étape est la modulation, qui consiste à placer le signal sur la fréquence f_c d'une porteuse. L'opération de modulation consiste alors à produire un signal de caractéristique

19. Pour Non Return to Zero Inverse.

20. Pour Gaussian Minimum Shift Keying.

21. Pour Frequency Shift Keying.

22. Pour Frequency Modulation.

spectrale adaptée à la transmission sur un canal de type passe-bande. Le signal modulé en GMSK peut s'écrire de la manière suivante :

$$s(t) = \Re\{e^{-j(2\pi f_c t + \phi(t))}\} = I(t) \cos(2\pi f_c t) - Q(t) \sin(2\pi f_c t), \quad (1.47)$$

avec $\Re\{\cdot\}$ la partie réelle du signal complexe, $\phi(t)$ la phase instantanée de $s_g(t)$, $I(t)$ (resp. $Q(t)$) module la porteuse en phase (resp. en quadrature de phase). Ici, nous parlons du canal réel I (resp. imaginaire Q).

Toutes les étapes de la modulation GMSK présentées ci-dessus, sont illustrées au niveau de la Fig. 1.4.

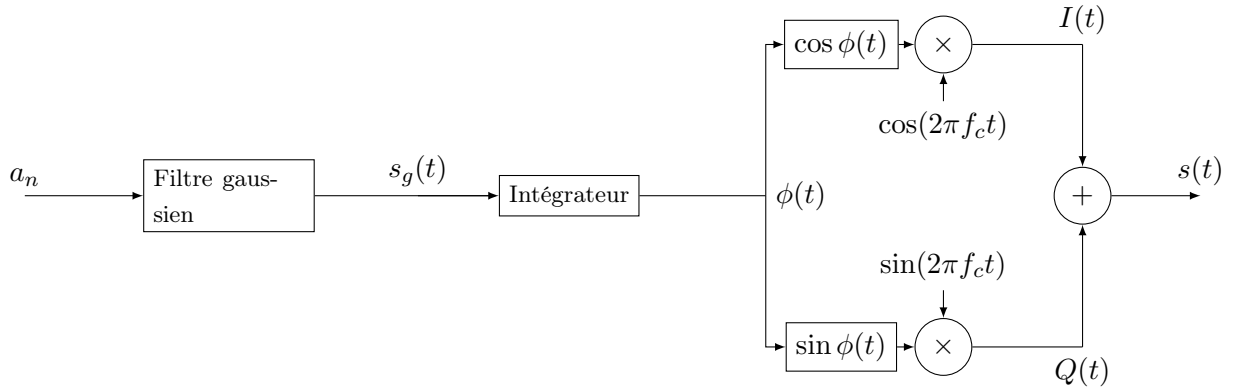


FIGURE 1.4 – Dessin synoptique résumant les différentes étapes de la modulation GMSK.

1.5 Réception des messages au niveau d'un récepteur satellitaire

1.5.1 Introduction

À l'échelle internationale, le trafic maritime a connu une évolution très importante, étant donné que les activités liées au transport de marchandises, au tourisme mais aussi à la contrebande de marchandises et à la clandestinité ont énormément augmenté au cours des dernières années. La réception de messages AIS par un navire ou une station de contrôle côtière atteint ses limites, et il est devenu nécessaire voire urgent d'adopter un nouveau système qui permet une surveillance à plus grande échelle (voire une surveillance mondiale) tout en préservant la fiabilité, l'efficacité et la robustesse du système original.

Cette problématique a donné lieu à une étude de faisabilité quant à l'élaboration d'un système de contrôle plus large des navires et qui assurerait une surveillance mondiale dans la profondeur et la réactivité [87]. Cette étude a été pilotée par le royaume de Norvège, et plus précisément par le FFI²³ puis mise à disposition des autres pays. L'idée qui gouvernait cette étude était d'étendre le système AIS à une surveillance depuis l'espace (au moyen d'un satellite) d'où la terminologie S – AIS (spatial - AIS). L'objectif visé par ce système, dont le principe est illustré au niveau de la Fig. 1.5 est l'amélioration de la surveillance maritime du grand-Nord et tout le long des côtes norvégiennes [133]. L'étude menée par le FFI a montré qu'un satellite en orbite terrestre

23. pour Forsvarets ForskningsInstitutt en Norvégien.

basse de type LEO²⁴ (situé à une distance ≤ 1000 km) équipé d'une antenne de récepteurs AIS pourrait permettre - moyennant quelques aménagements du système initial - la réception des messages AIS et donc le contrôle de l'énorme densité de trafic dans cette zone. Ce système serait beaucoup plus efficace car il permettrait de couvrir de larges zones de l'océan voire d'atteindre une couverture mondiale des navires de même qu'une localisation en temps réel de ceux-ci par les exploitants.

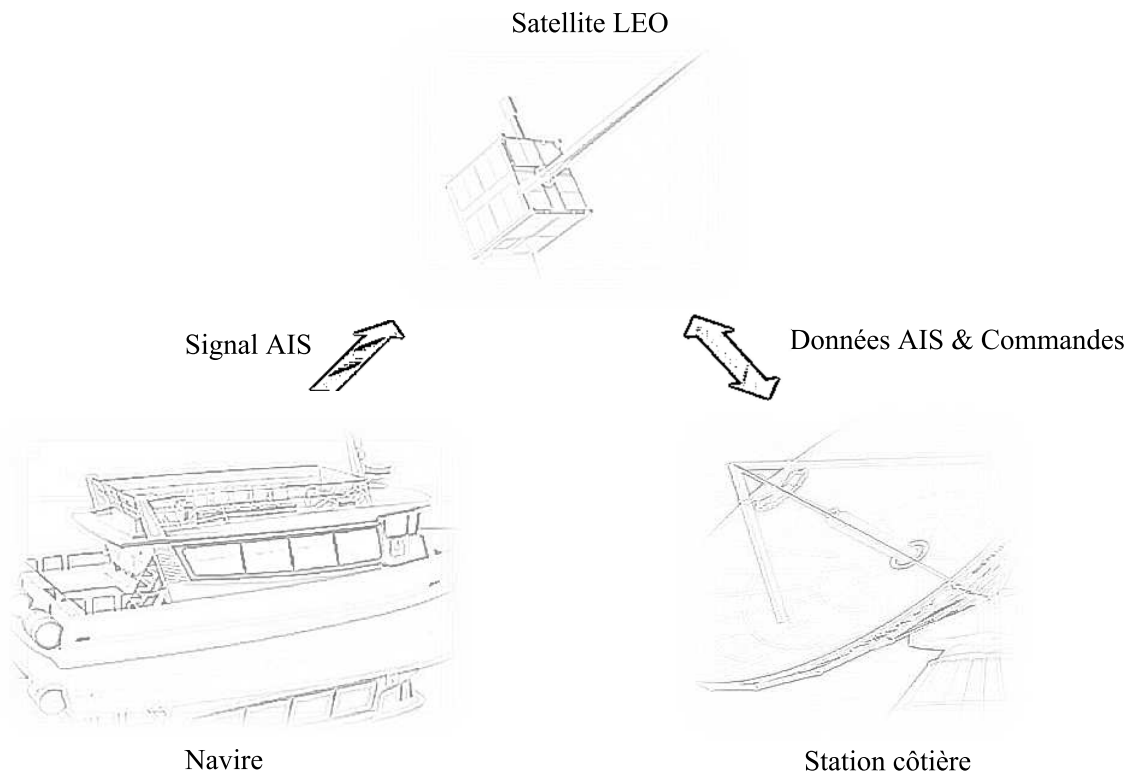


FIGURE 1.5 – Scénario de la détection des signaux AIS par un satellite de type LEO qui traite les signaux AIS et les envoie aux stations côtières à proximité.

En premier lieu, c'est la possibilité de recevoir des messages AIS depuis l'espace qui a été étudiée dans [235]. Le concept a ensuite été présenté dans [87] afin de valider le principe d'un récepteur AIS spatial lequel requiert des ajouts mineurs par rapport au système initial. Les analyses effectuées montrent que la réception du message d'origine en spatial est possible : en terme de puissance du signal, on observe une marge de 10 à 20 dB par rapport à la sensibilité d'un récepteur standard et la probabilité de détection est très élevée puisqu'elle peut même atteindre les 99 % pour environ 1000 navires.

1.5.2 Problèmes liés à une réception spatiale des messages

La généralisation du système AIS à l'espace met en exergue plusieurs phénomènes/problèmes qu'on ne rencontrait pas lorsque la surveillance des navires s'effectuait au niveau du sol soit depuis un autre navire soit depuis une station côtière à terre. Parmi ces phénomènes, on trouve :

24. pour Low Earth Orbit.

- l'effet Doppler engendré par la vitesse de déplacement des satellites qui crée des décalages en fréquence au niveau des signaux S – AIS,
- le délai de propagation des signaux et l'atténuation de ceux-ci du fait de l'altitude à laquelle sont situés les satellites,
- enfin comme une zone beaucoup plus large est couverte par le satellite, elle englobe donc forcément plusieurs cellules AIS traditionnelles. A cela s'ajoute également le fait que les temps de propagation des signaux transmis des navires au satellite peuvent fluctuer en fonction de la position des navires et de la zone de couverture maximale de l'antenne du satellite. L'effet conjugué de ces deux problèmes affecte le mécanisme d'organisation des signaux du système S – AIS. Il en résulte des risques de collision des données émises par des navires localisés dans des cellules AIS différentes, mais reçues au niveau de l'antenne d'un même satellite.

Dans ce manuscrit, nous nous focalisons principalement sur ce phénomène de collision des signaux AIS en proposant de nouvelles méthodes de “décollision” inspirées des méthodes de séparation aveugle de sources. L'effet Doppler, le délai et l'amplitude seront pris en considération au niveau de nos méthodes.

Effet Doppler, atténuation et collision des signaux S – AIS - Comme illustré au niveau de la Fig. 1.5, le système S – AIS reçoit les messages issus des navires au niveau d'un satellite LEO qui retransmet les informations reçues aux centres de surveillance (stations côtières) afin de surveiller les navires présents dans une zone beaucoup plus étendue. Ces signaux sont affectés par la propagation du navire au satellite ce qui induit d'une part une atténuation, un retard voire même l'entrée en collision des signaux S – AIS reçus au niveau de l'antenne du satellite, et d'autre part, un décalage en fréquence nommé “effet Doppler” causé par le mouvement du satellite. L'ensemble de ces phénomènes a un impact significatif sur ces signaux [39].

Le décalage de fréquence est relatif à la différence de vitesse entre l'émetteur et le récepteur. Dans le cas du S – AIS, la vitesse du navire est faible (ou négligeable) par rapport à celle du satellite. L'effet Doppler peut être calculé suivant la relation suivante, $\Delta f = \frac{V_{sat}}{\lambda}$, avec $\lambda = \frac{c}{f}$ ($= 1,86$ m) la longueur d'onde du signal AIS original et V_{sat} la composante de la vitesse du satellite dirigée vers le navire. Cette composante varie selon l'élévation (plan vertical) et l'azimut²⁵ de 0° (plan horizontal) du satellite ($\Delta f = 0$). Plus précisément, avec un satellite de type LEO tel qu'on le préconise dans le système S – AIS, le décalage Doppler maximal peut atteindre 4 kHz. En raison de la symétrie de la zone de couverture du satellite, le décalage Doppler est alors d'environ ± 4 kHz [39].

Pour le système AIS terrestre, la longueur des messages admettait un retard de propagation entre les différents navires d'une même cellule AIS de l'ordre de 2ms maximum. En revanche, dans le cas du système S – AIS, les retards de propagation qui varient en fonction de la position des navires et l'empreinte maximale de l'antenne du satellite LEO vont pouvoir atteindre jusqu'à 10ms [162]. Cela peut produire un dépassement du délai du buffer et ce même si les trames S – AIS ont été transmises dans des slots TDMA différents. Ce phénomène peut donc être à l'origine de collision entre les signaux émis par des navires différents.

Par ailleurs, le rayon de couverture d'une cellule SO – TDMA est d'environ 40 mm²⁶, il n'y avait donc pas de problème de collision des signaux AIS émis par les navire de la cellule. Dans le cas où la réception se fait au niveau du satellite, ce dernier couvre plusieurs cellules SO – TDMA, si bien

25. exprimé en degrés de 0° à 360° en partant du Nord et en tournant dans le sens des aiguilles d'une montre.

26. mm : mille marin est l'unité de mesure de distance utilisée pour la navigation maritime, 1 mm correspond à 1852 mètres.

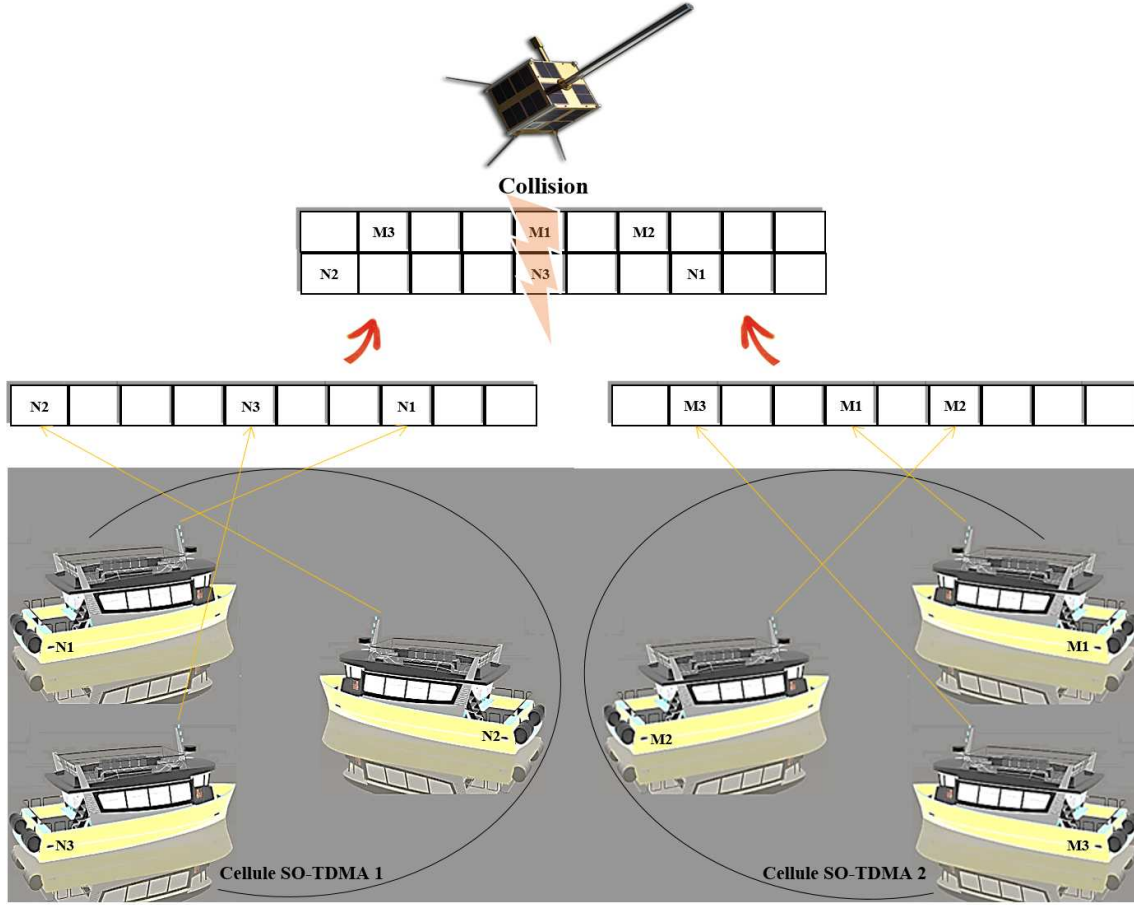


FIGURE 1.6 – Scénario de signaux AIS en provenance de cellules SO – TDMA différentes mais reçus par le satellite au même instant (collision).

que la structure du SO – TDMA en est perturbée : les signaux transmis par des navires situés dans des cellules différentes peuvent être reçus au même time-slot, d'où là encore une possible collision des signaux [231][251]. Ce phénomène est illustré au niveau de la Fig. 1.6.

1.5.3 Modélisation du problème de collision des messages AIS

Le signal reçu au niveau du k -ème capteur de l'antenne du satellite peut être modélisé de la façon suivante :

$$x_k(t) = \sum_{j=1}^J h_{kj} s_j(t - \tau_{kj}) e^{-i2\pi \Delta f_{kj} t} + n_k(t), \quad (1.48)$$

où

- h_{kj} sont les coefficients liés au canal de transmission du j -ème navire jusqu'au k -ème capteur de l'antenne embarquée dans le satellite
- $n_k(t)$ est un bruit additif supposé blanc, stationnaire gaussien, iid (indépendant identiquement distribué) et indépendant des messages émis s_j
- J est le nombre de signaux reçus (nombre de navires dans la zone surveillée)

- τ_{kj} (resp. Δf_{kj}) est le retard (resp. le décalage Doppler) associé au j -ème signal reçu sur le k -ème capteur.

Réécriture en un problème de séparation aveugle de sources

A notre connaissance, pour solutionner ce problème de collisions des messages, peu de travaux ont porté sur des méthodes de séparation aveugle de sources. Dans [263], Zhou *et al.* présentent un récepteur multi-utilisateurs équipé d'un réseau d'antennes embarqués dans des satellites de type LEO. Le principe de ce récepteur est d'exploiter le multiplexage spatial dans un contexte non-stationnaire asynchrone. En effet, les auteurs considèrent une trame de données contenant d messages $S - AIS$ asynchrones chacun issu d'un des d utilisateurs. Les signaux reçus sont démodulés en bande de base et échantillonnés à la période $T_s = T/P$, avec P le rapport de sur-échantillonnage. En regroupant les PN échantillons sur des périodes de N symboles, le modèle de la trame de données considéré s'écrit alors :

$$\mathbf{X} = \mathbf{H}\mathbf{G}(\mathbf{S} \odot \Phi) + \mathbf{N}, \quad (1.49)$$

où $\odot$ désigne l'opérateur de Schur-Hadamard, $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{PN}] \in \mathbb{C}^{M \times PN}$, $\mathbf{x}_n = \mathbf{x}(nT_s)$, $1 \leq n \leq PN$, est la matrice des observations, $\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_d] \in \mathbb{C}^{M \times d}$ est la matrice représentative de la réponse de l'antenne, $\mathbf{G} = \text{diag}\{g_1, g_2, \dots, g_d\} \in \mathbb{R}^{d \times d}$ contient la puissance des sources, $\mathbf{S} = [\mathbf{s}_1^H, \mathbf{s}_2^H, \dots, \mathbf{s}_d^H]^H \in \mathbb{C}^{d \times PN}$ est la matrice des sources et

$$\Phi = \begin{pmatrix} 1 & \varphi_1^1 & \dots & \varphi_1^{PN-1} \\ 1 & \varphi_2^1 & \dots & \varphi_2^{PN-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \varphi_d^1 & \dots & \varphi_d^{PN-1} \end{pmatrix}, \quad (1.50)$$

où $\varphi_k = e^{j2\pi\Delta f_k T_s}$ contient les fréquences Doppler des sources.

Cette méthode se décompose en plusieurs étapes. La première consiste à déterminer le début des trames de données en exploitant la technique dite du "suivi de rang"²⁷. Elle utilise un algorithme de suivi de sous-espace basé sur la décomposition de URV signée lequel a été proposé dans [262]. Il est noté SURV²⁸. Cependant les trames de données ainsi regroupées contiennent des messages $S - AIS$ asynchrones parfois incomplets ce qui peut nuire à la bonne convergence des algorithmes utilisés ensuite au niveau de l'étape de séparation. C'est pourquoi un traitement appelé "intersection de sous-espaces" est effectué afin de trouver les sous-espaces sans interférence et parvenir ainsi à éliminer les données incomplètes (en passant par plusieurs étapes de traitement dont une projection sur le sous-espace principal, une subdivision de la matrice obtenues en deux blocs puis à nouveau une décomposition SURV et une projection). Une nouvelle matrice d'observation ne contenant plus que des messages complets peut alors être construite afin de procéder à l'étape de séparation : celle-ci est basée sur l'algorithme algébrique de module constant encore noté ACMA (les signaux $S - AIS$ possédant une enveloppe constante) proposé par A.-J. Van der Veen dans [233]. Cet algorithme est lui-même fondé sur des techniques de diagonalisation conjointe d'ensembles de matrices afin de reconstruire les sources $S - AIS$.

Modélisation proposée : nous pouvons reformuler d'une manière plus simple le problème de la collision entre les signaux $S - AIS$ provenant des différents navires en le réécrivant comme un

²⁷. En anglais rank tracking.

²⁸. un algorithme de mise à jour de type Jacobi pour DVS.

problème de séparation aveugle de sources. Tout d'abord, précisons que nos développements sont basés sur les hypothèses suivantes :

Hyp₁. Les bruits $n_j(t)$ pour $j = 1, \dots, m$ sont stationnaires, blancs, de moyenne nulle. Ce sont des signaux aléatoires mutuellement décorrélés et indépendants des sources avec une variance σ_n^2 .

Hyp₂. Le nombre de capteurs m et le nombre de sources n sont supposés connus et $m \geq n$ pour conserver un modèle sur-déterminé.

Hyp₃. La distance entre les capteurs de l'antenne est très faible par rapport à la distance entre l'émetteur et le satellite. On suppose que la j -ème source arrive avec un retard et une fréquence Doppler sur les m capteurs : $\forall k = 1 \dots m, \tau_{kj} = \tau_j$ et $\Delta f_{kj} = \Delta f_j$.

L'Eq. (1.48) peut être réécrite sous une forme plus compacte en $\mathbf{x}(t) = \{\mathbf{H} *_t \mathbf{s}\}(t) + \mathbf{n}(t)$, où $\{\cdot *_t \cdot\}$ désigne l'opérateur de convolution. Toutefois, du fait de l'**Hyp₃**, cette équation va finalement être réécrite sous la forme simplifiée suivante :

$$\mathbf{x}(t) = \mathbf{H}\mathbf{s}'(t) + \mathbf{n}(t), \quad (1.51)$$

où $\mathbf{H} = (h_{kj})$ désigne la matrice de mélange de dimension $m \times n$, $\mathbf{s}'(t) = [s'_1(t), s'_2(t), \dots, s'_n(t)]^T$ est le vecteur de dimension $n \times 1$ des sources avec $s'_j(t) = s_j(t - \tau_j) \exp\{-i2\pi\Delta f_j t\}, \forall j = 1, \dots, n$ et $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_m(t)]^T$, $\mathbf{n}(t) = [n_1(t), n_2(t), \dots, n_m(t)]^T$ sont des vecteurs de dimension $m \times 1$ correspondant respectivement au vecteur des observations et à celui des bruits. La forme de l'Eq. (1.51) est alors similaire à celle de l'Eq. (1.1) tout en prenant en compte les différents paramètres (retard et effet Doppler) causés par l'éloignement du satellite considéré et son déplacement. Chaque source $s'_j(t)$ correspond en effet au message AIS d'origine affecté d'un retard τ_j et d'un effet Doppler $\exp\{-i2\pi\Delta f_j t\}$.

Notre objectif sera maintenant d'appliquer à ce modèle de données des techniques de SAS pour séparer (dé-collisionner) les signaux reçus au niveau du récepteur AIS spatial et ré-estimer chacune des sources (messages S – AIS) aux indéterminations de la SAS près. Précisons finalement que le système mélangeant que nous avons choisi de considérer pour l'application en surveillance maritime que nous étudierons au Chapitre 3 sera donc de type simple trajet (ou canal sans mémoire) en raison de nos hypothèses. Il correspond au cas d'un mélange instantané de signaux (pas de multi-trajet du fait de l'absence d'obstacles en plein océan). Toutefois, si les hypothèses que nous avons choisies de considérer s'avéraient trop restrictives conduisant à un modèle trop simplifié pour réellement s'appliquer en contexte réel, les algorithmes de (zéro) bloc-diagonalisation conjointe que nous allons développer au chapitre suivant sont très généraux et permettront donc de revenir à un modèle convolutif. Les traitements suggérés au Chapitre 3 pourront eux aussi se généraliser au modèle convolutif en utilisant pour ce faire les mêmes techniques que celles développées dans [107] pour d'autres classes de signaux (signaux déterministes, éventuellement corrélés).

1.6 Conclusion

Dans ce premier chapitre, nous avons fourni un aperçu général de ce qu'est la séparation aveugle de sources rappelant la manière de modéliser ce problème aussi bien dans le cas de mélanges instantanés que convolutifs. Nous avons également rappelé la manière de réécrire le modèle

convolutif comme un modèle instantané, de même que les hypothèses classiques de la SAS et ses indéterminations (les sources séparées ne sont estimées qu'à une permutation et à un facteur d'échelle près en modèle linéaire instantané ou à un filtrage près en modèle linéaire convolutif). Nous avons ensuite décrit brièvement certaines méthodes de SAS puis rappelé quelles étaient les principales approches de la littérature en contexte instantané en nous intéressant plus particulièrement à celles optimisant des fonctions de contrastes et utilisant d'une manière directe ou indirecte des décompositions matricielles conjointes (la DC et/ou la ZDC) d'ensembles judicieusement choisis de matrices ou de coupes de tenseurs. Par ailleurs, les approches développées en contexte convolutif ont également été évoquées qu'il s'agisse d'approches temporelles exploitant les SOD ou les SOS, ou d'approches fréquentielles qui transforment le modèle convolutif en un modèle instantané ou encore d'approches séquentielles qui se distinguent des approches précédentes par le fait que les sources ne sont pas extraites simultanément, mais les unes après les autres. Ensuite, nous avons rappelé que la SAS possède plusieurs domaines d'application. Nous avons alors introduit le contexte de la surveillance maritime et plus spécifiquement le système d'identification automatique spatial sur lequel nous allons concentrer nos travaux de recherche. Après avoir rappelé le principe du système d'identification automatique classique (en mer ou mer-sol) et en avoir détaillé les phases de transmission et de réception, nous sommes passés au système S – AIS dans lequel la surveillance est pilotée par un satellite de type LEO. Un certain nombre de problèmes apparaissent alors dont celui de la possible entrée en collision des messages S – AIS. Nous avons ensuite montré que ce problème pouvait se réécrire comme un problème de SAS en tenant compte des paramètres inhérents au mouvement du satellite (effet Doppler et retard dû à la propagation du navire au satellite).

Chapitre 2

Problèmes de la bloc-diagonalisation et de la zéro bloc-diagonalisation conjointe et nouveaux algorithmes

Avec l'utilisation d'antennes équipées de toujours plus de capteurs, mais aussi avec la mise en place de réseaux de capteurs (pour la surveillance ou le monitoring de l'environnement, du trafic, etc...), la quantité de données disponibles ne cesse de croître et des techniques de traitement du signal avancées (de même que les algorithmes qui en découlent) deviennent une quasi nécessité pour parvenir à extraire des informations pertinentes de ces grandes masses de données. D'une certaine manière, les algorithmes de décompositions conjointes d'ensembles de matrices (ou de tenseurs) répondent à cet impératif de pouvoir disposer d'outils toujours plus sophistiqués et efficaces pour traiter les données. C'est la raison pour laquelle le problème des décompositions conjointes d'ensembles de matrices ou de tenseurs a su attirer l'attention de la communauté scientifique du traitement du signal et de nombreux travaux lui ont donc été consacrés au cours des dernières années. Tout domaine d'application confondu, à partir du moment où le problème traité fini par se réécrire comme un problème :

- de traitement d'antenne et d'estimation des directions d'arrivées,
- ou encore de séparation aveugle de sources ou de démixage de signaux, d'images, etc...,
- de déconvolution aveugle multidimensionnelle,
- d'analyse ou de fouille de données,

ce type d'algorithmes est susceptible de présenter un intérêt.

Parce que traditionnellement, on utilisait une étape préalable de blanchiment spatial des observations, l'un des premiers problèmes de décompositions conjointes à avoir été considéré fut celui de la diagonalisation conjointe d'un ensemble de matrices donné sous contrainte d'unitarité de la matrice recherchée. Cela a conduit à deux algorithmes aujourd'hui bien connus : JADE [43] et SOBI [22] (§1.2.1). Les travaux qui suivirent ont traité le problème de la diagonalisation conjointe de tenseurs [68][171] ou celui de la diagonalisation conjointe d'ensembles de matrices sans la moindre contrainte d'unitarité cette fois [252][253][83][154][212][221][48] [208][169]. Il est possible d'obtenir un aperçu relativement exhaustif des différentes approches dans [50]. Ce premier cas particulier de décompositions matricielles conjointe s'est avéré utile dans la résolution de deux grandes catégories de problèmes : i) la localisation de sources et la radiogoniométrie²⁹, ii) la séparation aveugle de mélanges instantanés de sources.

29. La radiogoniométrie est la détermination de la direction d'arrivée d'une onde électromagnétique.

Parallèlement, plusieurs chercheurs ont commencé à s'intéresser à un second type de décompositions matricielles conjointes - à savoir - la zéro diagonalisation conjointe car il est fréquent de rencontrer ce problème dans différents domaines d'applications, parmi lesquels la SAS (et plus spécifiquement, lorsque l'on recourt à des méthodes fondées sur l'utilisation de représentations temps-fréquence spatiales quadratiques [21][24]), les télécommunications [47] et encore la cryptographie [236]. Les premiers algorithmes proposés reposaient sur une contrainte d'unitarité du zéro-diagonaliseur conjoint recherché [21], ce qui là encore impliquait de les appliquer après une étape préalable de blanchiment spatial des observations (§1.1.4). Cependant, comme J.-F. Cardoso et L. De Lathauwer l'ont remarqué en SAS [42][80], une telle étape influe sur les performances de séparation en les bornant, c'est pourquoi plusieurs solutions ont ensuite été proposées, parmi lesquelles [92][90][49][259], afin de lever cette contrainte d'unitarité.

Récemment, un problème encore plus général a été considéré : celui de la décomposition conjointe d'une combinaison d'ensembles de matrices complexes donnés pouvant potentiellement suivre différents types de décompositions [255][172][227]. À titre d'exemple, lorsque des signaux non circulaires complexes sont considérés et que l'on cherche à exploiter davantage d'informations statistiques, on peut être amené à utiliser simultanément deux ensembles de matrices : i) celui des matrices symétriques complexes et ii) celui des matrices Hermitiennes. Ce type de problématique se rencontre dans certains problèmes de traitement du signal, parmi lesquels l'identification aveugle ou encore la séparation aveugle de sources.

Nous nous consacrons la première partie de ce chapitre à un quatrième type de décompositions matricielles conjointes baptisé "bloc-diagonalisation conjointe (BDC)". C'est un problème relativement général puisqu'il englobe le problème de la DC, cité précédemment. Notons toutefois, qu'il pourrait n'être considéré que comme un sous problème du troisième type de décompositions, même si - à notre connaissance - jusqu'à aujourd'hui, seuls les problèmes de DC ont été considérés dans ce contexte là. Revenons au problème de la bloc-diagonalisation conjointe qui nous intéresse plus particulièrement : dans une telle décomposition, les matrices étudiées ont une structure algébrique bien spécifique puisque ce sont des matrices blocs diagonales. Rappelons qu'une matrice bloc-diagonale est une matrice bloc dont les blocs diagonaux sont des matrices carrées de n'importe quelle taille (éventuellement cette taille pourra être la même pour tous les blocs diagonaux, hypothèse couramment utilisée en SAS) tandis que les blocs hors diagonale sont des matrices nulles. Tout d'abord, ce problème fut traité en considérant le cas d'ensembles de matrices bloc-diagonales Hermitiennes et définies positives et d'un bloc-diagonaliseur conjoint unitaire. Dans ce cadre là, il a été suggéré de recourir à des algorithmes de type Jacobi [23][82]. Par la suite, plusieurs méthodes alternatives ont été proposées réalisant la même opération, mais cette fois-ci pour un bloc-diagonaliseur non nécessairement unitaire [108][113][249][110][184][148][220][55].

Dans ce chapitre, nous proposerons plusieurs nouveaux algorithmes itératifs de BDC non unitaire d'un ensemble donné de matrices (réelles ou complexes). Ils sont tous fondés sur l'optimisation d'une même fonction de coût quadratique de type moindres carrés. Cette fonction de coût est écrite en fonction de la matrice de séparation et non de la matrice de mélange, ceci dans un but de simplification des développements mathématiques afférents. Dans un premier temps, nous présenterons une solution basée sur l'algorithme du Gradient conjugué noté BDC_{GC} : il requiert le calcul exact du Gradient matriciel (ce calcul a été établi par H. Ghennioui dans sa thèse de doctorat et publié dans [113]). Plusieurs autres algorithmes nécessitant l'utilisation d'un préconditionneur seront ensuite proposés. C'est la raison pour laquelle, nous calculerons de manière exacte les matrices Hessiennes complexes. Ce calcul est réalisé, nous pourrions alors en déduire différents algorithmes lesquels dépendent du schéma d'optimisation utilisé. Nous présentons donc de nouveaux algorithmes pour la BDC_{NU} fondés respectivement sur l'utilisation de i) l'algorithme de Levenberg-Marquardt (LMA) baptisé $BDC_{LM_{HE}}$, ii) l'algorithme du Gradient conjugué précon-

ditionné baptisé $\text{BDC}_{\text{GCP}_{\text{HE}}}$ et finalement iii) sur une approche de type Quasi-Newton (QN) notée $\text{BDC}_{\text{QN}_{\text{HE}}}$. Parmi nos objectifs, le but est d'étudier l'impact voire l'intérêt d'un préconditionnement dans les algorithmes de BDC mais également d'établir quelle est la méthode qui offre le meilleur compromis entre efficacité et vitesse de convergence dans le contexte applicatif qui est le nôtre (à savoir la SAS et la déconvolution aveugle multidimensionnelle pour des problèmes de surveillance maritime par satellite).

Afin d'accélérer la convergence et d'assurer la stabilité numérique des différentes approches présentées, le pas d'adaptation "optimal" est également calculé de manière algébrique à chaque itération comme l'une des racines d'un polynôme d'ordre trois. L'avantage majeur et commun de tous ces algorithmes est qu'ils sont relativement généraux et qu'ils peuvent être appliqués à un ensemble quelconque de matrices (elles pourront être réelles ou complexes, ni définies positives ni à symétrie Hermitienne). Enfin, le bloc-diagonaliseur conjoint n'est pas supposé être une matrice unitaire.

Dans un second temps, nous nous intéresserons à un "nouveau" problème de décompositions matricielles conjointes d'un ensemble donné de matrices complexes que nous avons baptisé la ZBDC_{NU} du fait qu'il généralise la zéro diagonalisation conjointe non-unitaire [90][91]. Nous commençons par introduire la fonction de coût à minimiser pour estimer le zéro-bloc diagonaliseur conjoint. Nous passons ensuite au calcul des quantités nécessaires à l'établissement d'algorithmes permettant la réalisation de ce type de décomposition : gradient matriciel complexe et matrices Hessiennes complexes. Nous proposons ensuite un algorithme à base de Gradient conjugué noté ZBDC_{GC} . Cette méthode d'optimisation itérative a été largement utilisée dans divers domaines, dont les réseaux de neurones [120], le traitement d'antenne [32] et la SAS [257]. Elle nécessite le calcul exact du Gradient matriciel complexe. Dans un second temps, nous proposons un nouvel algorithme de ZBDC_{NU} basé sur une approche itérative de type LMA que l'on notera $\text{ZBDC}_{\text{LM}_{\text{HE}}}$. Cet algorithme repose sur le calcul exact des matrices Hessiennes - calcul que nous établirons -. Finalement, nous proposons une dernière solution basée sur le Gradient conjugué préconditionné baptisée $\text{ZBDC}_{\text{GCP}_{\text{HE}}}$. Nous rappelons que l'intérêt majeur de tous ces algorithmes réside dans le fait qu'ils sont plus généraux et qu'ils peuvent être appliqués à un ensemble quelconque de matrices.

Finalement, nous terminons ce chapitre sur des simulations numériques afin d'illustrer le comportement des approches proposées en BDC_{NU} et en ZBDC_{NU} et ce dans différentes situations (présence d'un bruit plus ou moins important, taille des ensembles à (zéro) bloc-diagonaliser conjointement, nombre de sources, etc...). Les matrices sont d'abord considérées comme exactement bloc-diagonales (respectivement zéro bloc-diagonales), puis elles seront ensuite progressivement perturbées par un bruit blanc additif de nature Gaussienne.

2.1 Bloc-diagonalisation conjointe d'un ensemble de matrices : problème et hypothèses

Nous rappelons dans ce paragraphe la formulation générale du problème de la bloc-diagonalisation conjointe. Nous commençons par en rappeler brièvement le principe avant d'introduire les fonctions de coût à optimiser qui ont été utilisées au niveau des différents algorithmes connus de BDC. Enfin, nous évoquons le lien existant entre BDC et DC.

2.1.1 Formulation du problème de la bloc-diagonalisation conjointe

Nous rappelons que le problème de la bloc-diagonalisation conjointe est formulé dans [108] de la manière suivante. Considérons un ensemble $\mathcal{M}$ de N_m ($N_m \in \mathbb{N}^*$) matrices carrées $\mathbf{M}_i \in \mathbb{C}^{M \times M}$ pour tout $i \in \{1, \dots, N_m\}$, admettant toutes la factorisation suivante :

$$\mathbf{M}_i = \mathbf{A} \mathbf{D}_i \mathbf{A}^H \quad (2.1)$$

où $(\cdot)^H$ est l'opérateur de transposition conjugaison et les N_m matrices carrées $\mathbf{D}_i$ de dimensions $N \times N$ sont des matrices bloc-diagonales *c.à.d* qu'elles possèdent toutes la forme suivante :

$$\mathbf{D}_i = \begin{pmatrix} \mathbf{D}_{i,11} & \mathbf{0}_{12} & \dots & \mathbf{0}_{1r} \\ \mathbf{0}_{21} & \mathbf{D}_{i,22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}_{r-1r} \\ \mathbf{0}_{r1} & \dots & \mathbf{0}_{rr-1} & \mathbf{D}_{i,rr} \end{pmatrix}, \text{ avec } r \text{ est le nombre de blocs considérés } (r \in \mathbb{N}^*), \text{ les}$$

matrices $\mathbf{D}_{i,jj}$, $i \in \{1, \dots, N_m\}$, $j \in \{1, \dots, r\}$ sont des matrices carrées de dimensions $n_j \times n_j$ avec $n_1 + \dots + n_r = N$ (dans l'optique d'une application en SAS, nous nous restreignons au cas où les matrices $\mathbf{D}_{i,jj}$ sont toutes de la même taille, impliquant en l'occurrence $N = r \times n_j$, $j \in \{1, \dots, r\}$), où $\mathbf{0}_{ij}$ désigne la matrice nulle de dimension $n_i \times n_j$. La matrice $\mathbf{A} \in \mathbb{C}^{M \times N}$ est une matrice de rang plein de dimension $M \times N$ ($M \geq N$) tandis que $\mathbf{A}^\dagger$ est la pseudo inverse de $\mathbf{A}$ (ou inverse de Moore-Penrose généralisée). C'est une matrice de dimension $N \times M$, notée $\mathbf{B}$ ($\mathbf{B} = \mathbf{A}^\dagger$). L'ensemble de N_m matrices carrées $\mathbf{D}_i \in \mathbb{C}^{N \times N}$ est noté par $\mathcal{D}$. La taille des blocs n_j pour tout $j = 1, \dots, r$ est supposée connue.

Précisons que dans un contexte de séparation aveugle de sources, $\mathbf{A}$ désigne la matrice de mélange alors que $\mathbf{B}$ désigne la matrice de séparation, N est le nombre de sources tandis que M est le nombre de capteurs. De nombreuses méthodes de SAS passent par une étape de décomposition conjointe de matrices afin d'estimer soit la matrice de mélange soit la matrice de séparation, même si d'autres approches ont également été considérées parmi lesquelles des méthodes directes ou par estimation successive des sources (déflation). En outre, ce problème de décomposition conjointe de matrices se révèle bien souvent être un problème de diagonalisation conjointe lorsque des mélanges instantanés de sources sont considérés ou de bloc-diagonalisation conjointe quand des mélanges convolutifs de sources sont considérés. Quand une contrainte d'unitarité sur les matrices $\mathbf{A}$ ou $\mathbf{B}$ recherchée est imposée, c'est qu'une étape préalable de blanchiment spatial des observations a été effectuée. Cependant, des études ont prouvé que dans le contexte de la SAS, cette étape bornait les performances des méthodes [42][80]. C'est la raison de l'engouement récent pour les méthodes s'affranchissant de cette contrainte.

2.1.2 Principe de la bloc-diagonalisation conjointe

Le problème général de la BDC non-unitaire consiste à estimer la matrice $\mathbf{A}$ et l'ensemble $\mathcal{D}$ des matrices bloc-diagonales, ce, uniquement, à partir des matrices de l'ensemble $\mathcal{M}$. Une méthode standard pour déterminer les matrices inconnues ($\mathbf{A}$ et $\{\mathbf{D}_i\}$) consiste à se ramener à un problème d'optimisation d'une fonction de coût judicieusement choisie. Un certain nombre de méthodes passent par la minimisation de la fonction de coût quadratique suivante :

$$C_{\text{BDC}}^{(1)}(\mathbf{A}, \{\mathbf{D}_i\}) = \sum_{i=1}^{N_m} \|\mathbf{M}_i - \mathbf{A} \mathbf{D}_i \mathbf{A}^H\|_F^2, \quad (2.2)$$

où $\|\cdot\|_F$ désigne la norme de Frobenius.

Néanmoins, il a été montré dans [110] que l'on peut procéder différemment et utiliser la fonction de coût quadratique suivante :

$$\mathcal{C}_{\text{BDC}}^{(2)}(\mathbf{B}, \{\widehat{\mathbf{D}}_i\}) = \sum_{i=1}^{N_m} \|\text{OffBdiag}_{(\mathbf{n})}\{\mathbf{B}\mathbf{M}_i\mathbf{B}^H\}\|_F^2 \stackrel{\text{def}}{=} \mathcal{C}_{\text{BDC}}(\mathbf{B}), \quad (2.3)$$

cela simplifie alors le calcul, puisqu'une seule matrice ($\mathbf{B}$) est finalement impliquée. Nous rappelons que les deux opérateurs matriciels $\text{Bdiag}_{(\mathbf{n})}\{\cdot\}$ et $\text{OffBdiag}_{(\mathbf{n})}\{\cdot\}$ sont respectivement définis par :

$$\text{Bdiag}_{(\mathbf{n})}\{\mathbf{M}\} = \begin{pmatrix} \mathbf{M}_{11} & \mathbf{0}_{12} & \dots & \mathbf{0}_{1r} \\ \mathbf{0}_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0}_{r1} & \mathbf{0}_{r2} & \dots & \mathbf{M}_{rr} \end{pmatrix}, \quad (2.4)$$

et

$$\text{OffBdiag}_{(\mathbf{n})}\{\mathbf{M}\} = \mathbf{M} - \text{Bdiag}_{(\mathbf{n})}\{\mathbf{M}\} = \begin{pmatrix} \mathbf{0}_{11} & \mathbf{M}_{12} & \dots & \mathbf{M}_{1r} \\ \mathbf{M}_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{M}_{r1} & \mathbf{M}_{r2} & \dots & \mathbf{0}_{rr} \end{pmatrix}, \quad (2.5)$$

avec $\mathbf{n} = (n_1, n_2, \dots, n_r)$, et la matrice carrée $\mathbf{M} = (M_{ij})$ ($N \times N$) possède la structure suivante :

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} & \dots & \mathbf{M}_{1r} \\ \mathbf{M}_{21} & \dots & \dots & \vdots \\ \vdots & \dots & \dots & \vdots \\ \mathbf{M}_{r1} & \mathbf{M}_{r2} & \dots & \mathbf{M}_{rr} \end{pmatrix}, \quad (2.6)$$

où les $\mathbf{M}_{ij}$ pour tous $i, j = 1, \dots, r$ sont des matrices ($n_i \times n_j$) (et $n_1 + \dots + n_r = N$).

Il est encore possible d'utiliser une autre fonction de coût, telle que celle définie dans [191] :

$$\mathcal{C}_{\text{BDC}}^{(3)}(\mathbf{B}) = \frac{1}{2} \sum_{i=1}^{N_m} \log \det \left(\text{Bdiag}_{(\mathbf{n})}\{\mathbf{B}\mathbf{M}_i\mathbf{B}^H\} \right) - \log \det (\mathbf{B}\mathbf{M}_i\mathbf{B}^H), \quad (2.7)$$

Succinctement, plusieurs solutions ont été proposées afin de résoudre le problème de la BDC. Nous trouvons celles basées sur l'optimisation de la fonction de coût définie au niveau de l'Eq. (2.2), tandis que d'autres sont basées sur l'optimisation de la fonction de coût définie au niveau de l'Eq. (2.3) [82][108]. Toutefois, L. De Lathauwer et al. dans [82] se sont restreints au cas d'une matrice $\mathbf{B}$ unitaire en proposant une méthode de type Jacobi. Ensuite, H. Ghennioui et al. ont considéré dans [108][110] le critère décrit au niveau de l'Eq. (2.3). Ce critère est réécrit dans le but d'effectuer une optimisation algébrique itérative impliquant une décomposition en valeurs singulières (DVS). Dans [113], la solution proposée par les auteurs est basée sur une approche de type Gradient. Or, cette méthode repose sur une approximation au niveau du calcul afin d'en déduire l'expression du Gradient matriciel complexe. Par la suite, le calcul exact a été effectué dans [110] et deux algorithmes de type gradient et gradient relatif en ont été déduits. Enfin, la fonction de coût $\mathcal{C}_{\text{BDC}}^{(3)}(\mathbf{B})$ a été introduite dans [191] afin de bloc-diagonaliser conjointement un ensemble de matrices définies positives et à symétrie Hermitienne sans contrainte d'unitarité sur la matrice recherchée.

2.1.3 Lien entre BDC et DC

Notons que lorsque les matrices $n_j \times n_j$ $\mathbf{D}_{i,jj}$ pour tout $j \in \{1, \dots, r\}$ et pour tout $i \in \{1, \dots, N_m\}$ se réduisent à des scalaires, ($n_j = 1$ pour tout $j \in \{1, \dots, r\}$ ce qui implique aussi que $r = N$), nous retombons alors sur un problème de diagonalisation conjointe non-unitaire d'un ensemble de matrices donné [50]. Les deux opérateurs matriciels susmentionnés se simplifient alors en $\text{OffDiag}\{\cdot\}$ et $\text{Diag}\{\cdot\}$:

$$\text{OffBdiag}_{(1)}\{\mathbf{M}\} = ((1 - \delta_{ij})M_{ij})\mathbf{1}_N = \text{OffDiag}\{\mathbf{M}\},$$

où $\mathbf{M} = (M_{ij})$, $\mathbf{1}_N$ est la matrice de dimension $(N \times N)$ dont toutes les composantes sont égales à 1 et $\delta_{ij} = 1$ si $i = j$ et 0 sinon et

$$\text{BDiag}_{(1)}\{\mathbf{M}\} = ((\delta_{ij})M_{ij})\mathbf{1}_N = \text{Diag}\{\mathbf{M}\} = \mathbf{M} - \text{OffDiag}\{\mathbf{M}\}.$$

Ce qui implique finalement de minimiser la fonction de coût suivante :

$$\mathcal{C}_D(\mathbf{B}) = \sum_{k=1}^{N_m} \|\text{OffDiag}\{\mathbf{B}\mathbf{M}_k\mathbf{B}^H\}\|_F^2. \quad (2.8)$$

Cette fonction a été introduite dans [43] et optimisée sous la contrainte d'une matrice $\mathbf{B}$ unitaire (*c.à.d* $M = N$ et $\mathbf{B}\mathbf{B}^H = \mathbf{B}^H\mathbf{B} = \mathbf{I}_N$, avec $\mathbf{I}_N$ une matrice identité de dimension $N \times N$). Les algorithmes proposés étaient fondés sur une généralisation de techniques de Jacobi [119]. Par la suite, cette fonction de coût a à nouveau été utilisée mais sans la contrainte de la recherche d'une matrice unitaire [252][190][92].

2.1.4 Quelques rappels sur l'optimisation et le préconditionnement

Dans la section précédente, nous avons rappelé que le problème de la BDC_{NU} posé dans de nombreuses applications de traitement du signal (traitement d'antennes, DOA, SAS, ...) pouvait être reformulé en un problème d'optimisation sans contrainte. Nous trouvons dans la littérature de nombreuses méthodes d'optimisation sans contrainte (méthodes de descente, Gradient, Gradient de la plus forte pente, Gradient relatif et Gradient conjugué) ont été étudiées et appliquées dans le contexte de la BDC [110][184]. Ces approches sont connues comme étant des méthodes du premier ordre, car elles ne dépendent que du Gradient de la fonction de coût et non pas de ses dérivées secondes (matrice Hessienne, notée $\mathbf{H}_\bullet$ sur laquelle nous allons revenir un peu plus tard dans ce manuscrit).

À notre connaissance, très peu d'études ont été menées sur le préconditionnement dans le contexte de la BDC. Vu notre application assez spécifique en télécommunications numériques et plus précisément en surveillance maritime ($\mathbf{S} - \mathbf{AIS}$), de nombreuses exigences doivent être prises en compte (vitesse de convergence, coût algorithmique (temps de calcul par itération), stabilité numérique, sensibilité/robustesse vis-à-vis du bruit et/ou des erreurs de modèle, problèmes d'initialisation, facilité de mise en œuvre en temps réel, etc.) l'objectif étant de choisir l'algorithme offrant le meilleur compromis parmi l'ensemble de ces exigences. C'est pourquoi nous avons choisi, ici, de nous intéresser plus particulièrement à l'utilisation d'un préconditionneur au niveau des algorithmes que nous développons afin de déterminer si le préconditionnement présente ou non un intérêt dans les problèmes de bloc-diagonalisation conjointe.

Généralement, le préconditionnement est une technique qui présente un réel intérêt si l'on se trouve dans l'un (ou plusieurs) des cas de figure suivants. Vouloir :

- permettre d’accélérer la convergence parce que les algorithmes du Gradient (notamment Gradient descendant, Gradient de la plus forte pente) sont réputés pour leur lenteur de convergence,
- améliorer la stabilité et la robustesse vis-à-vis d’erreurs (bruit ou modèle),
- parer au fait que les inconnues du problème peuvent avoir des amplitudes extrêmement différentes (la forme la plus basique du préconditionnement demeure la mise à l’échelle des variables au moyen de coefficients bien choisis (on utilise alors un préconditionneur diagonal)),
- faire face au fait que la fonction de coût peut évoluer très rapidement dans certaines directions et très lentement dans d’autres,
- être en mesure de résoudre des problèmes mal posés,
- assurer certaines contraintes telles que la non-négativité (voir par exemple, l’algorithme de Factorisation en Matrices Non négatives (NMF) proposé par Lee et Seung [152] basé sur une règle de mise à jour multiplicative qui à la base n’est rien d’autre qu’un algorithme de Gradient préconditionné).

Parfois le préconditionnement permet juste d’accélérer la vitesse de convergence mais il existe également certains cas/configurations difficiles qui ne peuvent être résolus sans un préconditionnement adéquat. Plusieurs travaux ont traité de ce problème de la détermination de préconditionneurs efficaces. En premier lieu, le préconditionnement a été utilisé dans des méthodes directes. Par la suite, ce concept a été étendu dans [232][88] au cas des méthodes itératives (accélération de Chebyshev). La question principale qui se pose reste bien évidemment la manière de choisir le préconditionneur que nous noterons dans la suite du manuscrit $\mathbf{P}$. Dans la littérature, nous pouvons trouver de nombreux ouvrages sur les différents types de préconditionneurs, parmi lesquels les préconditionneurs de Jacobi et de Gauss-Seidel [241]. Le premier consiste à utiliser la diagonale de la matrice recherchée tandis que le second repose sur la décomposition de la matrice de préconditionnement sous la forme d’un produit d’une matrice triangulaire inférieure, une matrice diagonale et une matrice triangulaire supérieure. Cependant, pour des problèmes non-symétriques, ces préconditionneurs restent limités en terme d’efficacité et de robustesse [19]. L’analyse de la vitesse de convergence de la méthode Gradient descendant préconditionné fournit une aide substantielle dans la manière de choisir un préconditionneur $\mathbf{P}$ adapté de même que les grandes lignes à suivre. Si la fonction de coût est dérivable deux fois dans le voisinage d’un minimiseur local, *i.e.* avec une matrice Hessienne définie comme :

$$\mathbf{H}_\bullet \stackrel{\text{def}}{=} \nabla_{\mathbf{B}}^2 \mathcal{C}_{\text{BDC}} \stackrel{\text{def}}{=} \frac{\partial^2 \mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B} \partial \mathbf{B}^*}, \quad (2.9)$$

la convergence la plus rapide s’obtient quand le préconditionneur $\mathbf{P}$ minimise le conditionnement noté κ du produit $\mathbf{P}\mathbf{H}_\bullet$.³⁰ Pour des fonctions de coût quadratiques, le préconditionneur idéal devrait donc être $\mathbf{P} = (\mathbf{H}_\bullet)^{-1}$ de sorte que $\mathbf{P}\mathbf{H}_\bullet = \mathbf{I}$ étant donné que la matrice identité $\mathbf{I}$ possède le conditionnement minimal (càd un conditionnement de 1). Néanmoins, il n’est pas toujours évident/facile de calculer $(\mathbf{H}_\bullet)^{-1}$ (surtout lorsque N et M tendent à devenir grands), c’est pourquoi, il peut être intéressant d’utiliser des préconditionneurs qui vont approximer $(\mathbf{H}_\bullet)^{-1}$ et qui restent peu coûteux à calculer, construire et appliquer [27][54].

Pour être en mesure de concevoir une matrice de préconditionnement $\mathbf{P}$ adéquate dans le contexte de la BDC, nous commençons donc par le calcul exact des quatres matrices Hessiennes complexes. Ensuite, nous serons en mesure d’en déduire quatre nouveaux algorithmes fondés respectivement sur :

30. Le conditionnement κ est le rapport entre la plus grande valeur singulière d’une matrice donnée, ici $\mathbf{P}\mathbf{H}_\bullet$, divisée par la plus petite valeur singulière de cette matrice, *c.à.d* $\kappa \stackrel{\text{def}}{=} \frac{\lambda_{\max}(\mathbf{P}\mathbf{H}_\bullet)}{\lambda_{\min}(\mathbf{P}\mathbf{H}_\bullet)}$.

- le gradient conjugué (dans ce cas, le préconditionneur est tout simplement la matrice identité),
- le gradient conjugué préconditionné,
- l’algorithme de Levenberg-Marquardt,
- les méthodes Quasi-Newton.

Rappelons encore que la méthode du Gradient conjugué est une méthode d’optimisation itérative très utilisée en optimisation non linéaire [186][122] et son efficacité peut parfois être améliorée grâce au choix d’une bonne matrice de préconditionnement [19].

2.1.5 Quatre nouveaux algorithmes de bloc-diagonalisation conjointe non unitaire

Pour pouvoir estimer la matrice complexe inconnue $\mathbf{B} \in \mathbb{C}^{N \times M}$, la fonction de coût donnée au niveau de l’Eq. (2.3) doit être minimisée. Dans le but d’utiliser les algorithmes d’optimisation itératifs déterministes classiques, il nous faut donc évaluer des grandeurs telles que le Gradient matriciel complexe voire les matrices Hessiennes complexes. En ce qui concerne le gradient matriciel, nous devons calculer la différentielle d’ordre un, notée $d\mathcal{C}_{\text{BDC}}(\mathbf{B})$ de la fonction de coût $\mathcal{C}_{\text{BDC}}(\mathbf{B})$, ce qui passe par le calcul des dérivées partielles $\frac{\partial}{\partial \mathbf{B}}$ de la fonction de coût quadratique $\mathcal{C}_{\text{BDC}}(\mathbf{B})$ par rapport à $\mathbf{B}$ et $\mathbf{B}^*$. C’est pourquoi, à partir de maintenant, la fonction de coût sera notée $\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)$ et non plus $\mathcal{C}_{\text{BDC}}(\mathbf{B})$ sachant que $\mathbf{B}^*$ désigne la matrice complexe conjuguée de la matrice $\mathbf{B}$. Pour les matrices Hessiennes complexes, il nous faudra évaluer la différentielle d’ordre deux, notée $d^2\mathcal{C}_{\text{BDC}}(\mathbf{B})$.

2.1.5.1 Gradient matriciel complexe

La matrice $N \times M$ du Gradient matriciel complexe de notre fonction de coût scalaire à valeur réelle donnée au niveau de l’Eq. (2.3), sera notée $\nabla_{\mathbf{B}}\mathcal{C}_{\text{BDC}}$. Elle est définie dans [129] (p. 63 et p. 77) par :

$$\nabla_{\mathbf{B}}\mathcal{C}_{\text{BDC}} \stackrel{\text{def}}{=} \frac{\partial \mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}^*}, \quad (2.10)$$

Ce calcul a été effectué dans [110] où il est démontré que cette matrice vaut :

$$\frac{\partial \mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}^*} = \sum_{i=1}^{N_m} \left[\text{OffBdiag}_{(\mathbf{n})}\{\mathbf{B}\mathbf{M}_i\mathbf{B}^H\}\mathbf{B}\mathbf{M}_i^H + \left(\text{OffBdiag}_{(\mathbf{n})}\{\mathbf{B}\mathbf{M}_i\mathbf{B}^H\} \right)^H \mathbf{B}\mathbf{M}_i \right]. \quad (2.11)$$

Nous notons que dans le cas de matrices $\mathbf{M}_i$ à symétrie hermitienne, le gradient matriciel se simplifie en :

$$\frac{\partial \mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}^*} = 2 \sum_{i=1}^{N_m} \left[\text{OffBdiag}_{(\mathbf{n})}\{\mathbf{B}\mathbf{M}_i\mathbf{B}^H\}\mathbf{B}\mathbf{M}_i^H \right]. \quad (2.12)$$

2.1.5.2 Matrices Hessiennes complexes

Soit la matrice complexe $\mathbf{H}$, matrice carrée de dimensions $2NM \times 2NM$ constituée de quatre matrices blocs carrées de dimension $NM \times NM$ encore appelées les matrices Hessiennes :

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{\mathbf{B}, \mathbf{B}^*} & \mathbf{H}_{\mathbf{B}^*, \mathbf{B}^*} \\ \mathbf{H}_{\mathbf{B}, \mathbf{B}} & \mathbf{H}_{\mathbf{B}^*, \mathbf{B}} \end{bmatrix}. \quad (2.13)$$

Pour calculer la matrice $\mathbf{H}$ que nous venons de mentionner, nous suivons là encore le même raisonnement que celui tenu dans [129] (les détails de ce calcul se trouvent au niveau de l'Annexe 2 de ce manuscrit). Pour résumer, nous devons obtenir l'expression de la différentielle du second second ordre de $\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)$ en l'écrivant sous la forme suivante :

$$d^2\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*) = \begin{bmatrix} d\text{vec}^T(\mathbf{B}^*) & d\text{vec}^T(\mathbf{B}) \end{bmatrix} \begin{bmatrix} \mathbf{A}_{00} & \mathbf{A}_{01} \\ \mathbf{A}_{10} & \mathbf{A}_{11} \end{bmatrix} \begin{bmatrix} d\text{vec}(\mathbf{B}) \\ d\text{vec}(\mathbf{B}^*) \end{bmatrix}. \quad (2.14)$$

Comme cela est démontré dans [129], les quatre matrices complexes $\mathbf{A}_{00}$, $\mathbf{A}_{01}$, $\mathbf{A}_{10}$ et $\mathbf{A}_{11}$, de dimensions $NM \times NM$, impliquées dans l'expression précédente sont reliées aux matrices Hessiennes par les équations suivantes :

$$\mathbf{H}_{\mathbf{B}, \mathbf{B}^*}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)) = \frac{\mathbf{A}_{00} + \mathbf{A}_{11}^T}{2} = (\mathbf{H}_{\mathbf{B}^*, \mathbf{B}}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)))^T, \quad (2.15)$$

$$\mathbf{H}_{\mathbf{B}^*, \mathbf{B}^*}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)) = \frac{\mathbf{A}_{01} + \mathbf{A}_{01}^T}{2}, \quad (2.16)$$

$$\mathbf{H}_{\mathbf{B}, \mathbf{B}}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)) = \frac{\mathbf{A}_{10} + \mathbf{A}_{10}^T}{2}. \quad (2.17)$$

Nous avons donc établi qu'elles valaient (se reporter à l'Annexe 2 pour le détail des calculs) :

$$\begin{aligned} \mathbf{A}_{00} &= (\mathbf{M}_i^T \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{\text{Boff}}^T (\mathbf{B}^* \mathbf{M}_i^* \otimes \mathbf{I}_N) + (\mathbf{M}_i^* \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{\text{Boff}}^T (\mathbf{B}^* \mathbf{M}_i^T \otimes \mathbf{I}_N) \\ &\quad + \mathbf{M}_i^* \otimes \text{OffBdiag}_{(n)}\{\mathbf{B} \mathbf{M}_i \mathbf{B}^H\} + \mathbf{M}_i^T \otimes \text{OffBdiag}_{(n)}\{\mathbf{B} \mathbf{M}_i^H \mathbf{B}^H\} \\ &= \mathbf{A}_{11}^*, \end{aligned} \quad (2.18)$$

$$\begin{aligned} \mathbf{A}_{10} &= \mathbf{K}_{N,M}^T (\mathbf{I}_N \otimes \mathbf{M}_i \mathbf{B}^H) \mathbf{T}_{\text{Boff}}^T (\mathbf{B}^* \mathbf{M}_i^* \otimes \mathbf{I}_N) \\ &\quad + \mathbf{K}_{N,M}^T (\mathbf{I}_N \otimes \mathbf{M}_i^H \mathbf{B}^H) \mathbf{T}_{\text{Boff}}^T (\mathbf{B}^* \mathbf{M}_i^T \otimes \mathbf{I}_N) = \mathbf{A}_{01}^*, \end{aligned} \quad (2.19)$$

où l'opérateur $\otimes$ désigne le produit de Kronecker, $\mathbf{K}_{N,M}$ est une matrice carrée de commutation de taille $NM \times NM$ et $\mathbf{T}_{\text{Boff}} = \mathbf{I}_{N^2} - \mathbf{T}_{\text{Diag}}$, est la matrice de dimensions $N^2 \times N^2$ dite de "transformation" avec $\mathbf{T}_{\text{Diag}} = \text{diag}\{\text{vec}\{\mathbf{B} \text{diag}\{\mathbf{1}_N\}\}\}$, $\mathbf{1}_N$ est la matrice $N \times N$ dont tous les éléments sont égaux à un et $\text{diag}\{\mathbf{a}\}$ est une matrice diagonale carrée qui contient les éléments du vecteur $\mathbf{a}$ sur sa diagonale, $\mathbf{I}_{N^2} = \text{Diag}\{\mathbf{1}_{N^2}\}$ est la matrice identité de taille $N^2 \times N^2$ et $\text{Diag}\{\mathbf{A}\}$ est la matrice diagonale carrée qui possède les mêmes éléments diagonaux que la matrice $\mathbf{A}$.

Nous pouvons vérifier que la matrice complexe $\mathbf{H}$ définie au niveau de l'Eq. (2.13) est Hermitienne et que deux parties peuvent être distinguées au niveau de cette matrice $\mathbf{H}$: i) sa partie bloc-diagonale ii) et sa partie hors bloc-diagonale. Selon [209], il est suggéré de ne considérer que la partie diagonale (partie réelle) de la matrice Hessienne (nous parlons des Eqs. (2.15) et (2.18)) ce qui permet de facilement calculer son inverse, et d'assurer la non croissance de la fonction de coût d'autre part. Nous suggérons donc d'utiliser comme matrice de préconditionnement $\mathbf{P}$, non pas l'inverse de la matrice Hessienne mais une approximation de dimension $NM \times NM$ de celle-ci dont l'expression est donnée ci-dessous :

$$\mathbf{P} = (\hat{\mathbf{H}})^{-1}, \quad (2.20)$$

$$\hat{\mathbf{H}} = \text{Diag}\{\mathbf{H}_{\mathbf{B}, \mathbf{B}^*}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*))\} = \text{Diag}\{\mathbf{H}_{\mathbf{B}^*, \mathbf{B}}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*))\}. \quad (2.21)$$

L'objectif est maintenant d'en déduire différents algorithmes (avec préconditionnement) et de parvenir à évaluer leur comportement.

2.1.6 Schéma d'optimisation déterministes itératifs classiques

Nous rappelons que notre objectif est l'estimation de la matrice complexe $\mathbf{B} \in \mathbb{C}^{N \times M}$, ce qui est réalisé au moyen de la minimisation de la fonction de coût donnée au niveau de l'Eq. (2.3). Différents schémas d'optimisation itératifs déterministes peuvent être considérés.

Partant d'une matrice $\mathbf{B}$ initiale, notée $\mathbf{B}^{(0)}$, à chaque itération m , nous devons ré-estimer cette matrice que nous notons alors $\mathbf{B}^{(m)}$. Cette matrice peut bien évidemment aussi être stockée dans un vecteur $\mathbf{b}^{(m)} = \text{vec}\{\mathbf{B}^{(m)}\}$, moyennant l'application de l'opérateur $\text{vec}\{\cdot\}$. Selon le schéma d'optimisation considéré, $\mathbf{B}^{(m)}$ pour tout m sera mise à jour selon l'une des règles d'adaptation suivante. Ainsi, quand l'approche de la plus forte pente est considérée (approche du Gradient si le pas est un scalaire positif fixe *i.e.* $\mu^{(m)} = \mu$ pour tout m) $\mathbf{B}$ est alors mise à jour à chaque itération m ($m = 1, 2, \dots$) selon la règle suivante :

$$\mathbf{B}^{(m+1)} = \mathbf{B}^{(m)} - \mu^{(m)} \nabla_{\mathbf{B}} \mathcal{C}_{\text{BDC}}(\mathbf{B}^{(m)}, \mathbf{B}^{*(m)}), \quad (2.22)$$

où de façon équivalente selon :

$$\mathbf{b}^{(m+1)} = \mathbf{b}^{(m)} - \mu^{(m)} \mathbf{g}_{\mathbf{B}^*}^{(m)}, \quad (2.23)$$

avec $\nabla_{\mathbf{B}} \mathcal{C}_{\text{BDC}}(\mathbf{B}^{(m)}, \mathbf{B}^{*(m)})$ est le gradient matriciel donné par l'Eq. (2.11) ou par sa forme de vectorielle $\mathbf{g}_{\mathbf{B}^*}^{(m)}$ *i.e.* $\mathbf{g}_{\mathbf{B}^*}^{(m)} = \left(\mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{BDC}}(\mathbf{B}^{(m)}, \mathbf{B}^{*(m)})) \right)^T$ si $\mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{BDC}}(\mathbf{B}^{(m)}, \mathbf{B}^{*(m)})) = \text{vec}^T \left(\frac{\partial \mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}^*} \right)$ et $\mu^{(m)}$ est une constante réelle positive appelée le pas (le problème du choix de ce pas est traité dans la section §2.1.6.4).

Cependant, un tel algorithme peut converger très lentement, c'est la raison pour laquelle nous allons maintenant nous intéresser à d'autres algorithmes parmi lesquels le gradient conjugué et les algorithmes utilisant un préconditionneur destiné à accélérer la vitesse de convergence.

2.1.6.1 Algorithmes du gradient conjugué non linéaire et du gradient conjugué préconditionné

La méthode du gradient conjugué modifie les directions de recherche afin d'assurer le fait qu'elles soient approximativement mutuellement conjuguées (*i.e.* orthogonales) par rapport à un produit scalaire lié à la matrice Hessienne de la fonction de coût. La direction de recherche $\mathbf{d}$ est alors donnée à la première itération par l'expression suivante :

$$\mathbf{d}^{(1)} = -\mathbf{P}^{(1)} \mathbf{g}_{\mathbf{B}^*}^{(1)}. \quad (2.24)$$

Ensuite, $\forall m = 1, 2, \dots$, la règle d'adaptation utilisée est :

$$\begin{cases} \mathbf{b}^{(m+1)} = \mathbf{b}^{(m)} + \mu^{(m)} \mathbf{d}^{(m)}, \\ \mathbf{d}^{(m+1)} = -\mathbf{P}^{(m+1)} \mathbf{g}_{\mathbf{B}^*}^{(m+1)} + \beta^{(m)} \mathbf{d}^{(m)}, \end{cases} \quad (2.25)$$

où l'expression du préconditionneur $\mathbf{P}$ est donnée au niveau des Eqs. (2.20)-(2.21). Ce dernier devrait être une matrice définie positive. C'est pourquoi les éléments diagonaux non positifs sont exclus : dans le cas où ce problème est rencontré, nous choisissons de ne plus préconditionner l'algorithme ($\mathbf{P}$ est alors remplacé par la matrice identité $\mathbf{I}_{NM}$ et nous revenons alors à l'algorithme classique du gradient conjugué non linéaire - notons que la version non préconditionnée de l'algorithme du gradient conjugué sera notée BDC_{GC}). En optimisation mathématique, la méthode de la recherche linéaire (exact line search en anglais) est l'une des deux approches (avec la

méthode des régions de confiance) pour forcer la convergence des algorithmes. Dans ce contexte, plusieurs expressions de β ont donc été établies et utilisées : i) la formule de Fletcher-Reeves (β_{FR}), ii) la formule de Polak-Ribière (β_{PR}) ii) et celle de Dai-Yuan (β_{DY}) [209][195][74] :

$$\beta_{\text{FR}}^{(m+1)} = \frac{(\mathbf{g}_{\mathbf{B}^*}^{(m+1)})^H \mathbf{P}^{(m+1)} \mathbf{g}_{\mathbf{B}^*}^{(m+1)}}{(\mathbf{g}_{\mathbf{B}^*}^{(m)})^H \mathbf{P}^{(m)} \mathbf{g}_{\mathbf{B}^*}^{(m)}}, \quad (2.26)$$

$$\beta_{\text{PR}}^{(m+1)} = \frac{(\mathbf{g}_{\mathbf{B}^*}^{(m+1)} - \mathbf{g}_{\mathbf{B}^*}^{(m)})^H \mathbf{P}^{(m+1)} \mathbf{g}_{\mathbf{B}^*}^{(m+1)}}{(\mathbf{g}_{\mathbf{B}^*}^{(m)})^H \mathbf{P}^{(m)} \mathbf{g}_{\mathbf{B}^*}^{(m)}}, \quad (2.27)$$

$$\beta_{\text{DY}}^{(m+1)} = \frac{(\mathbf{g}_{\mathbf{B}^*}^{(m+1)})^H \mathbf{P}^{(m+1)} \mathbf{g}_{\mathbf{B}^*}^{(m+1)}}{(\mathbf{d}^{(m)})^H \mathbf{P}^{(m)} (\mathbf{g}_{\mathbf{B}^*}^{(m+1)} - \mathbf{g}_{\mathbf{B}^*}^{(m)})}. \quad (2.28)$$

Les algorithmes qui en résultent seront alors notés $\text{BDC}_{\text{GCP}_{\text{HE}}}$ où HE signifie Hessienne exacte. Enfin, notons que comme cela a été remarqué dans [195], une réinitialisation ou “restart” de l’algorithme peut jouer un rôle important au niveau de la convergence, car en général, on ne peut garantir que les directions de recherche $\mathbf{d}$ soient effectivement des directions de descente. Toutefois, une étape de descente la plus rapide (steepest descent) est effectuée à chaque restart en imposant $\mathbf{d} = -\mathbf{g}_{\mathbf{B}^*}$. Un restart peut aussi conduire à des meilleures performances car l’algorithme converge plus vite. Dans notre cas ce “restart” est réalisé tous les N_0 iterations mais on peut aussi le forcer si $|\mathbf{g}_{\mathbf{B}^*}|$ vient à devenir trop grand.

2.1.6.2 Algorithme de Levenberg-Marquardt

Lorsque la matrice Hessienne $\mathbf{H}$ ou sa version approchée $\hat{\mathbf{H}}$ tend à perdre sa propriété de définie-positivité au cours des itérations, il arrive que l’on ne parvienne plus à construire les directions de descente. C’est la raison pour laquelle, une autre manière de stabiliser les algorithmes d’optimisation consiste à utiliser des techniques dites à région de confiance qui modifient alors $\hat{\mathbf{H}}$ en lui ajoutant un multiple de la matrice d’identité. C’est le principe de l’approche de Levenberg-Marquardt [160] qui suit la règle d’adaptation suivante :

$$\mathbf{b}^{(m+1)} = \mathbf{b}^{(m)} - \mu^{(m)} \left(\hat{\mathbf{H}}^{(m)} + \alpha^{(m)} \mathbf{I}_{NM} \right)^{-1} \mathbf{g}_{\mathbf{B}^*}^{(m)}, \quad (2.29)$$

avec $\alpha^{(m)}$ est un coefficient de relaxation ou d’amortissement. Le choix de ce facteur n’est souvent pas évident. En fonction de $\mathbf{b}$, le facteur d’amortissement est remplacé à chaque itération m soit par $\alpha^{(m+1)} = \alpha^{(m)}/\gamma$ ou $\alpha^{(m+1)} = \alpha^{(m)} \times \gamma$. D. W. Marquardt a recommandé dans [167] de commencer par une valeur $\alpha = \alpha^{(0)} = 10^{-2}$ et un paramètre $\gamma > 1$ (un panorama sur le choix de α et γ est donné dans [250][104]). Dans la même optique, pour un choix plus optimal de α , une autre mise à jour de ce facteur a été proposée dans [183] que nous exploitons dans cet algorithme (pour plus de détails sur la procédure de mise à jour utilisée, le lecteur peut consulter le résumé des algorithmes de BDC_{NU} dans §2.1.7).

Nous remarquons que si $\hat{\mathbf{H}} = \mathbf{I}_{NM}$ dans l’Eq. (2.29) (ou encore en considérant α comme très grand), nous retombons alors sur l’algorithme du gradient de l’Eq. (2.23). Cet algorithme peut également être considéré comme une variante de l’algorithme du gradient conjugué préconditionné donné au niveau de l’Eq. (2.25) pour une valeur de $\beta = 0$, et dans lequel le préconditionneur n’est pas directement l’inverse de la Hessienne, mais une version légèrement modifiée. En ce qui concerne $\hat{\mathbf{H}}$, nous pouvons utiliser notre solution proposée au niveau des Eqs. (2.20)-(2.21) ou la version approchée de la matrice Hessienne suivante (également utilisée dans les méthodes quasi-Newton notamment les algorithmes de Broyden-Fletcher-Goldfarb-Shanno (BFGS) et Davidon-Fletcher-Powell (DFP)) :

$$\begin{cases} \hat{\mathbf{H}}^{(m+1)} &= \hat{\mathbf{H}}^{(m)} + \frac{\Delta \mathbf{g}_{\mathbf{B}^*}^{(m)} (\Delta \mathbf{g}_{\mathbf{B}^*}^{(m)})^T}{\langle \Delta \mathbf{g}_{\mathbf{B}^*}^{(m)}, \Delta \mathbf{b}^{(m)} \rangle} - \frac{(\hat{\mathbf{H}}^{(m)} \Delta \mathbf{b}^{(m)}) (\hat{\mathbf{H}}^{(m)} \Delta \mathbf{b}^{(m)})^T}{\langle \hat{\mathbf{H}}^{(m)} \Delta \mathbf{b}^{(m)}, \Delta \mathbf{b}^{(m)} \rangle}, \\ \Delta \mathbf{b}^{(m)} &= \mathbf{b}^{(m+1)} - \mathbf{b}^{(m)}, \\ \Delta \mathbf{g}_{\mathbf{B}^*}^{(m)} &= \mathbf{g}_{\mathbf{B}^*}^{(m+1)} - \mathbf{g}_{\mathbf{B}^*}^{(m)}, \end{cases} \quad (2.30)$$

où $\langle \mathbf{u}, \mathbf{v} \rangle$ désigne le produit scalaire *i.e.* $\mathbf{u}^H \mathbf{v}$. Les algorithmes résultant sont notés $\text{BDC}_{\text{LM}_{\text{HE}}}$ ou $\text{BDC}_{\text{LM}_{\text{HA}}}$ (HE signifiant Hessienne Exacte et HA signifiant Hessienne Approchée).

2.1.6.3 Méthodes Quasi-Newton (BFGS et DFP)

Dans l'Eq. (2.29), si nous posons $\alpha = 0$ et nous considérons que le préconditionneur $\mathbf{P}$ est une matrice de dimensions $NM \times NM$ approchée de l'inverse de la matrice Hessienne donnée par l'Eq. (2.30), nous obtenons la règle d'adaptation de Broyden-Fletcher-Goldfarb-Shanno (BFGS). En utilisant le lemme d'inversion et en écrivant la quantité suivante $\rho = \frac{1}{(\Delta \mathbf{g}_{\mathbf{B}^*}^{(m)})^T \Delta \mathbf{b}^{(m)}}$, nous pouvons alors estimer l'inverse de l'approchée de la matrice Hessienne. Ainsi, l'algorithme (2.30) ci-dessus peut être réécrit comme :

$$\begin{cases} \mathbf{b}^{(m+1)} &= \mathbf{b}^{(m)} - \mu^{(m)} (\hat{\mathbf{H}}^{(m)})^{-1} \mathbf{g}_{\mathbf{B}^*}^{(m)}, \\ \Delta \mathbf{b}^{(m)} &= \mathbf{b}^{(m+1)} - \mathbf{b}^{(m)}, \\ \Delta \mathbf{g}_{\mathbf{B}^*}^{(m)} &= \mathbf{g}_{\mathbf{B}^*}^{(m+1)} - \mathbf{g}_{\mathbf{B}^*}^{(m)}, \\ (\hat{\mathbf{H}}^{(m+1)})^{-1} &= (\hat{\mathbf{H}}^{(m)})^{-1} + \rho [1 + \rho (\Delta \mathbf{g}^{(m)})^T (\hat{\mathbf{g}}^{(m)})^{-1} \Delta \mathbf{g}^{(m)}] \Delta \mathbf{b}^{(m)} (\Delta \mathbf{b}^{(m)})^T \\ &\quad - \rho \Delta \mathbf{b}^{(m)} (\Delta \mathbf{g}_{\mathbf{B}^*}^{(m)})^T (\hat{\mathbf{H}}^{(m)})^{-1} - \rho (\hat{\mathbf{H}}^{(m)})^{-1} \Delta \mathbf{g}_{\mathbf{B}^*}^{(m)} (\Delta \mathbf{b}^{(m)})^T. \end{cases} \quad (2.31)$$

En considérant que le préconditionneur $\mathbf{P}$ est directement l'approximation $NM \times NM$ de l'inverse de la matrice Hessienne donnée dans l'Eq. (2.32) :

$$\mathbf{P}^{(m+1)} = \mathbf{P}^{(m)} + \frac{\Delta \mathbf{b}^{(m)} (\Delta \mathbf{b}^{(m)})^T}{\langle \Delta \mathbf{g}_{\mathbf{B}^*}^{(m)}, \Delta \mathbf{b}^{(m)} \rangle} - \frac{(\mathbf{P}^{(m)} \Delta \mathbf{g}_{\mathbf{B}^*}^{(m)}) (\mathbf{P}^{(m)} \Delta \mathbf{g}_{\mathbf{B}^*}^{(m)})^T}{\langle \Delta \mathbf{g}_{\mathbf{B}^*}^{(m)}, \mathbf{P}^{(m)} \Delta \mathbf{g}_{\mathbf{B}^*}^{(m)} \rangle}, \quad (2.32)$$

nous obtenons alors la règle d'adaptation suivante comme dans l'algorithme Davidon-Fletcher-Powell (DFP) :

$$\begin{cases} \mathbf{b}^{(m+1)} &= \mathbf{b}^{(m)} - \mu^{(m)} \mathbf{P}^{(m)} \mathbf{g}_{\mathbf{B}^*}^{(m)}, \\ \Delta \mathbf{b}^{(m)} &= \mathbf{b}^{(m+1)} - \mathbf{b}^{(m)}, \\ \Delta \mathbf{g}_{\mathbf{B}^*}^{(m)} &= \mathbf{g}_{\mathbf{B}^*}^{(m+1)} - \mathbf{g}_{\mathbf{B}^*}^{(m)}, \\ \mathbf{P}^{(m+1)} &= \mathbf{P}^{(m)} + \frac{\Delta \mathbf{b}^{(m)} (\Delta \mathbf{b}^{(m)})^T}{\langle \Delta \mathbf{g}_{\mathbf{B}^*}^{(m)}, \Delta \mathbf{b}^{(m)} \rangle} - \frac{(\mathbf{P}^{(m)} \Delta \mathbf{g}_{\mathbf{B}^*}^{(m)}) (\mathbf{P}^{(m)} \Delta \mathbf{g}_{\mathbf{B}^*}^{(m)})^T}{\langle \Delta \mathbf{g}_{\mathbf{B}^*}^{(m)}, \mathbf{P}^{(m)} \Delta \mathbf{g}_{\mathbf{B}^*}^{(m)} \rangle}. \end{cases} \quad (2.33)$$

Dans tous les cas, cet algorithme est initialisé en utilisant une matrice $\mathbf{P}^{(1)}$ (ou $\hat{\mathbf{H}}^{(1)}$), de dimensions $NM \times NM$, symétrique et définie positive. Les algorithmes résultants sont notés $\text{BDC}_{\text{BFGS}_{\text{HA}}}$ ou $\text{BDC}_{\text{DFP}_{\text{HIA}}}$ (HIA désigne Hessienne Inverse Approchée). Lorsque l'expression donnée par l'Eq. (2.20) est utilisée à la place de (2.30) ou (2.32), l'algorithme résultant est noté $\text{BDC}_{\text{QN}_{\text{HE}}}$ (QN signifiant Quasi-Newton).

Nous signalons que l'algorithme de Levenberg-Marquardt a déjà fait l'objet d'une publication [55]. Par ailleurs, les autres algorithmes de BDC sont présentés dans un article soumis à la revue de Signal Processing [62].

2.1.6.4 Recherche du pas optimal

Il reste essentiel de trouver un pas μ dans la direction de recherche considérée $\mathbf{d}^{(m)}$ à chaque itération m . Ce pas $\mu^{(m)}$ peut être trouvé par différents moyens : i) fixe ou diminué avec les itérations, ii) ou nous pouvons opter pour une recherche linéaire globale avec ELS (Enhanced Line Search), iii) ou par une recherche linéaire approchée de type “backtracking” qui est une méthode localement optimale [37]. Ce paragraphe est dédié au calcul du pas optimal noté μ_{opt} . Ainsi lorsque nous considérons par exemple l’algorithme du gradient conjugué préconditionné, nous devons effectuer le calcul algébrique de la quantité suivante : $\mathcal{C}_{\text{BDC}}(\mathbf{B}^{(m+1)}) = \mathcal{C}_{\text{BDC}}(\mathbf{B}^{(m)} - \mu \mathbf{D}^{(m)})$, avec $\mathbf{D}^{(m)} = \text{unvec}(\mathbf{d}^{(m)})$ ce qui est aussi équivalent à $\mathbf{d}^{(m)} = \text{vec}(\mathbf{D}^{(m)})$ où le vecteur $\mathbf{d}^{(m)}$ est donné par l’Eq. (2.25) (et l’opérateur unvec permet de repasser de la forme vectorielle à la forme matricielle). Cette quantité est un polynôme d’ordre 4 dont l’expression est donnée par (pour simplifier les différentes expressions, nous choisissons d’omettre la dépendance vis-à-vis des itérations m) :

$$\mathcal{C}_{\text{BDC}}(\mathbf{B} - \mu \mathbf{D}) = a_0 + a_1\mu + a_2\mu^2 + a_3\mu^3 + a_4\mu^4, \quad (2.34)$$

où les cinq coefficients a_0, a_1, a_2, a_3 et a_4 valent (voir l’Annexe 2 pour plus de détails sur les calculs intermédiaires),

$$a_0 = \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_0^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_0 \} \right\}, \quad (2.35)$$

$$a_1 = - \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_1^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_0 \} + \mathbf{C}_0^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_1 \} \right\}, \quad (2.36)$$

$$a_2 = \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_2^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_0 \} + \mathbf{C}_1^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_1 \} + \mathbf{C}_0^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_2 \} \right\}, \quad (2.37)$$

$$a_3 = - \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_2^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_1 \} + \mathbf{C}_1^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_2 \} \right\}, \quad (2.38)$$

$$a_4 = \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_2^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_2 \} \right\}, \quad (2.39)$$

où

$$\mathbf{C}_0 = \mathbf{B} \mathbf{M}_i \mathbf{B}^H, \quad (2.40)$$

$$\mathbf{C}_1 = \mathbf{B} \mathbf{M}_i \mathbf{D}^H + \mathbf{D} \mathbf{M}_i \mathbf{B}^H, \quad (2.41)$$

$$\mathbf{C}_2 = \mathbf{D} \mathbf{M}_i \mathbf{D}^H. \quad (2.42)$$

Ensuite, on en déduit l’expression du polynôme d’ordre trois de l’Eq. (2.43) qui correspond à la dérivée par rapport à μ du polynôme de degré quatre donné par l’Eq. (2.34).

$$\frac{\partial \mathcal{C}_{\text{BDC}}(\mathbf{B} - \mu \mathbf{D})}{\partial \mu} = 4a_4\mu^3 + 3a_3\mu^2 + 2a_2\mu + a_1. \quad (2.43)$$

Ses trois racines peuvent alors être obtenues soit algébriquement soit numériquement. Le pas optimal μ_{opt} correspond à la racine qui conduit à la plus petite valeur de l’Eq. (2.34).

2.1.7 Algorithmes de BDC_{NU} en résumé

Les algorithmes de BDC_{NU} (notés par BDC_{·HE} ou BDC_{·H(l)A}) sont résumés au niveau du Tableau 2.1 :

Data : N_m matrices carrées $\mathbf{M}_1, \mathbf{M}_2, \dots, \mathbf{M}_{N_m}$, critère d'arrêt ϵ , pas μ , nb d'itérations max N_{max} , valeur utilisée pour le restart N_0

Result : Estimation de la matrice $\mathbf{B}$

initialiser : $\mathbf{B}^{(0)}$; $\mathbf{D}^{(0)}$; $m = 0$; facteur d'amortissement $\alpha^{(0)} = 10^{-2} \max \{ \text{Diag} \{ \mathbf{H}_{\mathbf{B}, \mathbf{B}^*} (\mathcal{C}_{\text{BDC}}(\mathbf{B}^{(0)}, (\mathbf{B}^{(0)})^*)) \} \}$; $e^{(0)} = \frac{1}{N_m} \mathcal{C}_{\text{BDC}}(\mathbf{B}^{(0)})$; $\gamma = 2$

repeat

if (pour les algorithmes du GC(P) seulement) $m \bmod N_0 = 0$ **then**

 | restart

else

 Calculer le pas optimal $\mu_{\text{opt}}^{(m)}$ (donné par (2.34) et (2.43))

 Calculer le Gradient matriciel $\nabla_{\mathbf{B}} \mathcal{C}_{\text{BDC}}(\mathbf{B}^{(m+1)}, (\mathbf{B}^{(m+1)})^*)$ donné par (2.11) (ou sa forme de vecteur $\mathbf{g}_{\mathbf{B}^*}^{(m+1)}$)

 Calculer le préconditionneur exact ou approximé $\mathbf{P}^{(m+1)} = (\hat{\mathbf{H}}^{(m+1)})^{-1}$ donné (2.20) ou (2.30) ou directement son inverse donné par (2.32)

 Calculer la matrice $\mathbf{B}^{(m+1)}$ à l'aide de (2.25)

 Pour les algorithmes du GC(P) seulement, calculer $\beta_{\text{FR}}^{(m)}$ ou $\beta_{\text{PR}}^{(m)}$ ou $\beta_{\text{DY}}^{(m)}$ donné par (2.26)-(2.28), Calculer la direction de recherche $\mathbf{D}^{(m+1)}$ à l'aide de (2.25)

 Pour l'algorithme de Levenberg-Marquardt seulement, calculer la matrice $\mathbf{B}^{(m+1)}$ à l'aide de (2.29), calculer l'erreur $e^{(m+1)} = \frac{1}{N_m} \mathcal{C}_{\text{BDC}}(\mathbf{B}^{(m+1)})$,

$\mathbf{d}_{\text{LM}}^{(m)} = \left((\hat{\mathbf{H}}^{(m)} + \alpha^{(m)} \mathbf{I}_{NM})^{-1} \mathbf{g}_{\mathbf{B}^*}^{(m)} \right)$, $dl = \frac{(\mathbf{d}_{\text{LM}}^{(m)})^H (\alpha^{(m)} \mathbf{d}_{\text{LM}}^{(m)} + \mathbf{g}_{\mathbf{B}^*}^{(m+1)})}{2}$, $\Delta e = e^{(m+1)} - e^{(m)}$

if $\Delta e > 0$, $dl > 0$ **then**

 | $\alpha^{(m+1)} = \alpha^{(m)} \max [1/3, 1 - (\frac{2\Delta e}{dl} - 1)^3]$, $e^{(m+1)} = e^{(m)}$, $\gamma = 2$

else

 | $\alpha^{(m+1)} = \gamma \alpha^{(m)}$, $\gamma = 2\gamma$

end

$m = m + 1$;

end

until ($\|\mathbf{B}^{(m+1)} - \mathbf{B}^{(m)}\|_F^2 \leq \epsilon$) ou ($m \geq N_{max}$);

TABLE 2.1 – Résumé des algorithmes de BDC_{NU}.

2.1.8 Complexité algorithmique

Pour le calcul du gradient matriciel $\nabla_{\mathbf{B}} \mathcal{C}_{\text{BDC}}(\mathbf{B}^{(m+1)}, (\mathbf{B}^{(m+1)})^*)$ dont l'expression est donnée par l'Eq. (2.11), la complexité algorithmique est :

$M \neq N$	$4N_m N M (M + N) + 2N^2 N_m$
$M \gg N$	$4N_m N M^2$
$M = N$	$8N_m N^3 + 2N_m N^2$

Pour calculer la matrice de préconditionnement selon $\mathbf{P} = \text{Diag} \{ \mathbf{H}_{\mathbf{B}, \mathbf{B}^*} (\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)) \}$, la complexité C_2 est :

$M \neq N$	$2N_m (MN(4M + 3N) + MN^4(M + N))$
$M \gg N$	$8N_m M^2 N + 2N_m M^2 N^4$
$M = N$	$8N_m N^3 + 2N^6$

Pour mettre à jour la direction de recherche donnée par l'Eq. (2.25), en choisissant la formule Polak-Ribière $\beta^{(m)}$ donnée par l'Eq. (2.27) pour tous les algorithmes proposés utilisant cette quantité, la complexité correspondante est alors :

$M \neq N$	$2M^3 N^3 + 4N_m (MN(4M + 3N) + 2MN^4(M + N)) + 2NM(M + N)$
$M \gg N$	$2M^3 N^3 + 16N_m M^2 N + 8N_m M^2 N^4 + 2NM^2$
$M = N$	$8N^6 + 28N_m N^3 + 4N^4$

Pour calculer les coefficients du polynôme d'ordre 4 $a_0^{(m)}, \dots, a_4^{(m)}$ selon les Eqs. (2.35)–(2.39) et en déduire le pas optimal $\mu_{\text{opt}}^{(m)}$, la complexité est de l'ordre de :

$M \neq N$	$24M N N_m (M + N) + 9N^2 N_m (1 + N)$
$M \gg N$	$24M^2 N N_m$
$M = N$	$57N^3 N_m + 9N^2 N_m$

La complexité totale pour chaque algorithme est alors transcrite au niveau du Tableau (2.2). Enfin, nous signalons que la complexité globale des différents algorithmes doit être multipliée par le nombre total d'itérations nécessaire N_i pour atteindre la convergence. Dans le cas d'applications pratiques, le temps de calcul nécessaire pour construire l'ensemble des matrices N_m doit également être pris en considération aussi.

Algorithme	Cas	Coût algorithmique
BDC _{GC}	$M \neq N$	$28N_mNM(M+N) + 11N^2N_m + 2N^3M^3 + NM(M+N) + 9N^3N_m$
	$M \gg N$	$11N^2N_m + 28N_mNM^2 + 2N^3M^3 + 9N^3N_m + NM^2$
	$M = N$	$17N_mN^3 + 11N^2N_m + 2N^6 + 2N^3 + 48N_mN^3$
BDC _{GCP_{HE}}	$M \neq N$	$2N_mNM(20M + 17N + 5N^3(M + N)) + 2N^3M^3 + 19N^2N_m + 9N^3N_m$
	$M \gg N$	$40N_mNM^2 + 10M^2N^4 + 2N^3M^3 + 19N^2N_m + 9N^3N_m$
	$M = N$	$83N_mN^3 + 20N^4 + 2N^6 + 19N^2N_m$
BDC _{LM_{HE}}	$M \neq N$	$2N_mNM(18M + 17N + N^3(M + N)) + 11N^2N_m + 2N^3M^3 + 9N^3N_m$
	$M \gg N$	$36N_mM^2N + 11N^2N_m + 2N_mM^2N^4 + 2N^3M^3 + 9N^3N_m$
	$M = N$	$88N_mN^3 + 4N_mN^6 + 2N^6 + 2N^2N_m$
BDC _{LM_{HA}}	$M \neq N$	$28NMN_m(M + N) + 11N^2N_m + 9N^3N_m + 25N^2M^2 + 2NM + N^3M^3$
	$M \gg N$	$28NM^2N_m + 11N^2N_m + 9N^3N_m + 25N^2M^2 + 2NM + N^3M^3$
	$M = N$	$65N^3N_m + 11N^2N_m + 25N^4 + 2N^2 + N^6$
BDC _{BFGS_{HA}}	$M \neq N$	$52N_mNM(M + N) + 4N^3M^3 + 20N^2N_m + 18N^3N_m + 10N^2M^2 + 4NM + 2NM(M + N)$
	$M \gg N$	$52N_mNM^2 + 4N^3M^3 + 20N^2N_m + 18N^3N_m + 10N^2M^2 + 4NM + 2NM^2 + N$
	$M = N$	$122N_mN^3 + 4N^6 + 20N^2N_m + 10N^4 + 4N^2 + 4N^3$
BDC _{DFP_{HA}}	$M \neq N$	$28N_mNM(M + N) + 11N^2N_m + 18N^2M^2 + 8NM + 4N^3M^3 + NM(M + N) + 9N^3N_m$
	$M \gg N$	$28N_mNM^2 + 11N^2N_m + 18N^2M^2 + 8NM + 4N^3M^3 + NM^2 + 9N^3N_m$
	$M = N$	$65N_mN^3 + 11N^2N_m + 18N^4 + 8N^2 + 4N^6 + 2N^3$
BDC _{QN_{BFGS} & DFP_{HA}}	$M \neq N$	$28N_mNM(M + N) + 11N^2N_m + 5N^2M^2 + 2NM + 9N^3N_m$
	$M \gg N$	$28N_mNM^2 + 11N^2N_m + 5N^2M^2 + 2NM + 9N^3N_m$
	$M = N$	$65N_mN^3 + 11N^2N_m + 5N^4 + 2N^2$
BDC _{QN_{HE}}	$M \neq N$	$2N_mNM(18M + 17N) + 11N^2N_m + 2N_mMN^4(M + N) + 9N^3N_m$
	$M \gg N$	$36N_mNM^2 + 11N^2N_m + 2N_mM^2N^4 + 9N^3N_m$
	$M = N$	$69N_mN^3 + 11N^2N_m + 4N_mN^6$

TABLE 2.2 – Complexité totale pour l'ensemble des algorithmes de BDC_{NU} étudiés.

2.2 Zéro bloc-diagonalisation conjointe d'un ensemble de matrices : problème et hypothèses

Dans ce paragraphe, nous introduisons un nouveau problème : celui de la décomposition matricielle conjointe d'un ensemble donné de matrices complexes (non nécessairement Hermitiennes) que nous allons choisir d'appeler "Zéro-Bloc Diagonalisation Conjointe Non-Unitaire" (ZBDC_{NU}). L'attribut non unitaire fait à nouveau référence au fait que la matrice que nous allons rechercher n'est pas forcément supposée unitaire. Nous montrerons que ce problème de zéro-bloc diagonali-

sation conjointe généralise celui de la ZDC non unitaire étudié dans [90][91] entre autres, et quels sont les liens qui unissent ces deux problèmes. Nous décrirons ensuite brièvement le principe de la zéro bloc-diagonalisation conjointe avant d'introduire la fonction de coût que nous allons chercher à minimiser pour estimer le zéro bloc-diagonaliseur conjoint (ou sa pseudo-inverse). Enfin, nous détaillerons les différents algorithmes proposés en réponse à ce problème de ZBDC_{NU} et nous illustrerons le comportement de chaque algorithme au moyen des simulations informatiques.

2.2.1 Formulation du problème de la zéro bloc-diagonalisation conjointe et ses hypothèses

Le problème de la zéro-bloc diagonalisation conjointe non unitaire (ZBDC_{NU}) s'énonce comme suit. Considérons un ensemble $\mathcal{N}$ de N_m ($N_m \in \mathbb{N}^*$) matrices carrées $\mathbf{X}_i \in \mathbb{C}^{M \times M}$, $i \in \{1, \dots, N_m\}$ se décomposant toutes de la manière suivante :

$$\mathbf{X}_i = \mathbf{A} \mathbf{Z}_i \mathbf{A}^H, \quad (2.44)$$

où $(\cdot)^H$ désigne l'opérateur de transposition conjugaison et les matrices

$$\mathbf{Z}_i = \begin{pmatrix} \mathbf{0}_{11} & \mathbf{Z}_{i,12} & \dots & \mathbf{Z}_{i,1r} \\ \mathbf{Z}_{i,21} & \mathbf{0}_{22} & \ddots & \mathbf{Z}_{i,2r} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{Z}_{i,r1} & \dots & \vdots & \mathbf{0}_{rr} \end{pmatrix} \quad \forall i \in \{1, \dots, N_m\},$$

sont des matrices zéro-bloc diagonales de dimension $N \times N$, où $r \in \mathbb{N}^*$ est le nombre de blocs considérés. Les matrices $\mathbf{Z}_{i,jk}$, $i \in \{1, \dots, N_m\}$, $j, k \in \{1, \dots, r\}$ sont des matrices rectangulaires de dimension $n_j \times n_k$ de telle sorte que $n_1 + \dots + n_r = N$ alors que $\mathbf{0}_{jj}$ désigne la matrice nulle de dimension $n_j \times n_j$. $\mathbf{A}$ est une matrice de rang plein de dimension $M \times N$ ($M \geq N$) tandis que sa pseudo-inverse (ou inverse généralisée de Moore-Penrose) de taille $N \times M$ est notée $\mathbf{B}$. L'ensemble des N_m matrices $\mathbf{Z}_i \in \mathbb{C}^{N \times N}$ zéro-bloc diagonales est noté $\mathcal{Z}$. Nous rappelons que la taille des blocs n_j pour tout $j = 1, \dots, r$ est supposée connue.

Dans un contexte SAS, $\mathbf{A}$ désigne toujours la matrice de mélange tandis que $\mathbf{B}$ représente la matrice de séparation, N est le nombre de sources et M le nombre de capteurs.

Le problème de la ZBDC_{NU} consiste alors à estimer l'ensemble des matrices zéro-bloc diagonales $\mathcal{Z}$ et la matrice $\mathbf{A}$ ou bien la matrice $\mathbf{B}$ et ce uniquement à partir des matrices de l'ensemble $\mathcal{N}$.

2.2.2 Principe de la zéro-bloc diagonalisation conjointe

En suivant le même raisonnement qu'en BDC, le problème de la ZBDC peut se ramener à un problème de minimisation. On peut en effet chercher à minimiser la distance au carré entre les matrices observées et leur approximation soit finalement la fonction de coût quadratique suivante :

$$\mathcal{C}_{\text{ZBDC}}^{(1)}(\mathbf{A}, \{\mathbf{Z}_i\}) = \sum_{i=1}^{N_m} \|\mathbf{X}_i - \mathbf{A} \mathbf{Z}_i \mathbf{A}^H\|_F^2, \quad (2.45)$$

où $\|\cdot\|_F$ désigne la norme de Frobenius.

On peut procéder de manière légèrement différente en multipliant à gauche (resp. à droite) la matrice $\mathbf{X}_i$ décrite au niveau de l'Eq. (2.44) par la pseudo-inverse $\mathbf{A}^\dagger$ de la matrice $\mathbf{A}$ (resp. par $(\mathbf{A}^\dagger)^H$). Ce qui conduit alors à :

$$\mathbf{B} \mathbf{X}_i \mathbf{B}^H = \mathbf{Z}_i, \quad \forall i \in \{1, \dots, N_m\}, \quad (2.46)$$

où $\mathbf{B} = \mathbf{A}^\dagger$. Pour estimer la matrice $\mathbf{B}$, nous proposons donc d'utiliser plutôt la fonction de coût suivante :

$$\mathcal{C}_{\text{ZBDC}}^{(2)}(\mathbf{B}, \{\mathbf{Z}_i\}) = \sum_{i=1}^{N_m} \|\mathbf{B}\mathbf{X}_i\mathbf{B}^H - \mathbf{Z}_i\|_F^2. \quad (2.47)$$

Pour des raisons de simplicité, c'est cette fonction de coût $\mathcal{C}_{\text{ZBDC}}^{(2)}$ que nous allons exploiter. Cela vient du fait que la minimisation de $\mathcal{C}_{\text{ZBDC}}^{(2)}$ par rapport aux matrices $\mathbf{Z}_i$, $i \in \{1, \dots, N_m\}$ possède une solution directe lorsque la matrice $\mathbf{B}$ est fixée. Au préalable, introduisons quelques notations utiles. Si nous considérons une matrice carrée $\mathbf{X}$ de taille $N \times N$, telle que :

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_{11} & \mathbf{X}_{12} & \dots & \mathbf{X}_{1r} \\ \mathbf{X}_{21} & \dots & \dots & \vdots \\ \vdots & \dots & \dots & \vdots \\ \mathbf{X}_{r1} & \mathbf{X}_{r2} & \dots & \mathbf{X}_{rr} \end{pmatrix}, \quad (2.48)$$

dont les blocs $\mathbf{X}_{ij}$ sont de taille $n_i \times n_j$, pour tout i et j allant de 1 à r (avec $n_1 + \dots + n_r = N$). En introduisant, le vecteur $\mathbf{n} = (n_1, n_2, \dots, n_r)$, l'opérateur matriciel $\text{Bdiag}_{(\mathbf{n})}\{\cdot\}$ est défini comme au niveau de l'Eq. (2.4) pour une matrice carrée $\mathbf{X} = (X_{ij})$.

Calculons maintenant :

$$\hat{\mathbf{Z}}_i = \arg \min_{\mathbf{Z}_i} \mathcal{C}_{\text{ZBDC}}^{(2)}(\mathbf{B}, \{\mathbf{Z}_i\}). \quad (2.49)$$

En nous appuyant sur les propriétés rappelées dans l'Annexe 1, nous trouvons assez facilement que :

$$\begin{aligned} \mathcal{C}_{\text{ZBDC}}^{(2)}(\mathbf{B}, \{\mathbf{Z}_i\}) &= \sum_{i=1}^{N_m} \|\mathbf{B}\mathbf{X}_i\mathbf{B}^H - \mathbf{Z}_i\|_F^2 = \sum_{i=1}^{N_m} \text{tr}\{(\mathbf{B}\mathbf{X}_i\mathbf{B}^H - \mathbf{Z}_i)^H(\mathbf{B}\mathbf{X}_i\mathbf{B}^H - \mathbf{Z}_i)\} \\ &= \sum_{i=1}^{N_m} \text{tr}\{\mathbf{B}\mathbf{X}_i^H\mathbf{B}^H\mathbf{B}\mathbf{X}_i\mathbf{B}^H - \mathbf{Z}_i^H\mathbf{B}\mathbf{X}_i\mathbf{B}^H - \mathbf{B}\mathbf{X}_i^H\mathbf{B}^H\mathbf{Z}_i + \mathbf{Z}_i^H\mathbf{Z}_i\}. \end{aligned} \quad (2.50)$$

De cette équation, on peut calculer, la dérivée partielle :

$$\frac{\partial}{\partial \mathbf{Z}_i} \mathcal{C}_{\text{ZBDC}}^{(2)}(\mathbf{B}, \{\mathbf{Z}_i\}) = -2(\mathbf{B}\mathbf{X}_i\mathbf{B}^H - \mathbf{Z}_i) = 0 \Rightarrow \hat{\mathbf{Z}}_i = \text{OffBdiag}_{(\mathbf{n})}\{\mathbf{B}\mathbf{X}_i\mathbf{B}^H\} \quad \forall i = 1, \dots, N_m, \quad (2.51)$$

parce que les matrices $\mathbf{Z}_i$ pour tout $i = 1, \dots, N_m$ sont des matrices zéro-bloc diagonales. Finalement en utilisant l'Eq. (2.4), on peut considérer la fonction de coût suivante :

$$\mathcal{C}_{\text{ZBDC}}^{(2)}(\mathbf{B}, \{\hat{\mathbf{Z}}_i\}) = \sum_{i=1}^{N_m} \|\text{Bdiag}_{(\mathbf{n})}\{\mathbf{B}\mathbf{X}_i\mathbf{B}^H\}\|_F^2 \stackrel{\text{def}}{=} \mathcal{C}_{\text{ZBDC}}(\mathbf{B}). \quad (2.52)$$

2.2.3 Lien entre ZBDC et ZDC

Notons que lorsque les matrices $\mathbf{Z}_{i,jj}$ pour tout $j \in \{1, \dots, r\}$ et pour tout $i \in \{1, \dots, N_m\}$ se réduisent à des scalaires, ($n_j = 1$ pour tout $j \in \{1, \dots, r\}$ impliquant par alors que $r = N$),

nous retombons sur un problème de zéro-diagonalisation conjointe non-unitaire d'un ensemble de matrices donné [90][91]. L'opérateur matriciel se simplifie alors en $\text{Diag}\{\cdot\}$:

$$\mathbf{B}\text{diag}_{(\mathbf{1})}\{\mathbf{X}\} = ((1 - \delta_{ij})X_{ij})\mathbf{1}_N = \text{Diag}\{\mathbf{X}\},$$

où $\mathbf{X} = (X_{ij})$, $\mathbf{1}_N$ est la matrice de dimension $N \times N$ dont toutes les composantes sont égales à 1 et $\delta_{ij} = 1$ si $i = j$ et 0 sinon.

Ce qui conduit à la minimisation de la fonction de coût suivante :

$$\mathcal{C}_{\text{ZD}}(\mathbf{B}) = \sum_{k=1}^{N_m} \|\text{Diag}\{\mathbf{B}\mathbf{X}_k\mathbf{B}^H\}\|_F^2. \quad (2.53)$$

Plusieurs approches ont été développées pour l'estimation de la matrice $\mathbf{A}$ (et les matrices $\mathbf{X}_i$) ou de la matrice $\mathbf{B}$. Le premier cas particulier à avoir été considéré est celui de la zéro-diagonalisation conjointe sous la contrainte d'une matrice $\mathbf{A}$ unitaire [21]. Par la suite, d'autres solutions ont été proposées dans [90][91][49][258][259] pour l'estimation d'une matrice $\mathbf{B}$ non unitaire sous différentes hypothèses sur l'ensemble des matrices $\mathcal{N}$. Récemment, le problème de la zéro diagonalisation conjointe non symétrique pour lequel les matrices à droite et gauche impliquées dans la décomposition ne sont plus nécessairement les mêmes, a été considéré dans [47].

2.2.4 Trois nouveaux algorithmes de zéro-bloc diagonalisation conjointe non unitaire

Ayant justifié le choix de la fonction de coût utilisée (celle donnée au niveau de l'Eq. (2.52), l'objectif est maintenant de la minimiser afin d'estimer la matrice $\mathbf{B} \in \mathbb{C}^{N \times M}$. Nous allons proposer de nouveaux algorithmes basés sur des approches itératives. Mais au préalable, il va nous falloir calculer certaines quantités : le gradient matriciel complexe de la fonction de coût mais aussi les matrices Hessiennes complexes. Pour ce faire, - comme précédemment - nous allons commencer par calculer la différentielle $d\mathcal{C}_{\text{ZBDC}}(\mathbf{B})$ de $\mathcal{C}_{\text{ZBDC}}(\mathbf{B})$ en passant par les dérivées partielles $\frac{\partial}{\partial \cdot}$ de la fonction de coût $\mathcal{C}_{\text{ZBDC}}(\mathbf{B})$ par rapport à $\mathbf{B}$ et $\mathbf{B}^*$ (la fonction de coût sera dorénavant notée $\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)$). Dans un second temps, nous nous intéresserons à la différentielle d'ordre deux afin d'en déduire l'expression des matrices Hessiennes complexes.

2.2.5 Gradient matriciel complexe

Le gradient matriciel complexe de la fonction de coût scalaire à valeur réelle donnée au niveau de l'Eq. (2.52) est noté $\nabla_{\mathbf{B}}\mathcal{C}_{\text{ZBDC}}$. Nous en définissons également une version vectorielle que nous notons $\mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*))$. Ces quantités sont définies comme dans [129], par :

$$\nabla_{\mathbf{B}}\mathcal{C}_{\text{ZBDC}} \stackrel{\text{def}}{=} \frac{\partial \mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}^*} \quad \text{et} \quad \mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)) = \text{vec}^T \left(\frac{\partial \mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}^*} \right), \quad (2.54)$$

où $\text{vec}^T(\mathbf{A}) = (\text{vec}(\mathbf{A}))^T$, si $\text{vec}(\cdot)$ est l'opérateur de vectorisation qui permet de ranger les colonnes d'une matrice $N \times M$ dans un vecteur de taille $(N \times M) \times 1$ et $(\cdot)^T$ désigne l'opérateur de transposition. La matrice du gradient complexe $\nabla_{\mathbf{B}}\mathcal{C}_{\text{ZBDC}}$ de la fonction de coût quadratique $\mathcal{C}_{\text{ZBDC}}^{(2)}(\mathbf{B}, \mathbf{B}^*)$ est calculée au niveau de l'Annexe 3. Nous y démontrons qu'elle vaut :

$$\frac{\partial \mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}^*} = \sum_{i=1}^{N_m} \left[\mathbf{B}\text{diag}_{(\mathbf{n})}\{\mathbf{B}\mathbf{X}_i\mathbf{B}^H\}\mathbf{B}\mathbf{X}_i^H + \left(\mathbf{B}\text{diag}_{(\mathbf{n})}\{\mathbf{B}\mathbf{X}_i\mathbf{B}^H\} \right)^H \mathbf{B}\mathbf{X}_i \right]. \quad (2.55)$$

2.2.5.1 Matrices Hessiennes complexes

La matrice $\mathbf{H}$ complexe définie comme au niveau de l'Eq. (2.13) est une matrice carrée de dimensions $2NM \times 2NM$ constituée de quatre blocs/matrices carrées de dimensions $NM \times NM$ à savoir les matrices Hessiennes.

Pour calculer cette matrice $\mathbf{H}$, nous suivons le même cheminement que précédemment mais en nous intéressant à $d\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)$ et non plus à $d\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)$ (le détail des calculs peut être trouvé au niveau de l'Annexe 3). Nous calculons ensuite la différentielle d'ordre deux de $\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)$ qui s'écrit sous la même forme que celle donnée au niveau de l'Eq. (2.14). Les quatre matrices complexes $\mathbf{A}_{00}$, $\mathbf{A}_{01}$, $\mathbf{A}_{10}$ et $\mathbf{A}_{11}$ de taille $NM \times NM$ données ci-dessous sont liées aux matrices Hessiennes par les Eqs. (2.15-2.17).

Nous avons établi qu'elles valaient (voir l'Annexe 3 pour le détail des calculs) :

$$\begin{aligned} \mathbf{A}_{00} &= (\mathbf{M}_i^T \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{\text{Bdiag}}^T (\mathbf{B}^* \mathbf{M}_i^* \otimes \mathbf{I}_N) + (\mathbf{M}_i^* \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{\text{Bdiag}}^T (\mathbf{B}^* \mathbf{M}_i^T \otimes \mathbf{I}_N) \\ &\quad + \mathbf{M}_i^* \otimes \text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{M}_i \mathbf{B}^H \} + \mathbf{M}_i^T \otimes \text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{M}_i^H \mathbf{B}^H \} \\ &= \mathbf{A}_{11}^*, \end{aligned} \quad (2.56)$$

$$\begin{aligned} \mathbf{A}_{10} &= \mathbf{K}_{N,M}^T (\mathbf{I}_N \otimes \mathbf{M}_i \mathbf{B}^H) \mathbf{T}_{\text{Bdiag}}^T (\mathbf{B}^* \mathbf{M}_i^* \otimes \mathbf{I}_N) \\ &\quad + \mathbf{K}_{N,M}^T (\mathbf{I}_N \otimes \mathbf{M}_i^H \mathbf{B}^H) \mathbf{T}_{\text{Bdiag}}^T (\mathbf{B}^* \mathbf{M}_i^T \otimes \mathbf{I}_N) = \mathbf{A}_{01}^*, \end{aligned} \quad (2.57)$$

où l'opérateur $\otimes$ désigne le produit de Kronecker, $\mathbf{K}_{N,M}$ est une matrice carrée de permutation de taille $(MN \times MN)$ et $\mathbf{T}_{\text{Bdiag}} = \text{diag}\{\text{vec}\{\text{Bdiag}\{\mathbf{1}_N\}\}\}$, $\mathbf{1}_N$ est la matrice $N \times N$ dont tous les éléments sont égaux à un et $\text{diag}\{\mathbf{a}\}$ est une matrice diagonale carrée qui contient les éléments du vecteur $\mathbf{a}$ sur sa diagonale.

Pour assurer la minimisation de la fonction de coût et réduire la complexité du calcul de l'inverse de la matrice Hessienne, nous considérons la diagonale de la matrice Hessienne (*i.e.* Eqs. (2.15) et (2.56)). La matrice $NM \times NM$ de préconditionnement, notée $\mathbf{P}$, que nous suggérons donc d'utiliser dans ce contexte, s'écrit sous la forme suivante :

$$\mathbf{P} = (\hat{\mathbf{H}})^{-1}, \quad (2.58)$$

$$\hat{\mathbf{H}} = \text{Diag}\{\mathbf{H}_{\mathbf{B}, \mathbf{B}^*}(\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*))\} = \text{Diag}\{\mathbf{H}_{\mathbf{B}^*, \mathbf{B}}(\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*))\}. \quad (2.59)$$

L'objectif est maintenant d'utiliser les quantités qui viennent d'être calculées afin d'en déduire différents algorithmes d'optimisation. Nous évaluerons ensuite leur comportement dans différents contextes.

2.2.5.2 Gradient conjugué et gradient conjugué préconditionné non linéaire

Nous rappelons que lorsque nous considérons l'algorithme du gradient conjugué préconditionné, la direction de recherche $\mathbf{d}$ à la première itération est celle donnée au niveau de l'Eq. (2.24). Ensuite, $\forall m = 1, 2, \dots$, la règle d'adaptation est celle de l'Eq. (2.25).

Les algorithmes résultant seront notés $\text{ZBDC}_{\text{GCPHE}}$ pour le gradient conjugué préconditionné et ZBDC_{GC} pour le gradient conjugué (cas où le préconditionneur est la matrice identité).

2.2.5.3 Levenberg-Marquardt

Pour garantir la construction possible des directions de descente, l'algorithme de Levenberg-Marquardt [160] peut être utilisé. Il modifie $\hat{\mathbf{H}}$ en lui ajoutant un multiple de la matrice d'identité comme au niveau de l'Eq. (2.29) :

$$\mathbf{b}^{(m+1)} = \mathbf{b}^{(m)} - \mu^{(m)} \left(\widehat{\mathbf{H}}^{(m)} + \alpha^{(m)} \mathbf{I}_{NM} \right)^{-1} \mathbf{g}_{\mathbf{B}^*}^{(m)}, \quad (2.60)$$

Le choix du paramètre $\alpha^{(m)}$ s'effectue comme dans §2.1.6.2 (pour plus de détails sur la procédure de mise à jour utilisée, le lecteur peut consulter le résumé des algorithmes de ZBDC_{NU} dans §2.2.6). L'algorithme résultant est noté ZBDC_{LMHE}.

Nous notons que tous ces algorithmes ont déjà fait l'objet de publications [57][56].

2.2.5.4 Recherche du pas optimal

Dans ce paragraphe, nous nous concentrons sur le calcul du pas optimal toujours noté μ_{opt} . Comme pour la BDC_{NU}, le calcul algébrique du pas optimal (quand l'algorithme du gradient conjugué préconditionné est considéré) se fait en développant la quantité suivante : $\mathcal{C}_{\text{ZBDC}}(\mathbf{B}^{(m+1)}) = \mathcal{C}_{\text{ZBDC}}(\mathbf{B}^{(m)} - \mu \mathbf{D}^{(m)})$, avec $\mathbf{D}^{(m)} = \text{unvec}(\mathbf{d}^{(m)})$ ou encore $\mathbf{d}^{(m)} = \text{vec}(\mathbf{D}^{(m)})$ où le vecteur $\mathbf{d}^{(m)}$ est donné dans l'Eq. (2.25). Il s'agit d'un polynôme d'ordre quatre dont l'expression est donnée par l'équation suivante (pour des raisons de simplification, la dépendance vis-à-vis de m est omise) :

$$\mathcal{C}_{\text{ZBDC}}(\mathbf{B} - \mu \mathbf{D}) = a_0 + a_1 \mu + a_2 \mu^2 + a_3 \mu^3 + a_4 \mu^4, \quad (2.61)$$

où les cinq coefficients a_0, a_1, a_2, a_3 et a_4 valent (voir l'Annexe 3 pour le détail des calculs),

$$a_0 = \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{K}_3^H \mathbf{B} \text{diag}_{(\mathbf{n})} \{ \mathbf{K}_3 \} \right\}, \quad (2.62)$$

$$a_1 = - \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{K}_1 \mathbf{B} \text{diag}_{(\mathbf{n})} \{ \mathbf{K}_3 \} + \mathbf{K}_3^H \mathbf{K}_2 \right\}, \quad (2.63)$$

$$a_2 = \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{K}_0^H \mathbf{B} \text{diag}_{(\mathbf{n})} \{ \mathbf{K}_3 \} + \mathbf{K}_3^H \mathbf{B} \text{diag}_{(\mathbf{n})} \{ \mathbf{K}_0 \} + \mathbf{K}_1 \mathbf{K}_2 \right\}, \quad (2.64)$$

$$a_3 = - \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{K}_0^H \mathbf{K}_2 + \mathbf{K}_1 \mathbf{B} \text{diag}_{(\mathbf{n})} \{ \mathbf{K}_0 \} \right\}, \quad (2.65)$$

$$a_4 = \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{K}_0^H \mathbf{B} \text{diag}_{(\mathbf{n})} \{ \mathbf{K}_0 \} \right\}, \quad (2.66)$$

où :

$$\begin{aligned} \mathbf{K}_0 &= \mathbf{D}_{\mathbf{B}}^{(m)} \mathbf{X}_i (\mathbf{D}_{\mathbf{B}}^{(m)})^H, \\ \mathbf{K}_1 &= \mathbf{D}_{\mathbf{B}}^{(m)} \mathbf{X}_i^H \mathbf{B}^H + \mathbf{B} \mathbf{X}_i^H (\mathbf{D}_{\mathbf{B}}^{(m)})^H, \\ \mathbf{K}_2 &= \mathbf{B} \text{diag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i (\mathbf{D}_{\mathbf{B}}^{(m)})^H + \mathbf{D}_{\mathbf{B}}^{(m)} \mathbf{X}_i \mathbf{B}^H \}, \\ \mathbf{K}_3 &= \mathbf{B} \mathbf{X}_i \mathbf{B}^H. \end{aligned}$$

Le polynôme d'ordre trois (Eq. (2.67)) correspondant à la dérivée par rapport à μ du polynôme d'ordre quatre donné par l'Eq. (2.61) vaut :

$$\frac{\partial \mathcal{C}_{\text{ZBDC}}(\mathbf{B} - \mu \mathbf{D})}{\partial \mu} = 4a_4 \mu^3 + 3a_3 \mu^2 + 2a_2 \mu + a_1. \quad (2.67)$$

Le pas optimal μ_{opt} correspond alors à la racine (2.67) qui minimise (2.61).

2.2.6 Algorithmes de ZBDC_{NU} en résumé

Les algorithmes de ZBDC_{NU} (notés par ZBDC_{EH}) sont résumés au niveau du Tableau 2.2.6 situé ci-dessous :

Data : N_m matrices carées $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{N_m}$, critère d'arrêt ϵ , pas μ , nombre d'itérations max N_{max} , nombre utilisé pour le restart N_0

Result : Estimation de la matrice $\mathbf{B}$

initialiser : $\mathbf{B}^{(0)}$; $\mathbf{D}^{(0)}$; $m = 0$; facteur d'amortissement $\alpha^{(0)} = 10^{-2} \max \{ \text{Diag} \{ \mathbf{H}_{\mathbf{B}, \mathbf{B}^*} (\mathcal{C}_{\text{ZBDC}}(\mathbf{B}^{(0)}, (\mathbf{B}^{(0)})^*)) \} \}$; $e^{(0)} = \frac{1}{N_m} \mathcal{C}_{\text{ZBDC}}(\mathbf{B}^{(0)})$; $\gamma = 2$

repeat

if (pour les algorithmes du GC(P) seulement) $m \bmod N_0 = 0$ **then**

 | restart

else

 | Calculer le pas optimal $\mu_{\text{opt}}^{(m)}$ (donné par (2.67) et (2.61))

 | Calculer le Gradient matriciel $\nabla_{\mathbf{B}} \mathcal{C}_{\text{ZBDC}}(\mathbf{B}^{(m+1)}, (\mathbf{B}^{(m+1)})^*)$ donné par (2.55) (ou sa forme de vecteur $\mathbf{g}_{\mathbf{B}^*}^{(m+1)}$)

 | Calculer le préconditionneur exact $\mathbf{P}^{(m+1)}$ donné par (2.58)

 | Calculer la matrice $\mathbf{B}^{(m+1)}$ à l'aide de (2.25)

 | Pour les algorithmes du GC(P) seulement, calculer $\beta_{\text{FR}}^{(m)}$ ou $\beta_{\text{PR}}^{(m)}$ ou $\beta_{\text{DY}}^{(m)}$ donné par (2.26)-(2.28), calculer la direction de recherche $\mathbf{D}^{(m+1)}$ à l'aide de (2.25)

 | Pour l'algorithme de Levenberg-Marquardt seulement, calculer la matrice $\mathbf{B}^{(m+1)}$ à l'aide de (2.60), calculer l'erreur $e^{(m+1)} = \frac{1}{N_m} \mathcal{C}_{\text{ZBDC}}(\mathbf{B}^{(m+1)})$,

 | $\mathbf{d}_{\text{LM}}^{(m)} = \left(\left(\hat{\mathbf{H}}^{(m)} + \alpha^{(m)} \mathbf{I}_{NM} \right)^{-1} \mathbf{g}_{\mathbf{B}^*}^{(m)} \right)$, $dl = \frac{(\mathbf{d}_{\text{LM}}^{(m)})^H (\alpha^{(m)} \mathbf{d}_{\text{LM}}^{(m)} + \mathbf{g}_{\mathbf{B}^*}^{(m+1)})}{2}$, $\Delta e = \frac{e^{(m+1)} - e^{(m)}}{2}$

 | **if** $\Delta e > 0$, $dl > 0$ **then**

 | $\alpha^{(m+1)} = \alpha^{(m)} \max [1/3, 1 - (\frac{2\Delta e}{dl} - 1)^3]$, $e^{(m+1)} = e^{(m)}$, $\gamma = 2$

 | **else**

 | $\alpha^{(m+1)} = \gamma \alpha^{(m)}$, $\gamma = 2\gamma$

 | **end**

 | $m = m + 1$;

end

until ($\|\mathbf{B}^{(m+1)} - \mathbf{B}^{(m)}\|_F^2 \leq \epsilon$) ou ($m \geq N_{max}$);

TABLE 2.3 – Résumé des algorithmes de ZBDC_{NU} proposés.

2.2.7 Complexité algorithmique

Le calcul le gradient matriciel $\nabla_{\mathbf{B}} \mathcal{C}_{\text{ZBDC}}(\mathbf{B}^{(m+1)}, (\mathbf{B}^{(m+1)})^*)$ dont l'expression est donnée au niveau de l'Eq. (2.55), du préconditionneur $\mathbf{P} = \text{Diag} \{ \mathbf{H}_{\mathbf{B}, \mathbf{B}^*} (\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)) \}$, de la direction de recherche donnée par l'Eq. (2.25) et des coefficients $a_0^{(m)}, \dots, a_4^{(m)}$ selon les Eqs. (2.62)–(2.66) afin de calculer le pas optimal $\mu_{\text{opt}}^{(m)}$, exigent exactement le même coût algorithmique que lorsque le problème de la bloc diagonalisation conjointe est considéré. La complexité totale des différents

algorithmes présentés est résumée ci-dessous (Tab. 2.4) :

Algorithme	Cas	Coût algorithmique
ZBDC _{GC}	$M \neq N$	$28N_mNM(M+N) + 11N^2N_m + 2N^3M^3 + NM(M+N) + 9N^3N_m$
	$M \gg N$	$11N^2N_m + 28N_mNM^2 + 2N^3M^3 + 9N^3N_m + NM^2$
	$M = N$	$17N_mN^3 + 11N^2N_m + 2N^6 + 2N^3 + 48N_mN^3$
ZBDC _{GCP_{HE}}	$M \neq N$	$2N_mNM(20M + 17N + 5N^3(M+N)) + 2N^3M^3 + 19N^2N_m + 9N^3N_m$
	$M \gg N$	$40N_mNM^2 + 10M^2N^4 + 2N^3M^3 + 19N^2N_m + 9N^3N_m$
	$M = N$	$83N_mN^3 + 20N^4 + 2N^6 + 19N^2N_m$
ZBDC _{LM_{HE}}	$M \neq N$	$2N_mNM(18M + 17N + N^3(M+N)) + 11N^2N_m + 2N^3M^3 + 9N^3N_m$
	$M \gg N$	$36N_mM^2N + 11N^2N_m + 2N_mM^2N^4 + 2N^3M^3 + 9N^3N_m$
	$M = N$	$88N_mN^3 + 4N_mN^6 + 2N^6 + 2N^2N_m$

TABLE 2.4 – Complexité totale des algorithmes de ZBDC_{NU} proposés.

Nous rappelons que la complexité globale des différents algorithmes doit être multipliée par le nombre total d'itérations nécessaires N_i pour atteindre la convergence. Dans le cas d'applications pratiques, le temps de calcul nécessaire pour construire l'ensemble des matrices N_m doit également être pris en considération.

2.3 Simulations informatiques et comparaisons

Afin d'étudier le comportement des différents algorithmes proposés en réponse au problème de la bloc diagonalisation conjointe non unitaire (resp. zéro bloc diagonalisation conjointe non unitaire), nous allons effectuer des simulations numériques sur données entièrement simulées dans différents contextes. Nous considérons des ensembles $\mathcal{D}$ (resp. $\mathcal{Z}$) (que nous construirons) de N_m ($N_m \in \mathbb{N}^*$) matrices carrées, de taille $N \times N$, complexes, exactement bloc diagonales (resp. zéro-bloc diagonales). Ces matrices sont générées selon un tirage aléatoire : les différentes composantes de ces matrices suivent une loi Gaussienne. Ensuite, ces matrices seront perturbées par un bruit additif afin nous écarter de l'hypothèse de bloc-diagonalité (resp. zéro-bloc diagonalité) et ainsi pouvoir apprécier l'impact d'une erreur de modèle au niveau des algorithmes. Dans ce cas là, une matrice complexe dont les termes aléatoires sont tirés selon une loi Gaussienne de moyenne nulle et de variance σ_b^2 est ajoutée à la matrice précédente. Le rapport signal à bruit (RSB) peut alors être défini en dB comme $\text{RSB} = 10 \log(\frac{1}{\sigma_b^2})$.

Pour juger de la qualité de l'estimation de la matrice $\mathbf{B}$ (bloc-diagonaliseur conjoint commun à toutes les matrices de l'ensemble $\mathcal{M}$ (resp. zéro-bloc diagonaliseur conjoint commun à toutes les matrices de l'ensemble $\mathcal{N}$)), nous utilisons un indice d'erreur que nous notons $\mathbf{l}_{\text{Conv}}(\mathbf{G})$. Il est fourni au niveau de l'Eq. (1.30). Nous rappelons de nouveau que la matrice produite $\mathbf{G}$ vaut, dans notre cas, $\mathbf{G} = \hat{\mathbf{B}}\mathbf{A}$. l'estimation est parfaite, quand l'indice d'erreur tend vers $-\infty$ en échelle logarithmique et/ou 0 en échelle linéaire. Quand $\mathbf{l}_{\text{Conv}}(\cdot)$ est donné en dB, il est défini comme $\mathbf{l}_{\text{Conv}}(\cdot) \text{ dB} = 10 \log(\mathbf{l}_{\text{Conv}}(\cdot))$. Nous précisons enfin que tous les résultats présentés dans la suite résultent du moyennage par méthode de Monte-Carlo sur 30 tirages aléatoires.

2.3.1 Cas général de la BDC d'un ensemble de matrices complexes

Nous étudions les algorithmes fondés sur différents schémas d'optimisation itératifs à savoir :

- le gradient conjugué (BDC_{GC}),
- le gradient conjugué préconditionné ($\text{BDC}_{\text{GCP}_{\text{HE}}}$, $\text{BDC}_{\text{GCP}_{\text{BFGS-HA}}}$ et $\text{BDC}_{\text{GCP}_{\text{DFP-HIA}}}$),
- l'algorithme de Levenberg-Marquardt ($\text{BDC}_{\text{LM}_{\text{HE}}}$ et $\text{BDC}_{\text{LM}_{\text{HA}}}$),
- et des méthodes quasi-Newton ($\text{BDC}_{\text{QN}_{\text{HE}}}$, $\text{BDC}_{\text{QN}_{\text{BFGS-HA}}}$ et $\text{BDC}_{\text{QN}_{\text{DFP-HA}}}$),

afin de déterminer quelle méthode offre le meilleur compromis entre efficacité et vitesse de convergence dans le contexte étudié. Tous ces algorithmes sont comparés à d'autres algorithmes de la littérature à savoir : celui de D. Nion basé lui aussi sur le gradient conjugué [184] ($\text{BDC}_{\text{GC}_{\text{Nion}}}$) mais utilisant la première fonction de coût introduite et non pas celle que nous utilisons et ceux proposés par H. Ghennioui utilisant la même fonction de coût plus "directe" et fondés sur un schéma d'optimisation algébrique (BDC_{Alg}) [109], ou sur des approches de type gradient (le calcul du gradient matriciel pouvant être approché comme dans $\text{BDC}_{\text{App-Grad}}$ [113] ou exact à pas optimal ($\text{BDC}_{\text{GradAbs}}$), un gradient relatif $\text{BDC}_{\text{GradR}}$ [112] peut également être utilisé à la place du gradient classique).

Cette étude comporte plusieurs cas de figures ou scénarios. Nous commençons, dans un premier temps, par l'initialisation à partir de la DVPG proposée par D. Nion dans [184]. Ensuite, nous initialisons tous les algorithmes à partir d'une même solution choisie de manière aléatoire (afin de partir plus loin de la solution) et d'étudier la robustesse des algorithmes vis-à-vis du choix de la solution initiale. Il faut mentionner aussi que le cas d'une matrice $\mathbf{A}$ (éventuellement $\mathbf{B}$) rectangulaire est pris en compte.

Nous considérons dans un premier temps le cas $M = N = 8$ (la matrice $\mathbf{A}$ est carrée) et $r = 4$ ($r = 4$ pour tout $j = 1, \dots, 4$). La matrice $\mathbf{A}$ est choisie de manière aléatoire selon une loi uniforme de moyenne nulle et de variance un sur l'intervalle $[-1, 1]$. L'ensemble $\mathcal{M}$ des matrices considérées $\mathbf{M}_i$, $i = 1, \dots, N_m$ est constitué de $N_m = 100$ (resp. 5) matrices complexes de taille $M \times M$. Nous notons que ces algorithmes sont initialisés à partir de la DVPG décrite dans [184].

Au niveau de la Fig. 2.1, nous traçons l'évolution de l'indice d'erreur $\text{I}_{\text{Conv}}(\cdot)$ en dB en fonction du nombre d'itérations dans le cas non bruité où les matrices de l'ensemble $\mathcal{D}$ sont quasiment bloc-diagonales ($\text{RSB} = 100$ dB). Par ailleurs, deux cas sont étudiés : i) sur la courbe de gauche, l'ensemble à bloc diagonaliser conjointement est constitué de $N_m = 100$ matrices tandis qu'à droite de cette même figure, l'ensemble étudié n'est que de $N_m = 5$ matrices. Au niveau de la Fig. 2.2, nous représentons l'évolution de l'indice d'erreur $\text{I}_{\text{Conv}}(\cdot)$ en fonction du nombre d'itérations pour les 14 algorithmes lorsque l'on s'écarte de l'hypothèse de bloc-diagonalité ($\text{RSB} = 30$ dB). À nouveau, pour la courbe de gauche, l'ensemble sur lequel nous travaillons est constitué de $N_m = 100$ matrices tandis qu'à droite nous ne considérons plus que $N_m = 5$ matrices.

Au regard de ces courbes, nous pouvons tirer plusieurs conclusions. Tout d'abord, nous observons que la vitesse de convergence des algorithmes reposant sur l'utilisation d'une matrice de préconditionnement utilisant le calcul exact des matrices Hessiennes complexes (BDC_{HE}) est plus rapide que celle des algorithmes fondés sur une approximation de ces matrices ($\text{BDC}_{\text{H(1)A}}$). Ensuite, en comparant la vitesse convergence des diverses approches du type gradient notamment BDC_{GC} , $\text{BDC}_{\text{GradR}}$, $\text{BDC}_{\text{App-Grad}}$, $\text{BDC}_{\text{GradAbs}}$ et $\text{BDC}_{\text{GC}_{\text{Nion}}}$ à l'algorithme $\text{BDC}_{\text{GCP}_{\text{HE}}}$, nous montrons l'intérêt d'utiliser un préconditionneur. De plus, nous constatons que parmi les algorithmes non préconditionnés, l'algorithme BDC_{GC} est plus rapide que les algorithmes $\text{BDC}_{\text{GradR}}$, $\text{BDC}_{\text{App-Grad}}$, $\text{BDC}_{\text{GradAbs}}$, BDC_{Alg} et que l'algorithme de D. Nion ($\text{BDC}_{\text{GC}_{\text{Nion}}}$) fondé quant à lui

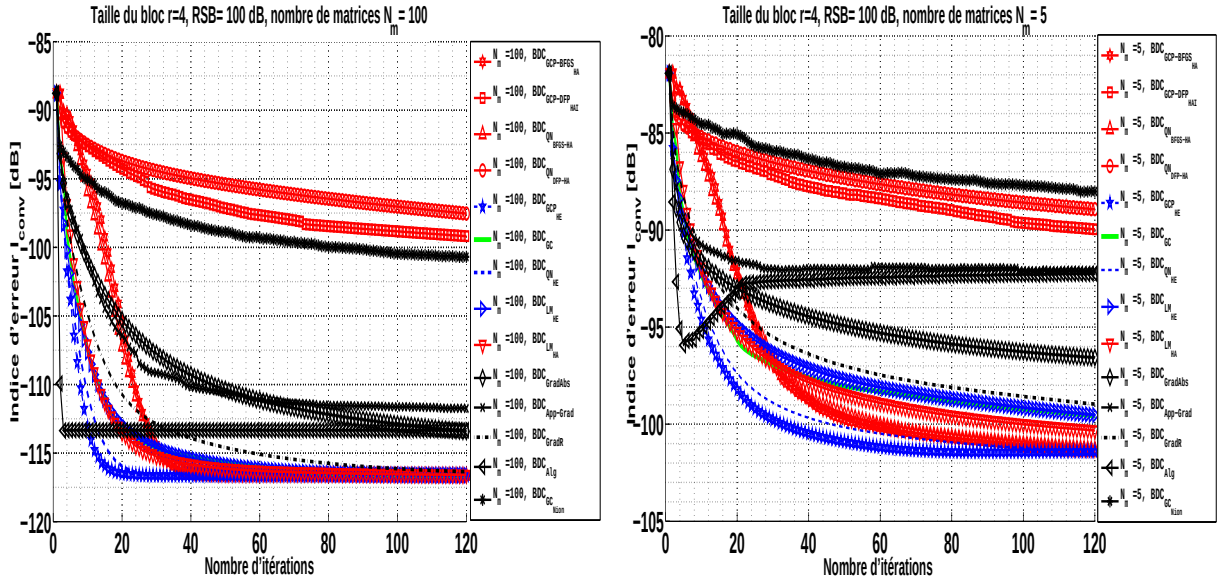


FIGURE 2.1 – Comparaison des algorithmes $BDC_{GCP_{BFGS-HA}}$, $BDC_{GCP_{DFP-HIA}}$, $BDC_{QN_{BFGS-HA}}$, $BDC_{QN_{DFP-HA}}$, $BDC_{GCP_{HE}}$, BDC_{GC} , $BDC_{QN_{HE}}$, $BDC_{LM_{HE}}$, $BDC_{LM_{HA}}$, BDC_{GradR} , $BDC_{App-Grad}$, $BDC_{GradAbs}$, BDC_{Alg} et $BDC_{GC_{Nion}}$ dans un contexte quasi non-bruité ($RSB = 100$ dB) pour le cas d’une matrice $\mathbf{A}$ carrée. Gauche : $I_{Conv}(\mathbf{G})$ en dB en fonction du nombre d’itérations pour un ensemble $\mathcal{M}$ de taille $N_m = 100$ matrices. Droite : $I_{Conv}(\mathbf{G})$ en dB en fonction du nombre d’itérations pour un ensemble $\mathcal{M}$ de taille $N_m = 5$ matrices.

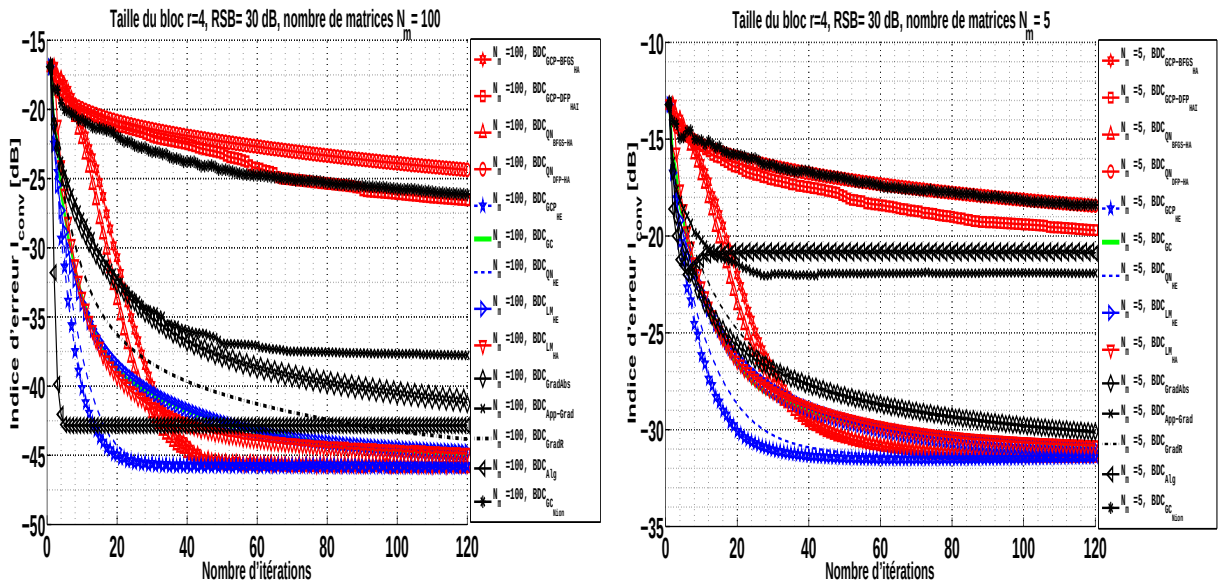


FIGURE 2.2 – Comparaison des algorithmes $BDC_{GCP_{BFGS-HA}}$, $BDC_{GCP_{DFP-HIA}}$, $BDC_{QN_{BFGS-HA}}$, $BDC_{QN_{DFP-HA}}$, $BDC_{GCP_{HE}}$, BDC_{GC} , $BDC_{QN_{HE}}$, $BDC_{LM_{HE}}$, $BDC_{LM_{HA}}$, BDC_{GradR} , $BDC_{App-Grad}$, $BDC_{GradAbs}$, BDC_{Alg} et $BDC_{GC_{Nion}}$ dans un contexte bruité ($RSB = 30$ dB) pour le cas d’une matrice $\mathbf{A}$ carrée. Gauche : $I_{Conv}(\mathbf{G})$ en dB en fonction du nombre d’itérations pour un ensemble $\mathcal{M}$ de taille $N_m = 100$ matrices. Droite : $I_{Conv}(\mathbf{G})$ en dB en fonction du nombre d’itérations pour un ensemble $\mathcal{M}$ de taille $N_m = 5$ matrices.

sur la même règle d'adaptation mais utilisant une autre fonction de coût que la nôtre. Enfin, nous remarquons que l'algorithme algébrique BDC_{Alg} est généralement le plus rapide, mais aussi le moins performant. Les meilleures performances sont généralement atteintes au moyen de l'algorithme $\text{BDC}_{\text{GCP}_{\text{HE}}}$ même en contexte difficile : avec du bruit et avec un ensemble constitué d'un nombre petit de matrices à décomposer conjointement. Nous observons encore que $\text{BDC}_{\text{GCP}_{\text{HE}}}$, $\text{BDC}_{\text{GradAbs}}$, BDC_{GC} , $\text{BDC}_{\text{GCP}_{\text{BFGS-HA}}}$, $\text{BDC}_{\text{QN}_{\text{HE}}}$, $\text{BDC}_{\text{LM}_{\text{HE}}}$, $\text{BDC}_{\text{LM}_{\text{HA}}}$ et $\text{BDC}_{\text{QN}_{\text{BFGS-HA}}}$ convergent approximativement tous vers la même solution : ce qui est logique étant donné qu'ils minimisent tous la même fonction de coût. Cependant, tous les algorithmes (exception faite de BDC_{Alg}) requièrent davantage d'itérations pour atteindre le même niveau de performances.

Nous conservons le même contexte de simulation, sauf que cette fois-ci, nous étudions leur comportement lorsqu'on les initialise plus loin de la solution au moyen d'une matrice aléatoire et non plus par la DVPG. Au niveau de la Fig. 2.3, nous avons à nouveau tracé l'évolution de l'indice d'erreur $\text{I}_{\text{Conv}}(\cdot)$ en dB en fonction du nombre d'itérations c'est à dire dans le cas non bruité où les matrices de l'ensemble $\mathcal{D}$ sont quasiment bloc-diagonales ($\text{RSB} = 100$ dB). Deux cas sont étudiés : à gauche cela correspond au cas où l'ensemble à bloc diagonaliser conjointement est constitué de $N_m = 100$ matrices tandis qu'à droite nous considérons un ensemble de $N_m = 10$ matrices. Au niveau de la Fig. 2.4, nous présentons une nouvelle fois l'évolution de l'indice d'erreur $\text{I}_{\text{Conv}}(\cdot)$ en fonction du nombre d'itérations pour tous les algorithmes lorsque l'on s'écarte de l'hypothèse de bloc-diagonalité (i.e. $\text{RSB} = 30$ dB). À nouveau, pour la courbe de gauche, l'ensemble sur lequel nous travaillons est constitué de $N_m = 100$ matrices tandis qu'à droite nous ne considérons plus que $N_m = 10$ matrices.

D'après l'allure de ces courbes, nous pouvons tirer des conclusions quasi similaires à celles tirées dans le cadre d'une initialisation par la DVPG surtout dans un contexte quasi-bruité. De façon tout à fait logique, nous constatons que tous les algorithmes prennent plus d'itérations pour converger du fait que nous partons d'un point plus éloigné de la solution ($\simeq 0$ dB). De plus, nous remarquons que les algorithmes $\text{BDC}_{\text{GCP}_{\text{DFP-HA}}}$ et $\text{BDC}_{\text{QN}_{\text{DFP-HA}}}$ conduisent à d'assez mauvaises performances ce qui démontre leur sensibilité vis-à-vis de l'initialisation.

Dans un second temps, nous considérons toujours les mêmes conditions que celles utilisées dans le tout premier scénario présenté, par contre nous nous intéressons maintenant au cas d'une matrice $\mathbf{A}$ rectangulaire de taille $M \times N$ ($M = 12$, $N = 8$), la taille du bloc est fixée égale à $r = 4$. Au niveau de la Fig. 2.5, nous comparons, dans le cadre d'une initialisation par DVPG, les différents algorithmes en suivant l'évolution de l'indice d'erreur $\text{I}_{\text{Conv}}(\cdot)$ en dB en fonction du nombre d'itérations. Les matrices de l'ensemble $\mathcal{D}$ sont d'abord considérées comme quasiment bloc-diagonales ($\text{RSB} = 100$ dB). À gauche, nous représentons les résultats obtenus sur un ensemble de $N_m = 100$ matrices tandis qu'à droite nous traçons ces mêmes résultats pour un ensemble constitué de $N_m = 5$ matrices. Au niveau de la Fig. 2.6, nous illustrons l'évolution de l'indice d'erreur $\text{I}_{\text{Conv}}(\cdot)$ en fonction du nombre d'itérations. Les matrices de l'ensemble $\mathcal{D}$ sont maintenant perturbées par un bruit Gaussien pour un $\text{RSB} = 30$ dB. À gauche, nous présentons les résultats obtenus sur un ensemble de $N_m = 100$ matrices tandis qu'à droite, ce sont les résultats obtenus sur un ensemble de $N_m = 5$ matrices.

Ces courbes étayent notre étude : nous observons que les trois algorithmes fondés sur l'utilisation d'un préconditionneur utilisant la diagonale d'une des matrices Hessiennes complexes exactes ($\text{BDC}_{\text{GCP}_{\text{HE}}}$, $\text{BDC}_{\text{QN}_{\text{HE}}}$ et $\text{BDC}_{\text{LM}_{\text{HE}}}$) conduisent à des meilleures performances que les algorithmes basés sur une approximation de celle-ci. Nous vérifions à nouveau que l'algorithme du

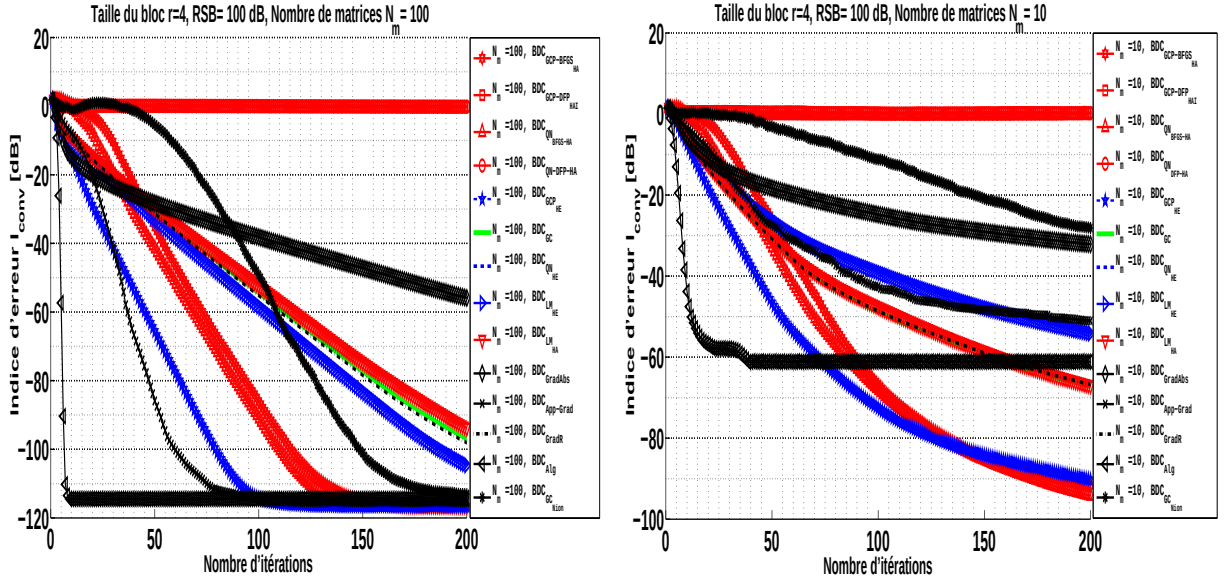


FIGURE 2.3 – Initialisation aléatoire : comparaison des algorithmes $BDC_{GCP-BFGS-HA}$, $BDC_{GCP-DFP-HA}$, $BDC_{QN-BFGS-HA}$, $BDC_{QN-DFP-HA}$, BDC_{GCP-HE} , BDC_{GC} , BDC_{QN-HE} , BDC_{LM-HE} , BDC_{LM-HA} , BDC_{GradR} , $BDC_{App-Grad}$, $BDC_{GradAbs}$, BDC_{Alg} et $BDC_{GC-Nion}$ dans un contexte quasi non-bruité ($RSB = 100$ dB) pour le cas d’une matrice $\mathbf{A}$ carrée. Gauche : $I_{Conv}(\mathbf{G})$ en dB en fonction du nombre d’itérations pour un ensemble $\mathcal{M}$ de taille $N_m = 100$ matrices. Droite : $N_m = 10$ matrices.

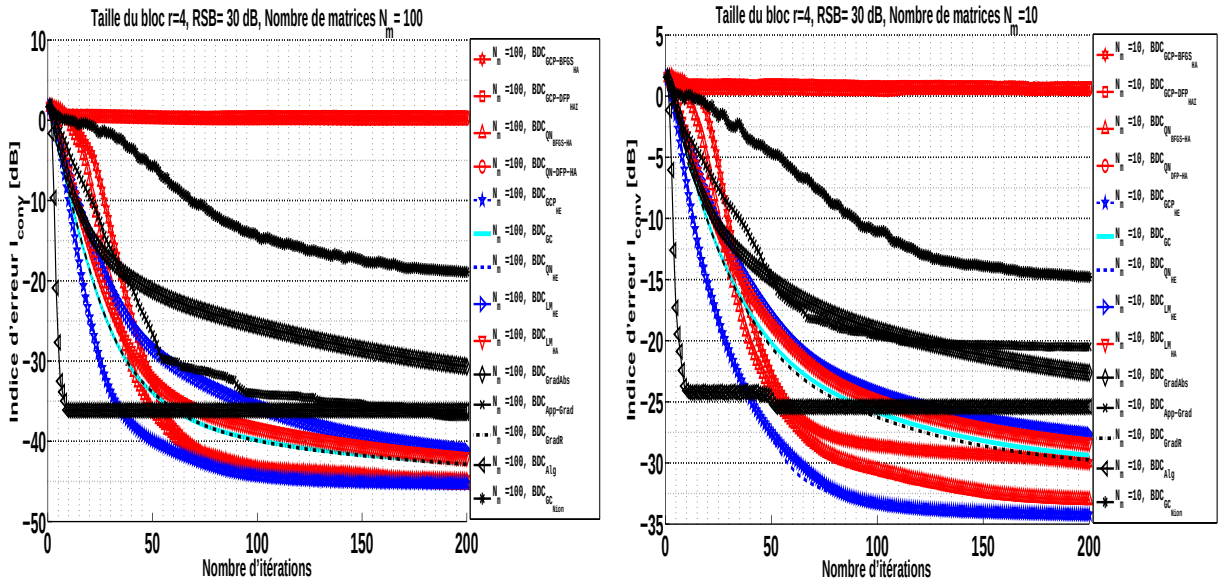


FIGURE 2.4 – Initialisation aléatoire : comparaison des algorithmes $BDC_{GCP-BFGS-HA}$, $BDC_{GCP-DFP-HA}$, $BDC_{QN-BFGS-HA}$, $BDC_{QN-DFP-HA}$, BDC_{GCP-HE} , BDC_{GC} , BDC_{QN-HE} , BDC_{LM-HE} , BDC_{LM-HA} , BDC_{GradR} , $BDC_{App-Grad}$, $BDC_{GradAbs}$, BDC_{Alg} et $BDC_{GC-Nion}$ dans un contexte bruité ($RSB = 30$ dB) pour le cas d’une matrice $\mathbf{A}$ carrée. Gauche : $I_{Conv}(\mathbf{G})$ en dB en fonction du nombre d’itérations pour un ensemble $\mathcal{M}$ de taille $N_m = 100$ matrices. Droite : $N_m = 10$ matrices.

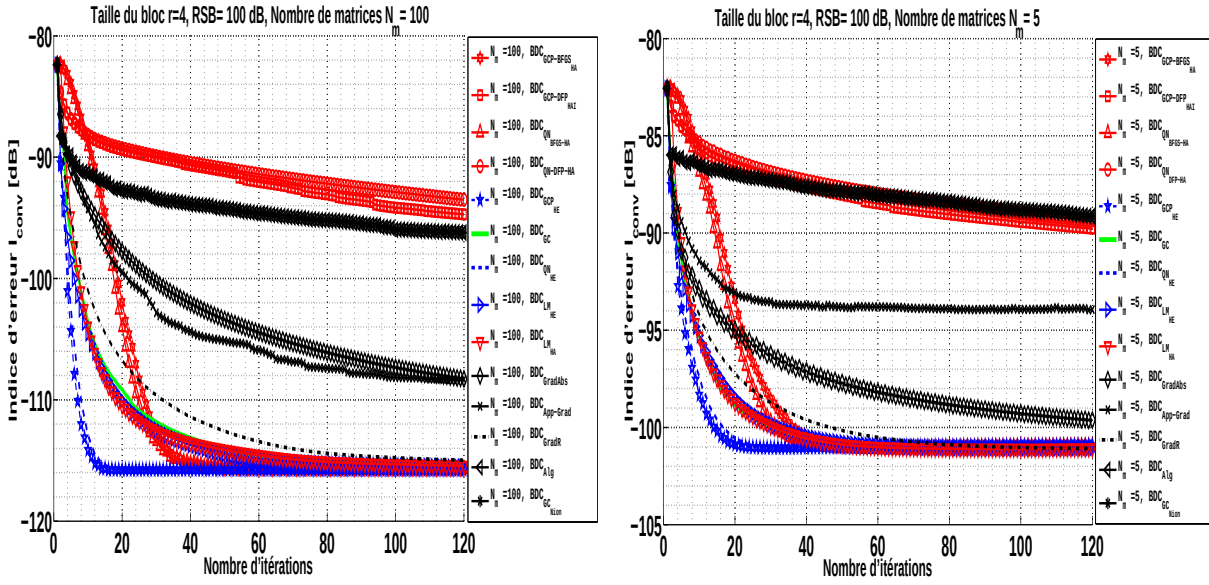


FIGURE 2.5 – Comparaison des algorithmes $\text{BDC}_{\text{GCP-BFGS-HA}}$, $\text{BDC}_{\text{GCP-DFP-HA}}$, $\text{BDC}_{\text{QN-BFGS-HA}}$, $\text{BDC}_{\text{QN-DFP-HA}}$, $\text{BDC}_{\text{GCP-HE}}$, BDC_{GC} , $\text{BDC}_{\text{QN-HE}}$, $\text{BDC}_{\text{LM-HE}}$, $\text{BDC}_{\text{LM-HA}}$, $\text{BDC}_{\text{GradR}}$, $\text{BDC}_{\text{App-Grad}}$, $\text{BDC}_{\text{GradAbs}}$ et $\text{BDC}_{\text{GC-Nion}}$ dans un contexte non-bruité dans le cas d'une matrice $\mathbf{A}$ rectangulaire. Gauche : $I_{\text{conv}}(\mathbf{G})$ en dB en fonction du nombre d'itérations pour un ensemble $\mathcal{M}$ de taille $N_m = 100$ matrices. Droite : $N_m = 5$ matrices.

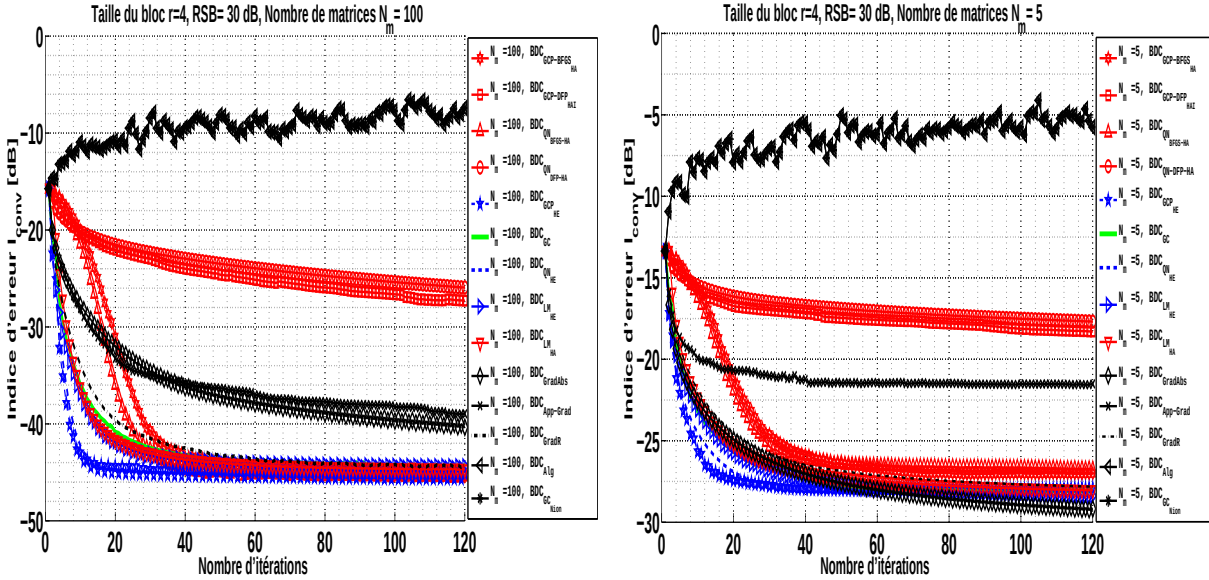


FIGURE 2.6 – Comparaison des algorithmes $\text{BDC}_{\text{GCP-BFGS-HA}}$, $\text{BDC}_{\text{GCP-DFP-HA}}$, $\text{BDC}_{\text{QN-BFGS-HA}}$, $\text{BDC}_{\text{QN-DFP-HA}}$, $\text{BDC}_{\text{GCP-HE}}$, BDC_{GC} , $\text{BDC}_{\text{QN-HE}}$, $\text{BDC}_{\text{LM-HE}}$, $\text{BDC}_{\text{LM-HA}}$, $\text{BDC}_{\text{GradR}}$, $\text{BDC}_{\text{App-Grad}}$, $\text{BDC}_{\text{GradAbs}}$ et $\text{BDC}_{\text{GC-Nion}}$ dans un contexte bruité ($\text{RSB} = 30$ dB) et pour une matrice $\mathbf{A}$ rectangulaire. Gauche : $I_{\text{conv}}(\mathbf{G})$ en dB en fonction du nombre d'itérations pour un ensemble $\mathcal{M}$ de taille $N_m = 100$ matrices. Droite : $N_m = 5$ matrices.

gradient conjugué préconditionné $\text{BDC}_{\text{GCPHE}}$ semble bien être le plus robuste vis-à-vis d'erreurs de modèles (écart par rapport à la bloc diagonalité) et d'une réduction du nombre de matrices à bloc-diagonaliser conjointement. Nous constatons également un gain dans les performances et la stabilité au niveau de l'algorithme à base du gradient conjugué (BDC_{GC}) utilisant la fonction de coût directe plutôt que celle utilisée par l'algorithme de D. Nion ($\text{BDC}_{\text{GC}_{\text{Nion}}}$). Nous observons encore (comme sur les exemples précédents) que les performances dépendent bien de la taille N_m de l'ensemble des matrices à bloc-diagonaliser. En augmentant cette taille N_m , nous améliorons les performances des algorithmes de BDC proposés.

2.3.1.1 Etude des performances en fonction du nombre de matrices N_m , du RSB et de la taille d'un bloc r

Afin de continuer à approfondir l'étude du comportement des algorithmes proposés, nous présentons d'autres simulations numériques consistant à étudier cette fois-ci l'évolution des performances en fonction de trois paramètres à savoir : i) le nombre N_m des matrices à bloc-diagonaliser, ii) le rapport signal à bruit qui correspond à l'écart vis-à-vis de l'hypothèse de bloc-diagonalité iii) et la taille d'un bloc r .

Nous considérons $M = N = 8$ (la matrice $\mathbf{A}$ est carrée complexe, sa partie réelle et sa partie imaginaire sont générées aléatoirement selon une loi uniforme sur l'intervalle $[-1, 1]$). L'ensemble $\mathcal{M}$ est constitué d'un nombre N_m variable de matrices allant de 2 à 100 matrices complexes de taille $M \times M$. Nous évaluons également le RSB qui vaut respectivement 25, 100 dB et la taille d'un bloc r pourra varier entre 1, 2 et 4.

Dans un premier temps, nous commençons par le comportement des algorithmes proposés en fonction du nombre de matrices N_m à bloc-diagonaliser conjointement, puis en fonction du rapport signal à bruit en fixant cette fois-ci la taille des blocs r . Ces algorithmes sont initialisés au moyen la méthode DVPG présentée dans [184]. Pour une taille de bloc $r = 4$ (resp. 2) donnée, l'indice d'erreur $\text{I}_{\text{Conv}}(\cdot)$ en dB est représenté au niveau de la Fig. 2.7 (resp. 2.9) en fonction de la taille N_m de l'ensemble de matrices considéré pour deux valeurs de RSB : 25 dB (resp. 30 dB) (à gauche) et 100 dB (resp. 100 dB) (à droite).

Pour les mêmes tailles de bloc, l'indice d'erreur $\text{I}_{\text{Conv}}(\cdot)$ en dB est représenté au niveau de la Fig. 2.8 et de la Fig. 2.10 en fonction du rapport signal à bruit pour deux tailles d'ensemble $\mathcal{M}$ de matrices à bloc diagonaliser conjointement à savoir $N_m = 5$ (à gauche) et $N_m = 100$ matrices (à droite).

Nous constatons que plus le RSB est favorable, meilleurs sont les résultats. Sur des ensembles de matrices quasi bloc diagonales i.e. RSB = 100 dB, nous convergions vers $I \approx -120$ dB : nous estimons donc parfaitement le bloc-diagonaliseur conjoint commun à toutes les matrices de l'ensemble $\mathcal{M}$. Nous observons également que plus que le nombre N_m de matrices à bloc-diagonaliser conjointement est grand, meilleurs sont les résultats. Finalement, de la même manière que sur les simulations précédentes, les algorithmes fondés sur des matrices Hessiennes exactes sont les plus stables et les plus adaptés à des situations défavorables dans lesquelles nous disposons de peu de matrices au niveau de l'ensemble $\mathcal{M}$ à décomposer conjointement, ces matrices s'écartant de plus de l'hypothèse de bloc-diagonalité.

Étudions maintenant le comportement des algorithmes proposés en fonction de la taille d'un bloc r dans un contexte quasi non-bruité (resp. bruité (RSB = 30 dB)). Les résultats obtenus sont

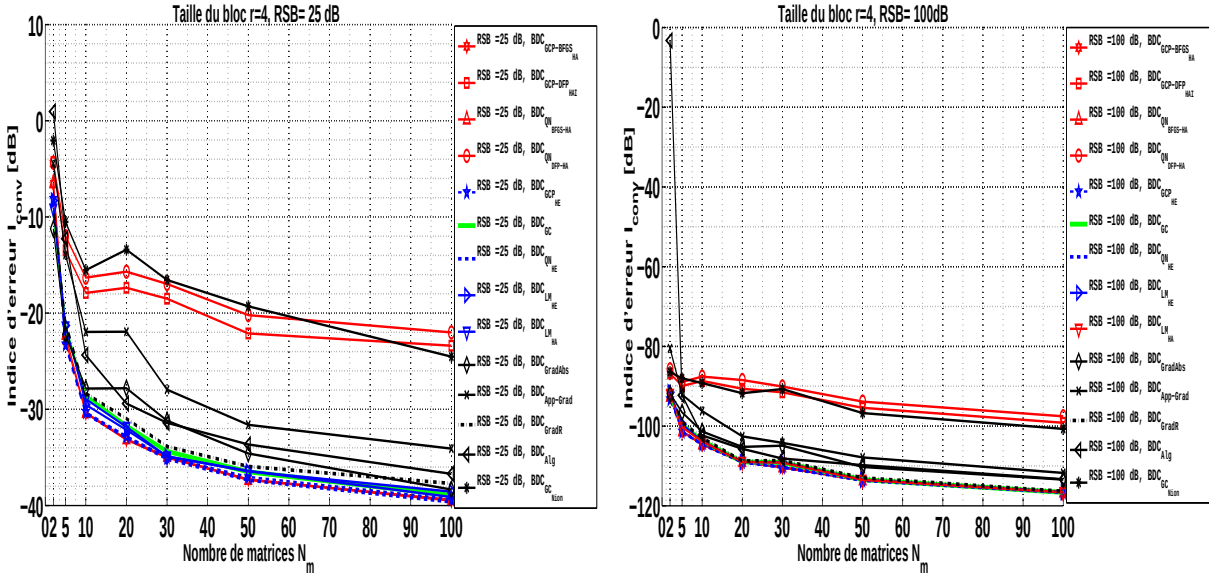


FIGURE 2.7 – Comparaison des algorithmes ($BDC_{GCP_{BFGS-HA}}$, $BDC_{GCP_{DFP-HIA}}$, $BDC_{QN_{BFGS-HA}}$, $BDC_{QN_{DFP-HA}}$, $BDC_{GCP_{HE}}$, BDC_{GC} , $BDC_{QN_{HE}}$, $BDC_{LM_{HE}}$, $BDC_{LM_{HA}}$, BDC_{GradR} , $BDC_{App-Grad}$, $BDC_{GradAbs}$, BDC_{Alg} et $BDC_{GC_{Nion}}$). À gauche : $RSB = 25$ dB. À droite : $RSB = 100$ dB.

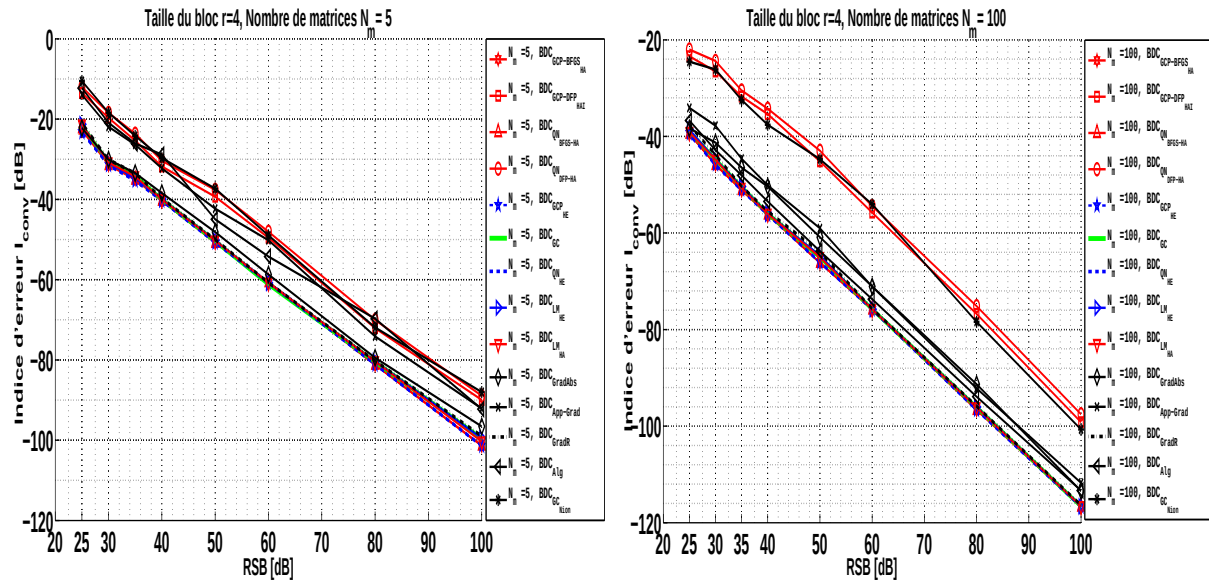


FIGURE 2.8 – Taille d'un bloc $r = 4$: performance en fonction du rapport signal à bruit des algorithmes $BDC_{GCP_{BFGS-HA}}$, $BDC_{GCP_{DFP-HIA}}$, $BDC_{QN_{BFGS-HA}}$, $BDC_{QN_{DFP-HA}}$, $BDC_{GCP_{HE}}$, BDC_{GC} , $BDC_{QN_{HE}}$, $BDC_{LM_{HE}}$, $BDC_{LM_{HA}}$, BDC_{GradR} , $BDC_{App-Grad}$, $BDC_{GradAbs}$, BDC_{Alg} et $BDC_{GC_{Nion}}$. Gauche : $N_m = 5$ matrices. Droite : $N_m = 100$ matrices.

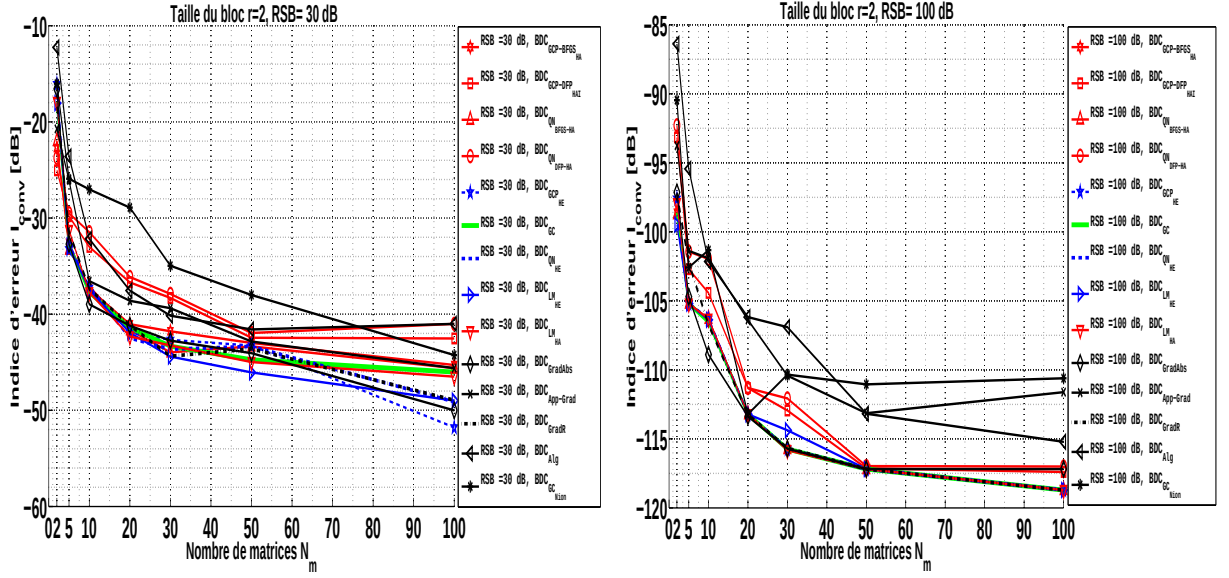


FIGURE 2.9 – Pour une taille de bloc fixée à $r = 2$, performances en fonction du nombre de matrices N_m pour les différents algorithmes ($BDC_{GCP-BFGS-HA}$, $BDC_{GCP-DFP-HIA}$, $BDC_{QN-BFGS-HA}$, $BDC_{QN-DFP-HA}$, BDC_{GCP-HE} , BDC_{GC} , BDC_{QN-HE} , BDC_{LM-HE} , BDC_{LM-HA} , BDC_{GradR} , $BDC_{App-Grad}$, $BDC_{GradAbs}$, BDC_{Alg} et $BDC_{GC-Nion}$). A gauche : $RSB = 30$ dB. A droite : $RSB = 100$ dB.

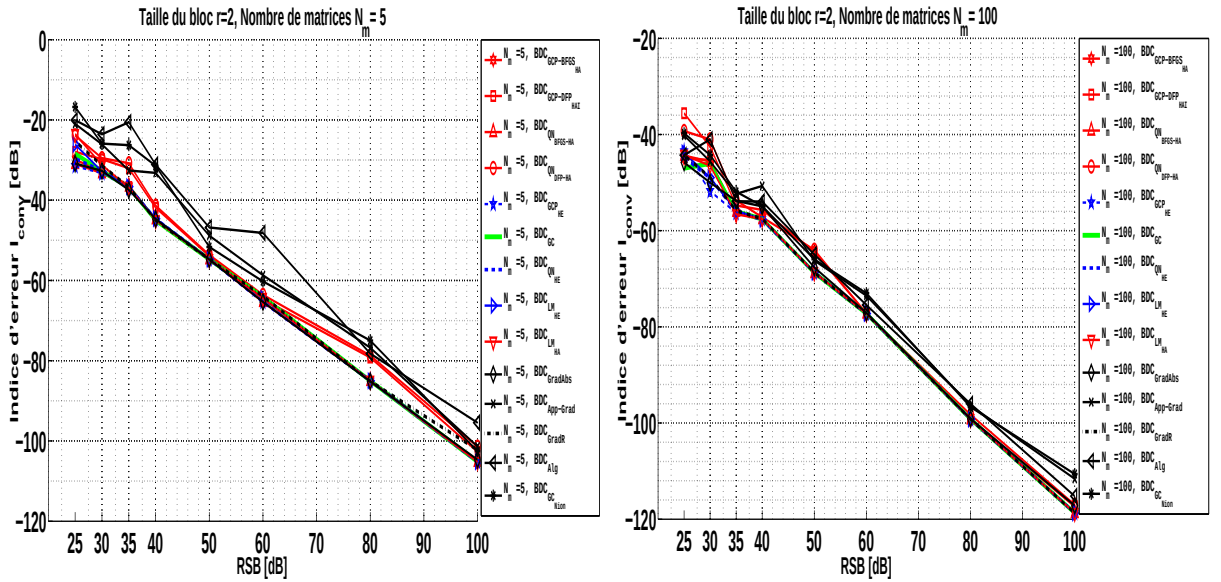


FIGURE 2.10 – Pour une taille de bloc fixée à $r = 2$, performances en fonction du rapport signal à bruit pour les algorithmes $BDC_{GCP-BFGS-HA}$, $BDC_{GCP-DFP-HIA}$, $BDC_{QN-BFGS-HA}$, $BDC_{QN-DFP-HA}$, BDC_{GCP-HE} , BDC_{GC} , BDC_{QN-HE} , BDC_{LM-HE} , BDC_{LM-HA} , BDC_{GradR} , $BDC_{App-Grad}$, $BDC_{GradAbs}$, BDC_{Alg} et $BDC_{GC-Nion}$. À gauche : $N_m = 5$ matrices. À droite : $N_m = 100$ matrices.

présentés au niveau de la Fig. 2.11 (resp. Fig. 2.12). La taille N_m de l'ensemble $\mathcal{M}$ et le RSB sont alors fixés. Sur ces figures, nous constatons que les performances des algorithmes de BDC proposés dépendent de la taille d'un bloc r considéré ou bien du nombre de blocs $n_j = \frac{N}{r}$ ³¹. Nous constatons de plus que les résultats obtenus auparavant restent valables même si l'on change la taille d'un bloc : l'algorithme $\text{BDC}_{\text{GCP}_{\text{HE}}}$ est toujours celui qui conduit aux meilleures performances, suivi des algorithmes $\text{BDC}_{\text{LM}_{\text{HE}}}$ et $\text{BDC}_{\text{QN}_{\text{HE}}}$. Les performances sont d'autant meilleures que l'on diminue la valeur de la taille d'un bloc r . Notons que dans le cas particulier où $r = 1$, nous retombons alors le problème de la diagonalisation conjointe d'un ensemble de matrices. Comme nous l'avons déjà mentionné précédemment, les quatre algorithmes de BDC proposés peuvent aussi servir à résoudre le problème de la DC. Autrement dit, les algorithmes $\text{DC}_{\text{GCP}_{\text{HE}}}$, DC_{GC} , $\text{DC}_{\text{QN}_{\text{HE}}}$, $\text{DC}_{\text{LM}_{\text{HE}}}$ constituent alors quatre nouvelles solutions au problème de la diagonalisation conjointe non unitaire.

2.3.2 Cas particulier : la DC d'un ensemble de matrices complexes

Nous présentons un autre exemple illustrant l'intérêt des algorithmes de BDC proposés et leur efficacité à résoudre un sous-problème de la BDC : celui de la DC. En nous plaçant dans le même contexte de simulations, nous évaluons l'évolution de l'indice d'erreur $\text{l}_{\text{Conv}}(\cdot)$ ³² en dB. Les résultats obtenus sont donnés au niveau de la Fig. 2.13 en fonction de la taille N_m de l'ensemble de matrices considéré pour deux valeurs de RSB = 30 dB (à gauche) et RSB = 100 dB (à droite). Quand $r = 1$, il est mentionné dans [114] que l'algorithme BDC_{Alg} équivaut alors à l'algorithme DC_{Alg} présenté dans [92] et l'algorithme $\text{BDC}_{\text{App-Grad}}$ équivaut alors à l'algorithme DOMUNG proposé dans [253].

L'évolution de l'indice d'erreur $\text{l}_{\text{Conv}}(\cdot)$ en dB est tracée au niveau de la Fig. 2.14 en fonction du rapport signal à bruit pour deux tailles d'ensembles considérés : $N_m = 5$ matrices (à gauche) et $N_m = 100$ matrices (à droite).

Nous tirons exactement les mêmes conclusions dans le cas particulier de la DC que celles tirées dans le cadre de la BDC_{NU} .

2.3.3 Cas général de la ZBDC d'un ensemble de matrices complexes

Nous étudions le comportement des nouveaux algorithmes de décompositions matricielles introduits au niveau de la section §2.2 lesquels sont fondés sur différents schémas d'optimisation itératifs de type :

- le gradient conjugué (ZBDC_{GC}),
- le gradient conjugué préconditionné ($\text{ZBDC}_{\text{GCP}_{\text{HE}}}$),
- l'algorithme de Levenberg-Marquardt ($\text{ZBDC}_{\text{LM}_{\text{HE}}}$).

Nous considérons deux types de scénario selon le type d'initialisation utilisé. Tout d'abord, nous commençons par une initialisation à partir de la DVPG [184] puis nous partons d'une initialisation aléatoire. Le cas d'une matrice $\mathbf{A}$ rectangulaire est également étudié.

Dans un premier temps, nous fixons $M = N = 8$ (la matrice $\mathbf{A}$ est carrée) et $r = 4$. La matrice $\mathbf{A}$ est choisie de manière aléatoire selon une loi uniforme de moyenne nulle et de variance un sur l'intervalle $[-1, 1]$. L'ensemble $\mathcal{N}$ des matrices considérées $\mathbf{X}_i$, $i = 1, \dots, N_m$ est constitué de

31. En séparation de sources, le nombre de blocs n_j représente le nombre de signaux sources n .

32. Lorsque $r = 1$, $\text{l}_{\text{Conv}}(\cdot)$ n'est rien d'autre que l'indice d'erreur proposé dans [171].

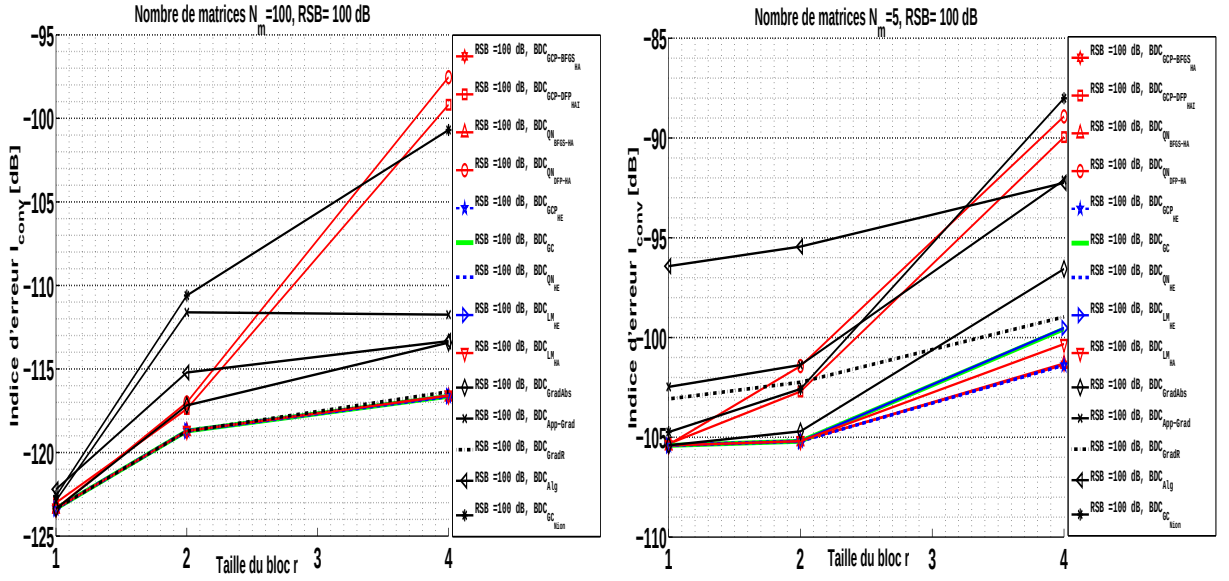


FIGURE 2.11 – Comparaison des algorithmes $BDC_{GCP_{BFGS-HA}}$, $BDC_{GCP_{DFP-HIA}}$, $BDC_{QN_{BFGS-HA}}$, $BDC_{QN_{DFP-HA}}$, $BDC_{GCP_{HE}}$, BDC_{GC} , $BDC_{QN_{HE}}$, $BDC_{LM_{HE}}$, $BDC_{LM_{HA}}$, BDC_{GradR} , $BDC_{App-Grad}$, $BDC_{GradAbs}$, BDC_{Alg} et $BDC_{GC_{Nion}}$ dans un contexte quasi non-bruité. Gauche : $I_{Conv}(G)$ en dB en fonction de la taille d'un bloc r pour un ensemble $\mathcal{M}$ de taille $N_m = 100$ matrices. Droite : $N_m = 5$ matrices.

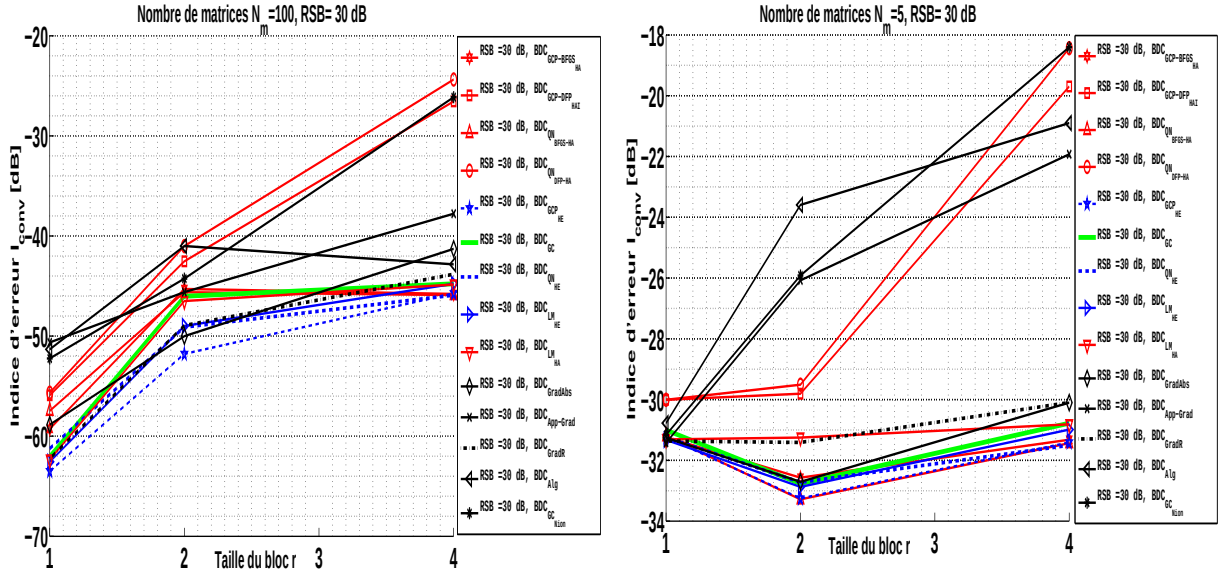


FIGURE 2.12 – Comparaison des algorithmes $BDC_{GCP_{BFGS-HA}}$, $BDC_{GCP_{DFP-HIA}}$, $BDC_{QN_{BFGS-HA}}$, $BDC_{QN_{DFP-HA}}$, $BDC_{GCP_{HE}}$, BDC_{GC} , $BDC_{QN_{HE}}$, $BDC_{LM_{HE}}$, $BDC_{LM_{HA}}$, BDC_{GradR} , $BDC_{App-Grad}$, $BDC_{GradAbs}$, BDC_{Alg} et $BDC_{GC_{Nion}}$ dans un contexte bruité ($RSB = 30$ dB). Gauche : $I_{Conv}(G)$ en dB en fonction de la taille d'un bloc r pour un ensemble $\mathcal{M}$ de taille $N_m = 100$ matrices. Droite : $N_m = 5$ matrices.

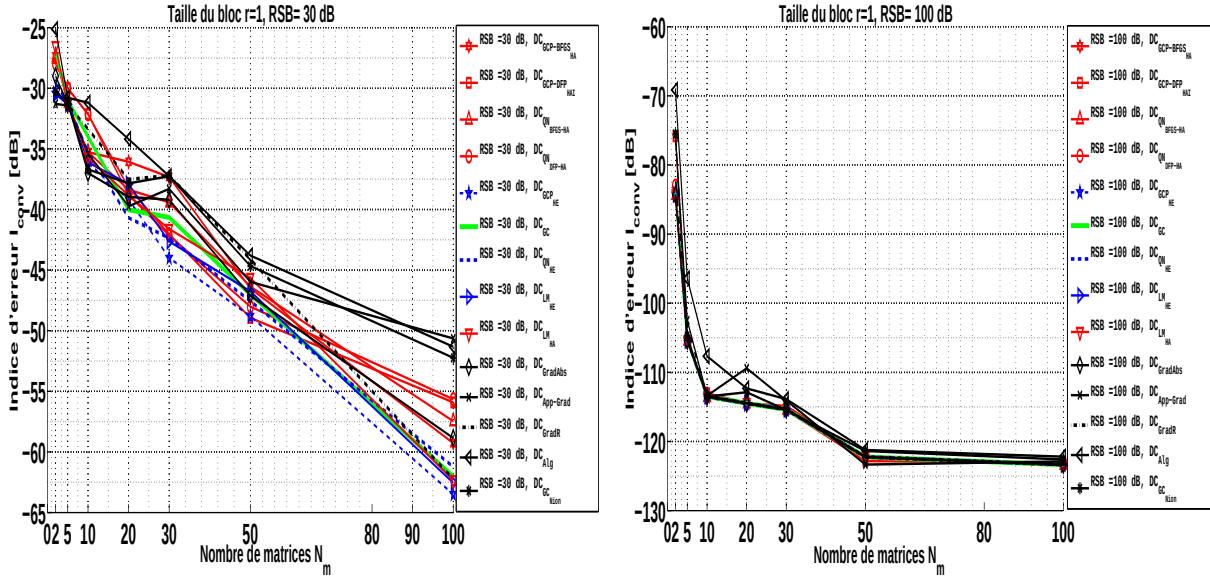


FIGURE 2.13 – Diagonalisation conjointe ($r = 1$) : performances en fonction du nombre de matrices N_m pour les algorithmes $DC_{GCP-BFGS-HA}$, $DC_{GCP-DFP-HIA}$, $DC_{QN-BFGS-HA}$, $DC_{QN-DFP-HA}$, DC_{GCP-HE} , DC_{GC} , DC_{QN-HE} , DC_{LM-HE} , DC_{LM-HA} , DC_{GradR} , $DC_{App-Grad}$, $DC_{GradAbs}$, DC_{Alg} et $DC_{GC-Nion}$. À gauche : $RSB = 30$ dB. À droite : $RSB = 100$ dB.

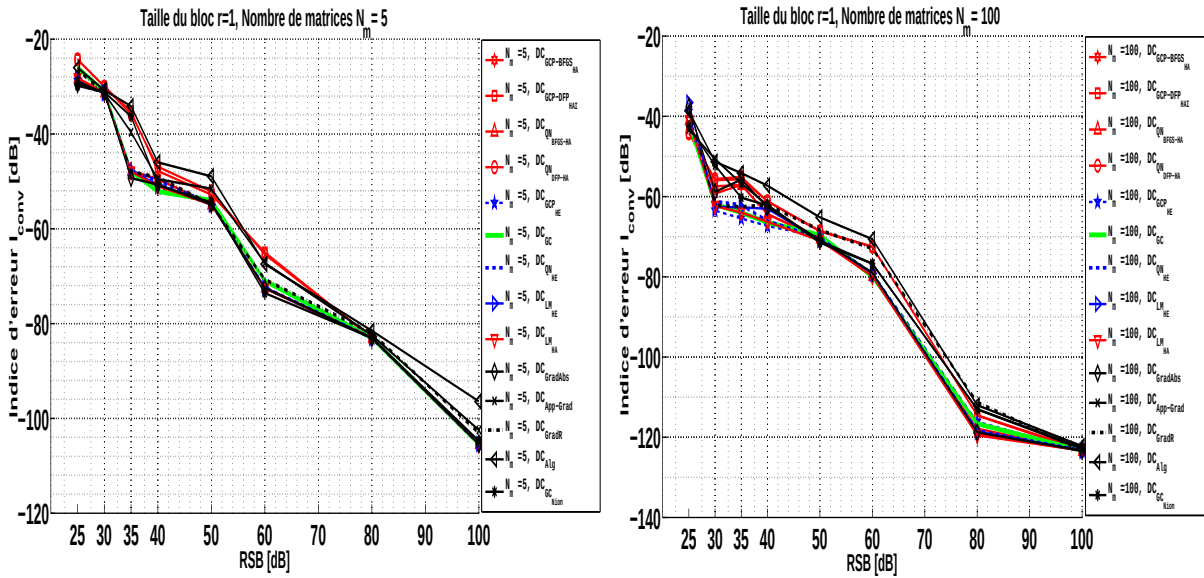


FIGURE 2.14 – Diagonalisation conjointe ($r = 1$) : performances en fonction du rapport signal à bruit pour les algorithmes $DC_{GCP-BFGS-HA}$, $DC_{GCP-DFP-HIA}$, $DC_{QN-BFGS-HA}$, $DC_{QN-DFP-HA}$, DC_{GCP-HE} , DC_{GC} , DC_{QN-HE} , DC_{LM-HE} , DC_{LM-HA} , DC_{GradR} , $DC_{App-Grad}$, $DC_{GradAbs}$, DC_{Alg} et $DC_{GC-Nion}$. À gauche : $N_m = 5$ matrices. À droite : $N_m = 100$ matrices.

$N_m = 100$ (resp. 5) matrices complexes de taille $M \times M$. Les trois algorithmes sont initialisés de la même manière par la DVPG.

Au niveau de la Fig. 2.15, nous présentons l'évolution de l'indice d'erreur $l_{\text{Conv}}(\cdot)$ en dB en fonction du nombre d'itérations dans le cas non bruité où les matrices de l'ensemble $\mathcal{N}$ sont quasiment zéro-bloc diagonales ($\text{RSB} = 100$ dB). Deux cas sont étudiés : pour la courbe de gauche l'ensemble à zéro-bloc diagonaliser conjointement est constitué de $N_m = 100$ matrices tandis que, pour la courbe de droite, il n'est plus constitué que de $N_m = 5$ matrices. Au niveau de la Fig. 2.16, nous présentons l'évolution de l'indice d'erreur $l_{\text{Conv}}(\cdot)$ en fonction du nombre d'itérations pour les trois algorithmes lorsque l'on introduit des erreurs de modèle et que l'on s'écarte de l'hypothèse de zéro-bloc diagonalité ($\text{RSB} = 30$ dB). Pour la courbe de gauche, nous travaillons sur un ensemble de $N_m = 100$ matrices tandis que, pour la courbe de droite, nous ne considérons plus que $N_m = 5$ matrices.

Sur ces courbes, nous vérifions toujours que la vitesse de convergence des algorithmes utilisant un préconditionneur ($\text{ZBDC}_{\text{LM}_{\text{HE}}}$ et $\text{ZBDC}_{\text{GCP}_{\text{HE}}}$) est plus rapide que celle d'approches non préconditionnées de type gradient conjugué (ZBDC_{GC}). Nous remarquons aussi que l'algorithme $\text{ZBDC}_{\text{GCP}_{\text{HE}}}$ est généralement plus rapide et plus performant que les autres (relative robustesse vis-à-vis d'erreurs de modèle et d'une diminution de la taille de l'ensemble des matrices à zéro-bloc diagonaliser conjointement). Nous remarquons encore que $\text{ZBDC}_{\text{GCP}_{\text{HE}}}$ et $\text{ZBDC}_{\text{LM}_{\text{HE}}}$ convergent quasiment vers la même solution. Cependant, l'algorithme ZBDC_{GC} requiert plus d'itérations pour atteindre les mêmes performances.

En conservant les mêmes paramètres au niveau de nos simulations, nous étudions maintenant, au niveau de la Fig. 2.17 le comportement de ces algorithmes lorsqu'on les initialise plus loin de la solution c'est-à-dire à partir d'une matrice choisie de manière aléatoire. Nous traçons donc l'indice d'erreur $l_{\text{Conv}}(\cdot)$ en dB en fonction du nombre d'itérations dans le cas non bruité où les matrices de l'ensemble $\mathcal{N}$ sont quasiment zéro bloc-diagonales ($\text{RSB} = 100$ dB). En ce qui concerne la courbe de gauche, l'ensemble à zéro-bloc diagonaliser conjointement est constitué de $N_m = 100$ matrices tandis que pour la courbe de droite, nous ne considérons plus que $N_m = 20$ matrices. Au niveau de la Fig. 2.18, nous présentons de nouveau l'évolution de l'indice d'erreur $l_{\text{Conv}}(\cdot)$ en fonction du nombre d'itérations pour les trois algorithmes lorsque l'on s'écarte de l'hypothèse de zéro-bloc diagonalité ($\text{RSB} = 30$ dB) : la courbe de gauche correspond à un ensemble constitué de $N_m = 100$ matrices, tandis que celle de droite correspond à un ensemble de $N_m = 20$ matrices.

Dans un second temps, nous considérons toujours les mêmes conditions que celles de l'exemple précédent, par contre nous étudions maintenant le cas d'une matrice $\mathbf{A}$ rectangulaire de taille $M \times N$ ($M = 12$, $N = 8$), la taille d'un bloc est cette fois-ci prise égale à $r = 4$. Dans le cadre d'une initialisation par la DVPG, nous comparons, au niveau de la Fig. 2.19, les résultats obtenus au moyen des différents algorithmes en fonction du nombre d'itérations. Les matrices de l'ensemble $\mathcal{Z}$ sont considérées quasi zéro-bloc diagonales ($\text{RSB} = 100$ dB). À gauche, nous présentons les résultats obtenus sur un ensemble de $N_m = 100$ matrices tandis qu'à droite nous fournissons les mêmes résultats pour un ensemble constitué de $N_m = 10$ matrices. Au niveau de la Fig. 2.20, nous illustrons l'évolution de l'indice d'erreur $l_{\text{Conv}}(\cdot)$ en fonction du nombre d'itérations. Les matrices de l'ensemble $\mathcal{Z}$ sont maintenant perturbées par un bruit Gaussien et le $\text{RSB} = 25$ dB. À gauche, nous donnons les résultats obtenus sur un ensemble de $N_m = 100$ matrices tandis qu'à droite, ce sont les résultats obtenus sur un ensemble de $N_m = 10$ matrices.

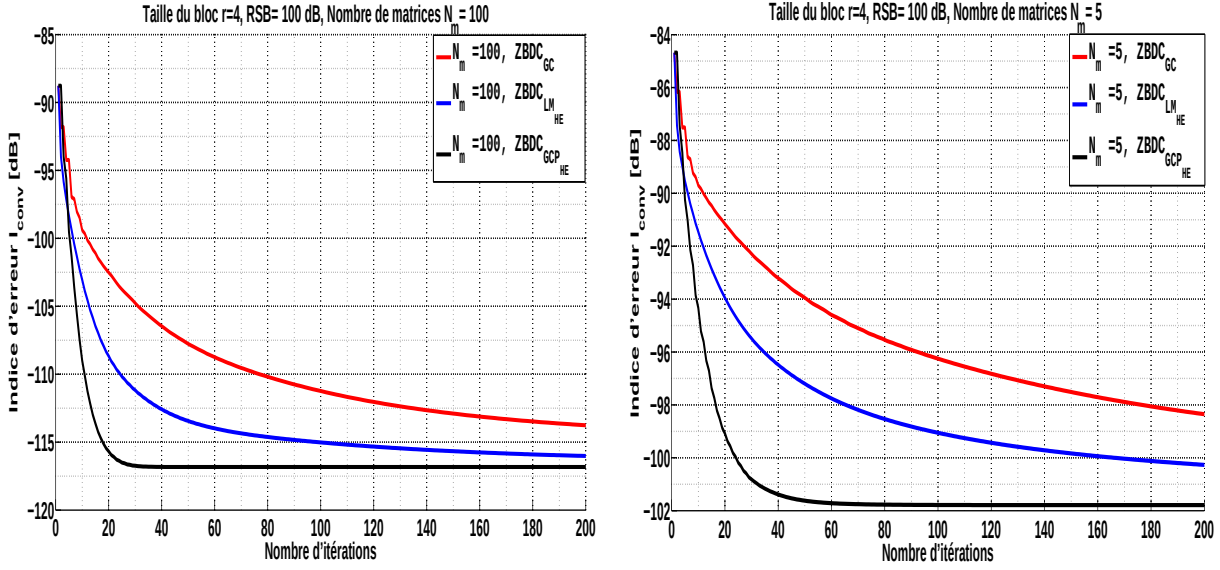


FIGURE 2.15 – Comparaison des algorithmes ZBDC_{GC}, ZBDC_{LM_{HE}}, ZBDC_{GCP_{HE}} dans un contexte quasi non-bruité (RSB = 100 dB) pour le cas d’une matrice $\mathbf{A}$ carrée. Gauche : $I_{\text{Conv}}(\mathbf{G})$ en dB en fonction du nombre d’itérations pour un ensemble $\mathcal{N}$ de taille $N_m = 100$ matrices. Droite : $I_{\text{Conv}}(\mathbf{G})$ en dB en fonction du nombre d’itérations pour un ensemble $\mathcal{N}$ de taille $N_m = 5$ matrices à droite.

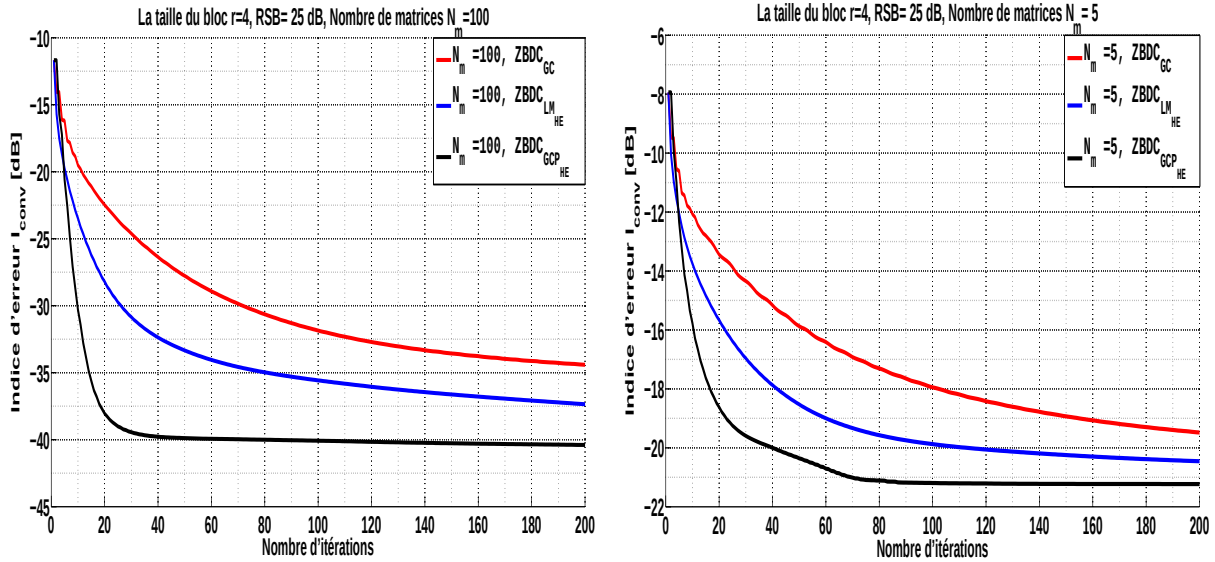


FIGURE 2.16 – Comparaison des algorithmes ZBDC_{GC}, ZBDC_{LM_{HE}}, ZBDC_{GCP_{HE}} dans un contexte bruité (RSB = 30 dB) pour le cas d’une matrice $\mathbf{A}$ carrée. Gauche : $I_{\text{Conv}}(\mathbf{G})$ en dB en fonction du nombre d’itérations pour un ensemble $\mathcal{N}$ de taille $N_m = 100$ matrices. Droite : $I_{\text{Conv}}(\mathbf{G})$ en dB en fonction du nombre d’itérations pour un ensemble $\mathcal{N}$ de taille $N_m = 5$ matrices.

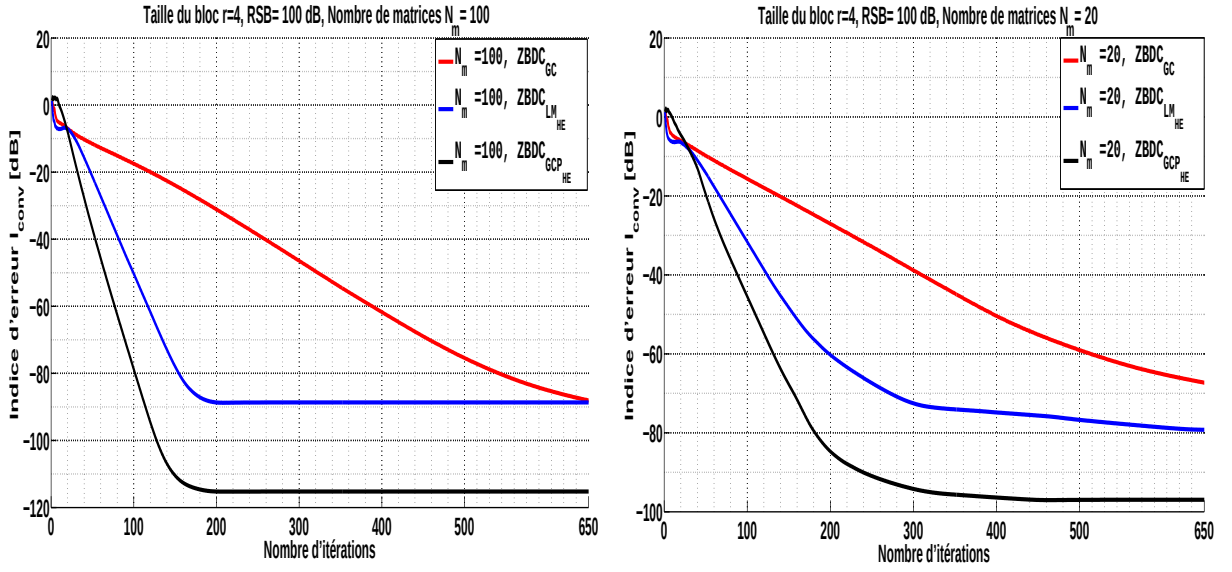


FIGURE 2.17 – Initialisation aléatoire : comparaison des algorithmes ZBDC_{GC}, ZBDC_{LM_{HE}}, ZBDC_{GCP_{HE}} dans un contexte non-bruité (RSB = 100 dB) pour le cas d’une matrice $\mathbf{A}$ carrée. Gauche : $I_{\text{Conv}}(\mathbf{G})$ en dB en fonction du nombre d’itérations pour un ensemble $\mathcal{N}$ de taille $N_m = 100$ matrices . Droite : $N_m = 20$ matrices.

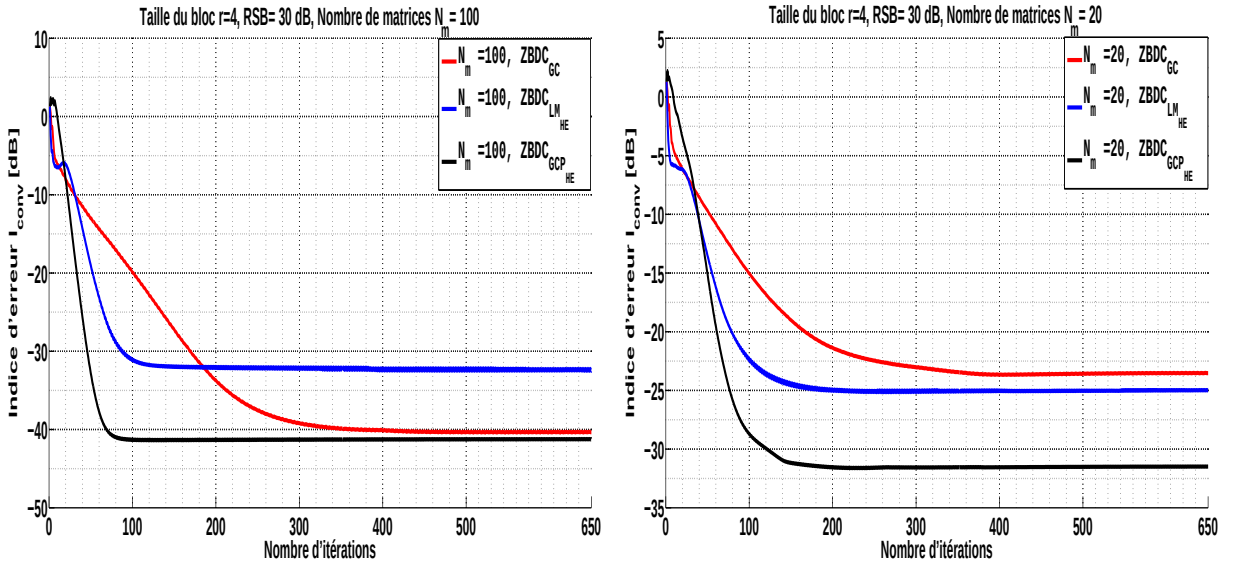


FIGURE 2.18 – Initialisation aléatoire : comparaison des algorithmes ZBDC_{GC}, ZBDC_{LM_{HE}}, ZBDC_{GCP_{HE}} dans un contexte bruité (RSB = 30 dB) pour le cas d’une matrice $\mathbf{A}$ carrée. Gauche : $I_{\text{Conv}}(\mathbf{G})$ en dB en fonction du nombre d’itérations pour un ensemble $\mathcal{M}$ de taille $N_m = 100$ matrices. Droite : $N_m = 20$ matrices.

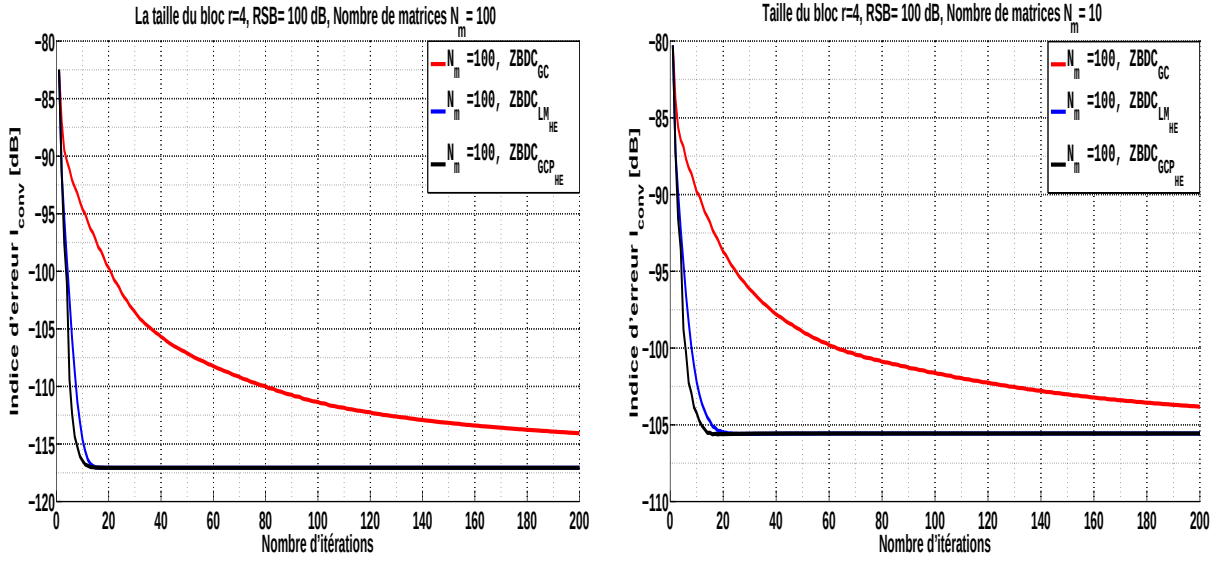


FIGURE 2.19 – Comparaison des algorithmes ZBDC_{GC} , $\text{ZBDC}_{\text{LM}_{\text{HE}}}$ et $\text{ZBDC}_{\text{GCP}_{\text{HE}}}$ dans un contexte non-bruité dans le cas d’une matrice $\mathbf{A}$ rectangulaire. Gauche : $I_{\text{Conv}}(\mathbf{G})$ en dB en fonction du nombre d’itérations pour un ensemble $\mathcal{M}$ de taille $N_m = 100$ matrices. Droite : $N_m = 10$ matrices.

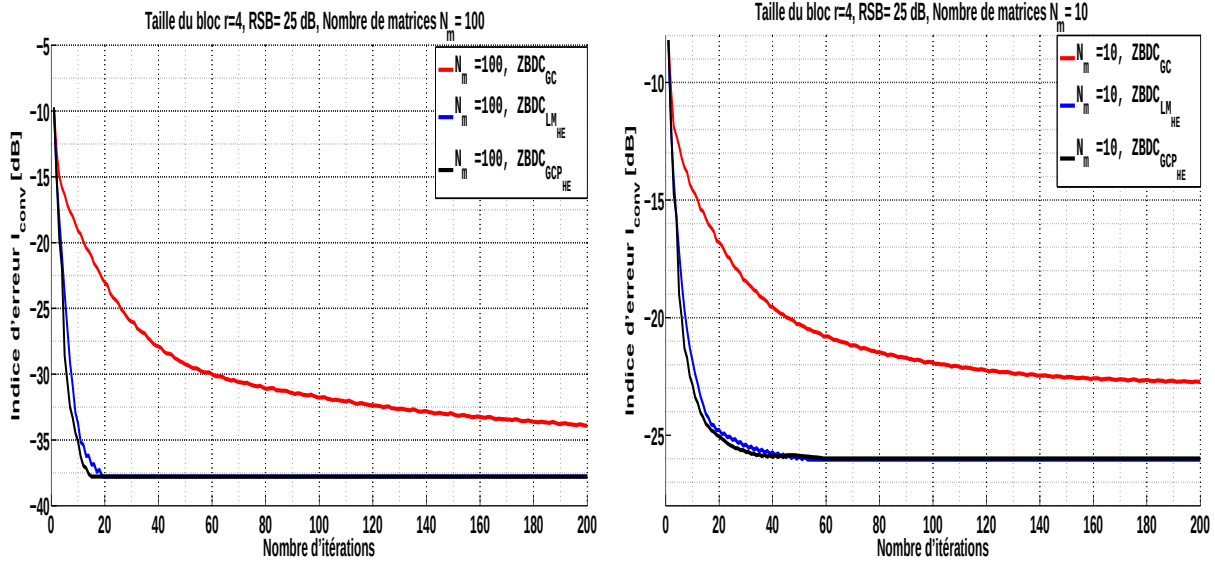


FIGURE 2.20 – Comparaison des algorithmes ZBDC_{GC} , $\text{ZBDC}_{\text{LM}_{\text{HE}}}$ et $\text{ZBDC}_{\text{GCP}_{\text{HE}}}$ dans un contexte bruité ($\text{RSB} = 25$ dB) et pour une matrice $\mathbf{A}$ rectangulaire. Gauche : $I_{\text{Conv}}(\mathbf{G})$ en dB en fonction du nombre d’itérations pour un ensemble $\mathcal{M}$ de taille $N_m = 100$ matrices. Droite : $N_m = 10$ matrices.

Ces graphes semblent indiquer que les algorithmes fondés sur un préconditionneur reposant sur un calcul exact des matrices Hessiennes complexes ($\text{ZBDC}_{\text{GCP}_{\text{HE}}}$ et $\text{ZBDC}_{\text{LM}_{\text{HE}}}$) conduisent à des meilleures performances qu'un algorithme non préconditionné tel que celui du gradient conjugué présenté ici. Ce dernier prend plus d'itérations pour atteindre les mêmes performances. Nous remarquons à nouveau que $\text{ZBDC}_{\text{GCP}_{\text{HE}}}$ est bien le plus robuste vis-à-vis d'un écart par rapport à la zéro-bloc diagonalité et d'une réduction du nombre de matrices à zéro-bloc diagonaliser conjointement. Nous constatons enfin (comme sur les exemples précédents) que les performances dépendent bien de la taille N_m de l'ensemble des matrices à zéro-bloc diagonaliser. En augmentant la valeur de N_m , nous améliorons les performances des algorithmes de ZBDC proposés.

2.3.3.1 Etude des performances en fonction du nombre de matrices N_m , du RSB et de la taille d'un bloc r

Dans un souci d'approfondissement de nos résultats, nous étudions le comportement des algorithmes proposés en présentant d'autres simulations numériques qui traitent de l'évolution des performances en dB en fonction de trois paramètres à savoir : i) le nombre N_m de matrices à zéro-bloc diagonaliser, ii) les erreurs de modèles modélisées par un rapport signal à bruit et finalement iii) la taille d'un bloc r .

Nous considérons $M = N = 8$ c-à-d que la matrice $\mathbf{A}$ est carrée complexe, sa partie réelle et sa partie imaginaire étant générées aléatoirement selon une loi uniforme sur l'intervalle $[-1, 1]$. L'ensemble $\mathcal{N}$ est constitué d'un nombre N_m variable de matrices allant de 2 à 100 matrices complexes de taille $M \times M$. Nous évaluons aussi le RSB qui va de 5 jusqu'à 100 dB et la taille d'un bloc r peut varier de 1, 2 à 4.

En premier lieu, nous étudions le comportement des algorithmes proposés en fonction du nombre de matrices N_m à zéro bloc-diagonaliser conjointement. En second lieu, nous passons à l'étude des performances en fonction du RSB à taille de bloc r fixée tout en initialisant par la DVPG. Pour une taille de bloc $r = 4$ (resp. 2) donnée, l'évolution de l'indice d'erreur $\text{l}_{\text{Conv}}(\cdot)$ en dB est fournie au niveau de la Fig. 2.21 (resp. 2.23) en fonction de la taille N_m de l'ensemble de matrices considéré et ce pour deux valeurs de RSB : 20 dB (resp. 30 dB) (à gauche) et 100 dB pour $r = 4$ (resp. 2) (à droite).

Pour des tailles de bloc identiques, l'indice d'erreur $\text{l}_{\text{Conv}}(\cdot)$ en dB est représenté au niveau des Fig. 2.22 et 2.24 en fonction du rapport signal à bruit pour deux tailles d'ensemble $\mathcal{N}$ de matrices à zéro-bloc diagonaliser conjointement : $N_m = 5$ (à gauche) et $N_m = 100$ matrices (à droite).

Nous remarquons bien évidemment que les meilleurs résultats sont obtenus lorsque le RSB est favorable (grand). Pour des ensembles de matrices quasi zéro-bloc diagonales (i.e. RSB = 100 dB), nous ré-estimons parfaitement le zéro bloc-diagonaliseur conjoint commun à toutes les matrices de l'ensemble $\mathcal{N}$ puisqu'à la convergence nous atteignons $I \approx -120$ dB. Nous vérifions toujours que plus que le nombre N_m de matrices à décomposer conjointement est grand, meilleures sont les performances. Finalement, comme précédemment, les algorithmes préconditionnés sont les plus performants, les plus rapides et les plus adaptés à des situations défavorables (peu de matrices au niveau de l'ensemble $\mathcal{N}$ d'une part, erreurs de modèle d'autre part).

Nous nous intéressons maintenant au comportement des algorithmes proposés en fonction de la taille des blocs, r , en contexte quasi non-bruité (resp. bruité (RSB = 30 dB)). Les résultats obtenus sont présentés au niveau de la Fig. 2.25 (resp. Fig. 2.26). La taille N_m de l'ensemble $\mathcal{N}$ et le RSB sont fixés. Les performances des algorithmes de ZBDC proposés dépendent de la taille

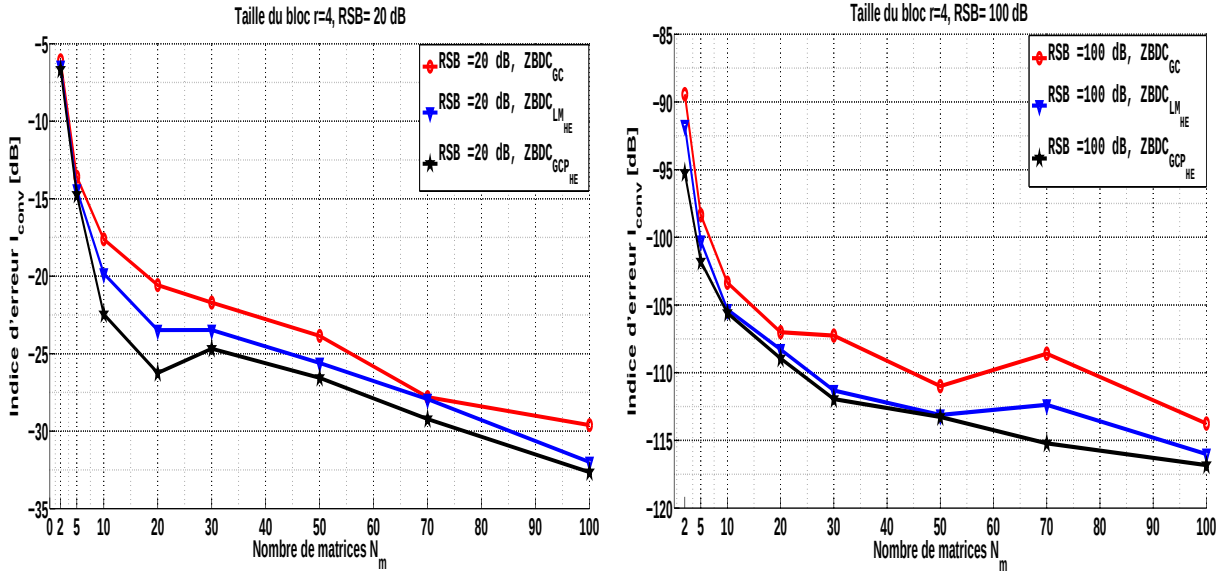


FIGURE 2.21 – Comparaison des algorithmes (ZBDC_{GC}, ZBDC_{LM_{HE}}, ZBDC_{GCP_{HE}}). À gauche : RSB = 20 dB. À droite : RSB = 100 dB.

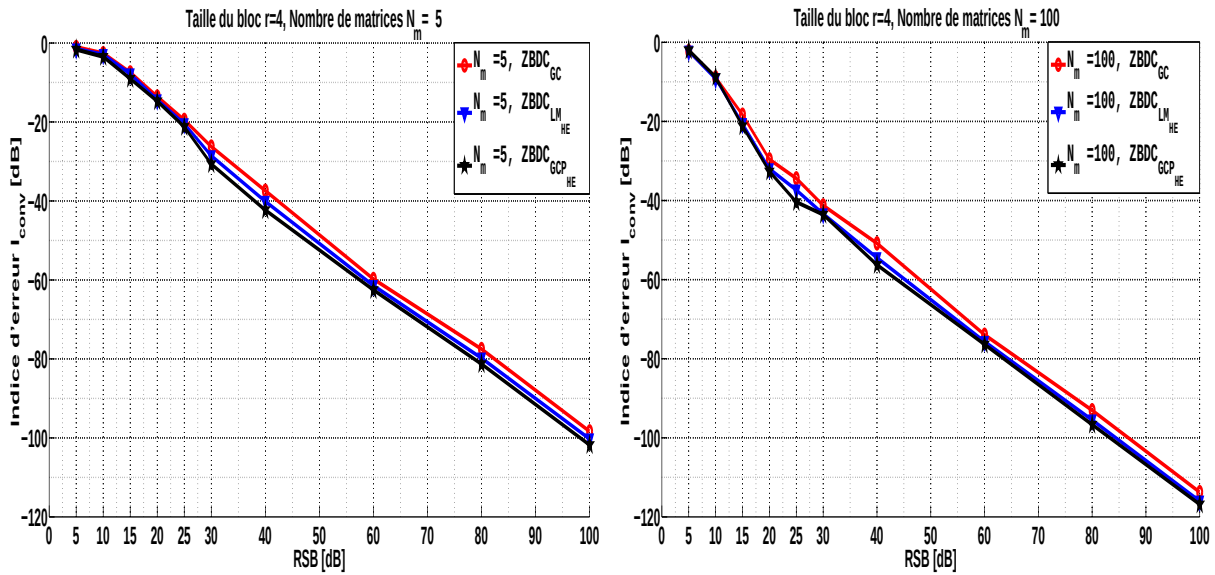


FIGURE 2.22 – Taille d'un bloc $r = 4$: performances en fonction du rapport signal à bruit des algorithmes ZBDC_{GC}, ZBDC_{LM_{HE}}, ZBDC_{GCP_{HE}}. Gauche : $N_m = 5$ matrices. Droite : $N_m = 100$ matrices.

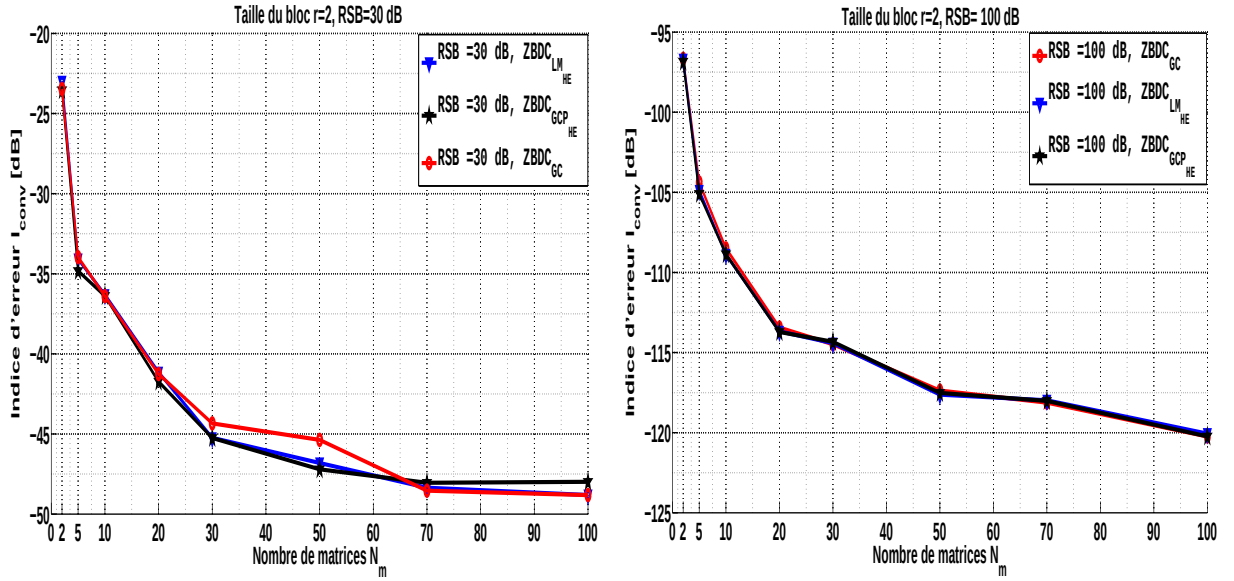


FIGURE 2.23 – Pour une taille de bloc fixée à $r = 2$, performances en fonction du nombre de matrices N_m pour les différents algorithmes (ZBDC_{GC}, ZBDC_{LM_{HE}}, ZBDC_{GCP_{HE}}). À gauche : RSB = 30 dB. À droite : RSB = 100 dB.

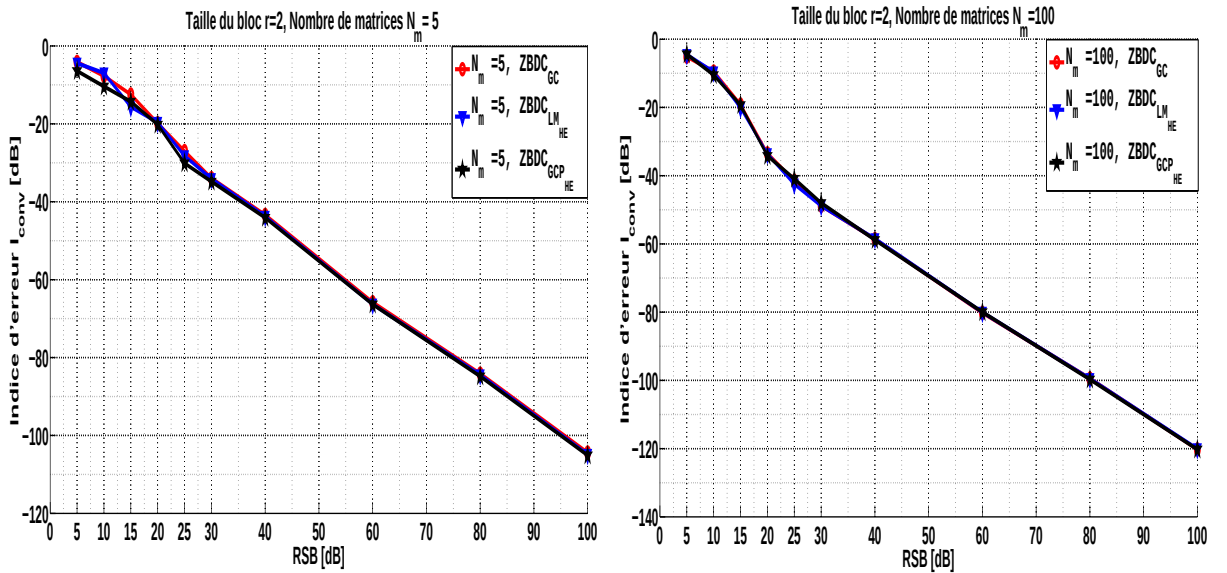


FIGURE 2.24 – Pour une taille de bloc fixée à $r = 2$, performances en fonction du rapport signal à bruit pour les algorithmes ZBDC_{GC}, ZBDC_{LM_{HE}}, ZBDC_{GCP_{HE}}. À gauche : $N_m = 5$ matrices. À droite : $N_m = 100$ matrices.

r de bloc considérée. mais qu'en dépit de cela, toutes les conclusions auxquelles nous sommes précédemment arrivés restent valides : elles sont donc indépendantes de la taille des blocs. Enfin, l'algorithme du gradient conjugué préconditionné ($\text{ZDC}_{\text{GCP}_{\text{HE}}}$) reste celui qui permet d'atteindre les meilleures performances, suivi des algorithmes $\text{ZBDC}_{\text{LM}_{\text{HE}}}$ et ZBDC_{GC} . Ces performances sont d'autant meilleures que l'on diminue la valeur de la taille r du bloc. Rappelons que dans le cas particulier où l'on considère que $r = 1$, nous retombons sur le problème de la zéro-diagonalisation conjointe d'un ensemble de matrices complexes (les algorithmes que nous proposons pouvant aussi servir à résoudre un sous-problème à savoir celui de la ZDC). Trois nouvelles solutions au problème de la zéro-diagonalisation conjointe non unitaire peuvent donc être suggérées : nous les noterons respectivement $\text{ZDC}_{\text{GCP}_{\text{HE}}}$, ZDC_{GC} , et $\text{ZDC}_{\text{LM}_{\text{HE}}}$.

2.3.4 Cas particulier : la ZDC d'un ensemble de matrices complexes

Dans cette dernière partie, nous montrons que les algorithmes de ZBDC précédents présentent aussi un intérêt lorsqu'ils sont utilisés dans le contexte plus simple de la zéro-diagonalisation conjointe. Nous conservons les mêmes paramètres pour nos simulations que ceux utilisés au paragraphe précédent (initialisation par DVPG) et évaluons l'évolution de l'indice d'erreur $\text{I}_{\text{Conv}}(\cdot)$ en fonction de la taille N_m de l'ensemble de matrices considérées. Ces résultats sont fournis au niveau de la Fig. 2.27 pour deux valeurs de $\text{RSB} = 30$ dB (à gauche) et $\text{RSB} = 100$ dB (à droite).

Ensuite, l'évolution de l'indice d'erreur $\text{I}_{\text{Conv}}(\cdot)$ en dB est tracée au niveau de la Fig. 2.28 en fonction du rapport signal à bruit pour deux tailles d'ensembles considérées : $N_m = 5$ matrices (à gauche) et $N_m = 100$ matrices (à droite).

Nous comparons maintenant le comportement des méthodes itératives proposées à celui d'une méthode directe suggérée dans [49] et que nous noterons $\text{ZDC}_{\text{Chabriel}}$. Nous opérons toujours sur un ensemble donné, $\mathcal{N}$, de matrices complexes. Par ailleurs, la taille N_m de cet ensemble et le RSB seront fixés. Nous fournissons dans le Tableau 2.5 (resp. 2.6) les performances atteintes par les méthodes itératives proposées en les comparant mais aussi les initialisant par la méthode de [49]. Nous utilisons pour nos simulations un $\text{RSB} = 20$ dB (resp. $\text{RSB} = 100$ dB) et deux tailles N_m d'ensembles respectivement égales à 5 et 100.

Algorithme	RSB = 20 dB	Performances
ZDC_{GC}	$N_m = 5$	$\text{I}_{\text{Conv}}(\mathbf{G}) = -24.0781$ dB
	$N_m = 100$	$\text{I}_{\text{Conv}}(\mathbf{G}) = -29.84$ dB
$\text{ZDC}_{\text{LM}_{\text{HE}}}$	$N_m = 5$	$\text{I}_{\text{Conv}}(\mathbf{G}) = -24.5816$ dB
	$N_m = 100$	$\text{I}_{\text{Conv}}(\mathbf{G}) = -31.02$ dB
$\text{ZDC}_{\text{GCP}_{\text{HE}}}$	$N_m = 5$	$\text{I}_{\text{Conv}}(\mathbf{G}) = -27.794$ dB
	$N_m = 100$	$\text{I}_{\text{Conv}}(\mathbf{G}) = -35.52$ dB
$\text{ZDC}_{\text{Chabriel}}$	$N_m = 5$	$\text{I}_{\text{Conv}}(\mathbf{G}) = -23.670$ dB
	$N_m = 100$	$\text{I}_{\text{Conv}}(\mathbf{G}) = -28.88$ dB

TABLE 2.5 – Cas de la zéro-diagonalisation conjointe ($r = 1$) : comparaison des performances des algorithmes itératifs $\text{ZDC}_{\text{GCP}_{\text{HE}}}$, ZDC_{GC} , $\text{ZDC}_{\text{LM}_{\text{HE}}}$ à celles d'une méthode directe $\text{ZDC}_{\text{Chabriel}}$ pour un $\text{RSB} = 20$ dB et deux tailles d'ensembles $N_m = 5$ et 100.

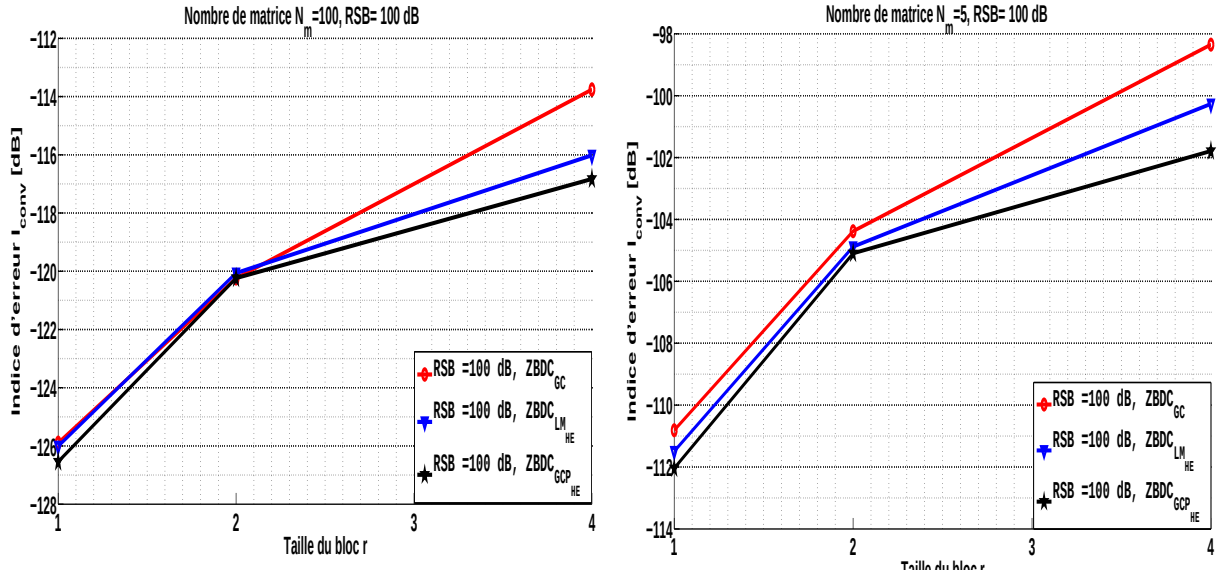


FIGURE 2.25 – Comparaison des algorithmes $\text{ZBDC}_{\text{GCP}_{\text{HE}}}$, ZBDC_{GC} et $\text{ZBDC}_{\text{LM}_{\text{HE}}}$ dans un contexte quasi non-bruité. Gauche : $I_{\text{Conv}}(\mathbf{G})$ en dB en fonction de la taille d'un bloc r pour un ensemble $\mathcal{N}$ de taille $N_m = 100$ matrices (resp. $N_m = 5$ matrices à droite).

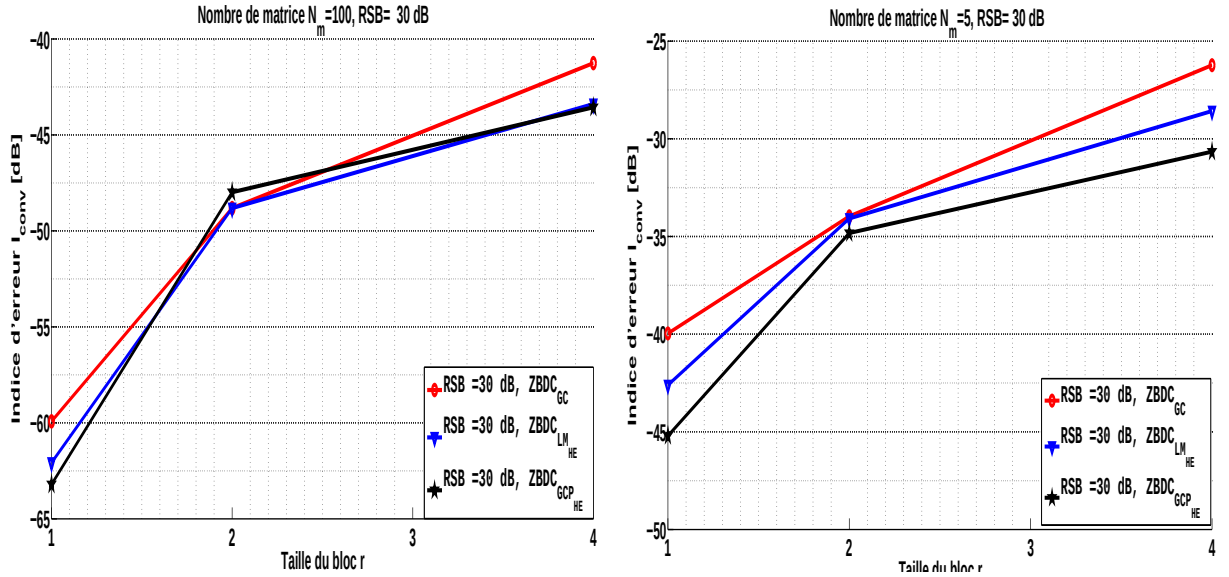


FIGURE 2.26 – Comparaison des algorithmes $\text{ZBDC}_{\text{GCP}_{\text{HE}}}$, ZBDC_{GC} et $\text{ZBDC}_{\text{LM}_{\text{HE}}}$ dans un contexte bruité ($\text{RSB} = 30$ dB). Gauche : $I_{\text{Conv}}(\mathbf{G})$ en dB en fonction de la taille d'un bloc r pour un ensemble $\mathcal{N}$ de taille $N_m = 100$ matrices (resp. $N_m = 5$ matrices à droite).

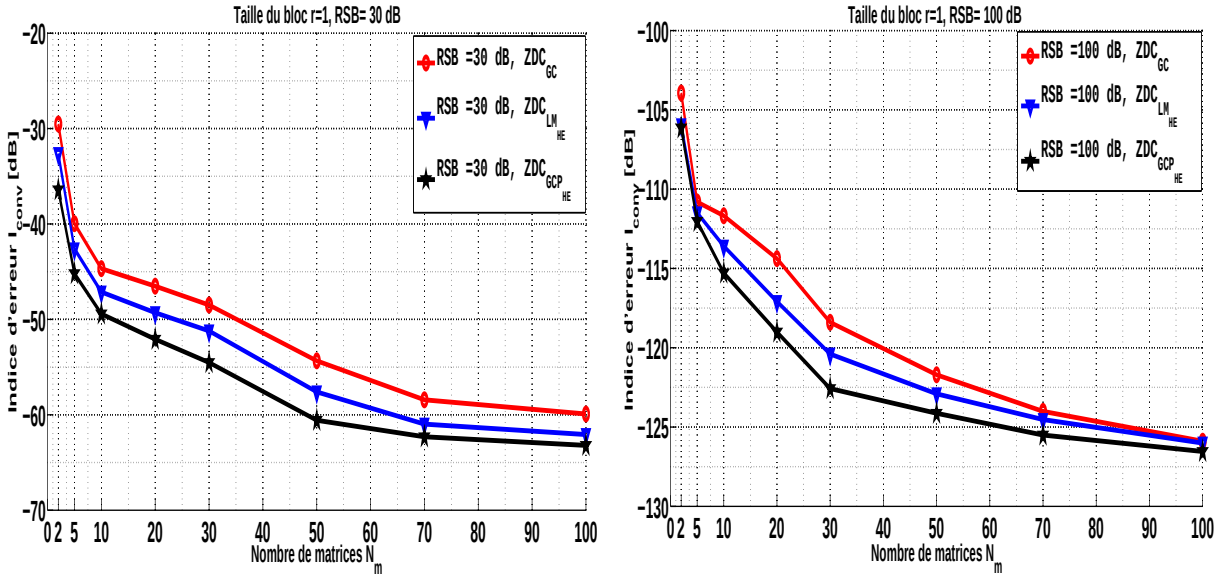


FIGURE 2.27 – Zéro-diagonalisation conjointe ($r = 1$) : performances en fonction du nombre de matrices N_m pour les algorithmes $\text{ZDC}_{\text{GCP}_{\text{HE}}}$, ZDC_{GC} , $\text{ZDC}_{\text{LM}_{\text{HE}}}$. À gauche : $\text{RSB} = 30$ dB. À droite : $\text{RSB} = 100$ dB.

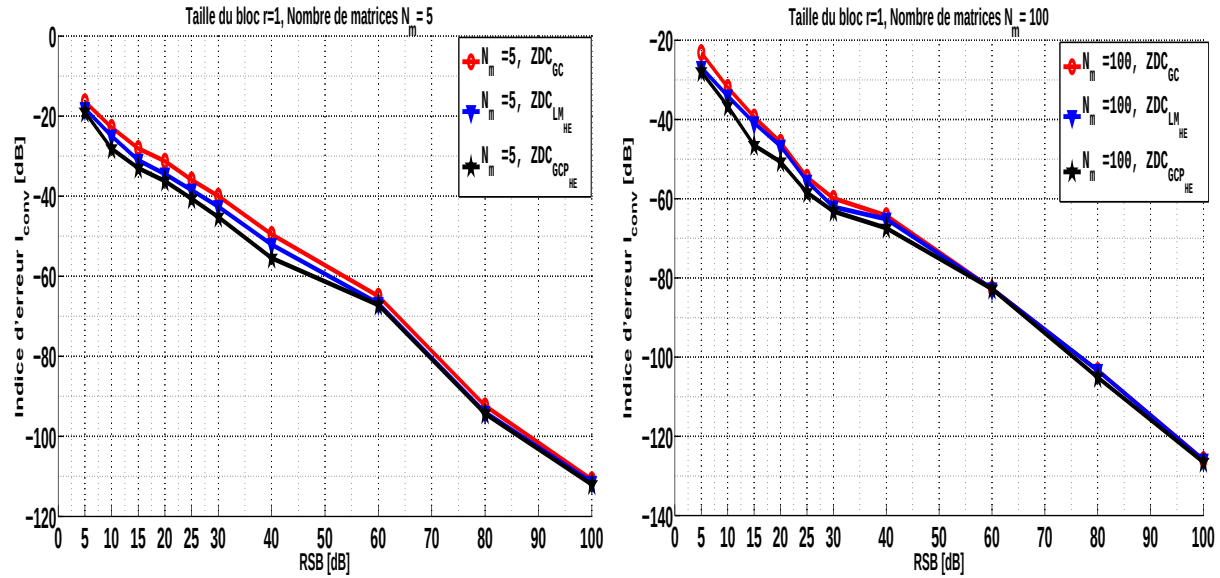


FIGURE 2.28 – Zéro-diagonalisation conjointe ($r = 1$) : performances en fonction du rapport signal à bruit pour les algorithmes $\text{ZDC}_{\text{GCP}_{\text{HE}}}$, ZDC_{GC} , $\text{ZDC}_{\text{LM}_{\text{HE}}}$. À gauche : $N_m = 5$ matrices. À droite : $N_m = 100$ matrices.

Algorithme	RSB =100 dB	Performances
ZDC _{GC}	$N_m=5$	$l_{\text{Conv}}(\mathbf{G}) = -109.01 \text{ dB}$
	$N_m=100$	$l_{\text{Conv}}(\mathbf{G}) = -123.019 \text{ dB}$
ZDC _{LM_{HE}}	$N_m=5$	$l_{\text{Conv}}(\mathbf{G}) = -109.4 \text{ dB}$
	$N_m=100$	$l_{\text{Conv}}(\mathbf{G}) = -123.13 \text{ dB}$
ZDC _{GCP_{HE}}	$N_m=5$	$l_{\text{Conv}}(\mathbf{G}) = -110.11 \text{ dB}$
	$N_m=100$	$l_{\text{Conv}}(\mathbf{G}) = -124.212 \text{ dB}$
ZDC _{Chabriel}	$N_m=5$	$l_{\text{Conv}}(\mathbf{G}) = -109.10 \text{ dB}$
	$N_m=100$	$l_{\text{Conv}}(\mathbf{G}) = -122.74 \text{ dB}$

TABLE 2.6 – Zéro-diagonalisation conjointe ($r = 1$) : comparaison des performances des algorithmes itératifs ZDC_{GCP_{HE}}, ZDC_{GC}, ZDC_{LM_{HE}} à celles d’une méthode directe ZDC_{Chabriel} pour un RSB =100 dB et deux tailles d’ensembles $N_m = 5$ et 100 .

Nous pouvons retirer exactement les mêmes conclusions présentées dans le cadre de la ZBDC_{NU} dans ce cas particulier qui n’est rien d’autre que la ZDC.

2.4 Conclusion

Dans ce chapitre, nous avons traité deux types de décompositions matricielles conjointes. Tout d’abord, nous nous sommes intéressés au problème de la bloc-diagonalisation conjointe non-unitaire d’un ensemble donné de matrices. Nous avons présenté quatre nouveaux algorithmes permettant de résoudre ce problème. Ils sont tous itératifs et passent par la minimisation d’une même fonction de coût quadratique de type moindres carrés. Le premier algorithme, noté BDC_{GC}, repose sur un schéma d’optimisation de type gradient conjugué et reprend les résultats sur le gradient matriciel complexe établi dans [113]. Le second algorithme de type Levenberg-Marquardt, baptisé BDC_{LM_{HE}}, repose sur un calcul exact des matrices Hessiennes complexes. Ce calcul théorique nous a conduit à développer deux autres algorithmes utilisant un préconditionneur : BDC_{GCP_{HE}} et BDC_{QN_{HE}} basés respectivement sur le gradient conjugué préconditionné et sur des méthodes Quasi-Newton. Le préconditionneur que nous choisissons d’utiliser est relié aux matrices Hessiennes complexes calculées. Cela nous a ensuite permis d’étudier l’impact et l’intérêt d’un préconditionnement. À l’aide des simulations informatiques, nous avons montré que l’algorithme BDC_{GCP_{HE}} offre le meilleur compromis entre efficacité et vitesse de convergence même en contexte difficile (peu de matrices dans l’ensemble à décomposer conjointement et erreurs de modèle relativement importantes). Nous avons également montré que ces algorithmes peuvent être utilisés dans la résolution du problème de la diagonalisation conjointe non unitaire.

Dans un deuxième temps, nous nous sommes efforcé de généraliser le problème de la zéro-diagonalisation conjointe en considérant un nouveau problème de décompositions matricielles conjointes que nous avons choisi d’intituler “zéro bloc diagonalisation conjointe non-unitaire” d’un ensemble donné de matrices complexes. Nous avons commencé par présenter et justifier notre choix de fonction de coût à minimiser pour estimer le zéro-bloc diagonaliseur conjoint. Nous avons proposé un nouvel algorithme basé sur l’approche du gradient conjugué noté ZBDC_{GC}. Cette méthode d’optimisation itérative nécessite le calcul exact du gradient matriciel complexe - calcul que nous avons effectué au cours de ce chapitre. Nous avons également proposé deux autres algo-

rithmes : l'un baptisé $\text{ZBDC}_{\text{LM}_{\text{HE}}}$ basé sur une approche itérative de type Levenberg-Marquardt et fondé sur un calcul exact des matrices Hessiennes, l'autre sur l'algorithme du gradient conjugué préconditionné baptisée $\text{ZBDC}_{\text{GCP}_{\text{HE}}}$. Finalement, nous avons illustré le comportement des approches de ZBDC_{NU} proposées à l'aide de simulations numériques dans différentes situations. Ces simulations ont montré que l'algorithme $\text{ZBDC}_{\text{GCP}_{\text{HE}}}$ est le plus performant surtout en situation difficile (peu de matrices et beaucoup de bruits) ce qui semble conforter l'intérêt et le choix du préconditionneur que nous avons fait. Nous avons également montré que ces algorithmes peuvent être utilisés dans la résolution du problème de la zéro-diagonalisation conjointe non unitaire.

Toutes ces méthodes de BDC_{NU} et/ou ZBDC_{NU} trouvent des applications intéressantes dans différents problèmes de traitement du signal, et plus spécifiquement en séparation aveugle de sources et en déconvolution multidimensionnelle aveugle. C'est l'objet du chapitre suivant que de montrer comment ces méthodes peuvent s'appliquer dans le contexte applicatif qui est le nôtre à savoir la surveillance maritime par satellite (S – AIS). Les matrices de l'ensemble à (bloc)-diagonaliser, à zéro-(bloc) diagonaliser conjointement pourront alors être des matrices de corrélation ou bien des matrices de corrélation cyclique estimées, ou encore des matrices issues des distributions temps-fréquence spatiales quadratiques ou bien des spectres temps-fréquence spatiaux quadratiques, ou des matrices issues des fonctions d'ambiguïtés moyennes généralisées, etc...

Chapitre 3

Nouvelles méthodes de séparation en surveillance maritime par satellite : cas sur-déterminé

Nous présentons dans ce chapitre de nouvelles approches afin de traiter le problème considéré. Ces approches sont prévues pour des modèles sur-déterminés dans lequel les n signaux $S - AIS$ (les sources) sont enregistrés au niveau de m capteurs (le nombre de capteurs étant plus grand que le nombre de sources *i.e.* $m > n$). Dans un tel contexte, nous proposons donc des méthodes de séparation exploitant principalement les algorithmes de décompositions matricielles conjointes proposés dans le chapitre précédent. En effet, comme indiqué au niveau des sous-sections §2.1.3 et §2.2.3, les algorithmes de bloc diagonalisation et zéro-bloc diagonalisation conjointe peuvent aussi solutionner le problème de la séparation de mélanges instantanés de sources. Dans ce chapitre, nous allons donc montrer comment les algorithmes développés au chapitre précédent vont nous permettre de traiter le problème de la dé-collision des sources $S - AIS$ - problème que nous avons réécrit à l'issue du premier chapitre, comme un problème de séparation aveugle de mélanges instantanés de sources -. Pour ce faire, les signaux considérés étant des signaux aléatoires cyclo-stationnaires, nous nous appuierons sur les méthodes fondées sur l'utilisation de Distributions Spatiales Quadratiques. Etant donnée l'application en surveillance maritime à laquelle nous nous intéressons et le caractère très spécifique des signaux que nous traitons (aléatoires et cyclo-stationnaires), nous choisirons d'utiliser soit la fonction d'Ambiguïté Moyenne Spatiale (Généralisée), soit des Spectres Temps-Fréquence Spatiaux Quadratiques (STFSQ). Nous montrerons ensuite comment la (zéro-)diagonalisation conjointe d'ensembles de matrices issus de l'une ou l'autre de ces distributions spatiales quadratiques, nous permettra la restitution des sources $S - AIS$. Nous rappellerons que ces ensembles de matrices sont construits à l'aide de critères de sélection automatique de "bons" points dans l'espace au niveau duquel on travaille, dans notre cas il s'agira en fait de points "fréquence Doppler-retard". Les points retenus correspondent à des matrices aux propriétés algébriques bien particulières. Enfin, au moyen de simulations informatiques, nous illustrerons le comportement et l'intérêt des approches proposées et les comparerons à d'autres méthodes du même type reposant sur des décompositions matricielles de type DC ou ZDC.

3.1 Approches basées sur les algorithmes de décomposition matricielle conjointe dans le cas sur-déterminé

Nous proposons dans cette section des méthodes permettant de résoudre le problème de séparation de mélanges sur-déterminés de sources $S - AIS$. Ces méthodes combinent généralement une transformation du domaine d'observation que l'on opère au moyen d'un opérateur bilinéaire (passage du domaine temps-capteurs au domaine temps-fréquence-capteurs par exemple). Dans ce nouveau domaine, on peut alors utiliser un détecteur automatique de points utiles, lequel va permettre la construction de l'ensemble des matrices à décomposer conjointement. Une fois cet ensemble construit, on peut appliquer un algorithme de décomposition matricielle conjointe du type de ceux proposés au chapitre précédent. Les principaux avantages de ce genre de méthodes résident dans le fait qu'elles s'appliquent à un large éventail de signaux (du moment qu'ils sont non stationnaires : signaux déterministes corrélés ou non, signaux aléatoires non stationnaires corrélés ou non) et quelles sont assez directes, en ce sens qu'elles ne requièrent pas d'étape préalable de blanchiment des signaux reçus au niveau des capteurs. Nous pouvons décomposer ces approches en trois étapes [24] qui consistent en :

- 1) l'application d'un opérateur bilinéaire caractérisé par son noyau $\mathbf{K}(\alpha, \beta) = \mathbf{K}_{\alpha, \beta}$ aux observations non blanchies et l'obtention de matrices $\mathbf{D}_{\mathbf{xx}}(\alpha, \beta; \mathbf{K}_{\alpha, \beta})$ en chacun des points (α, β) ,
- 2) la sélection de certains points (α, β) permettant la construction d'ensembles de matrices possédant toutes une structure algébrique bien particulière (cf. Annexe 4, consacrée majoritairement à ce problème de la sélection automatique de "bons" points (α, β)),
- 3) l'application d'un algorithme de décomposition matricielle conjointe spécifique par exemple un algorithme de DC et/ou de ZDC de l'ensemble (ou des ensembles) précédent(s) lorsque l'on travaille en contexte de mélange instantané afin d'estimer le système mélangeant de même que les sources $S - AIS$.

Remarque 1. Dans le cas d'un vecteur $\mathbf{x}(t)$ de signaux non-stationnaires ou cyclo-stationnaires, on peut se ramener à quatre domaines d'observation liés les uns aux autres par transformée de Fourier (ou transformée de Fourier inverse)[131][217]. Le couple (α, β) peut être un couple :

- "temps - retard" (t, τ) , et l'on travaille alors avec des matrices (fonctions dans le cas scalaire) de covariance ou d'autocorrélation instantanée que l'on pourra par exemple noter $\Gamma_{\mathbf{xx}}(t, \tau)$,
- "temps - fréquence" (t, ν) , et l'on travaille alors avec des distributions (resp. des spectres) temps-fréquence spatiales(aux) quadratiques selon qu'il s'agit de signaux déterministes ou de signaux aléatoires. Au Chapitre 1, nous les avons notés $\mathbf{D}_{\mathbf{xx}}(t, \nu; \mathbf{K})$. On peut alors citer l'exemple de la distribution de Wigner-Ville spatiale ou du spectre de Wigner-Ville spatial, notés plus simplement $\mathbf{WV}_{\mathbf{xx}}(t, \nu)$ ou $\mathbf{SWV}_{\mathbf{xx}}(t, \nu)$,
- "fréquence Doppler - retard" (μ, τ) , et l'on travaille alors avec des distributions spatiales quadratiques notées $\mathbf{D}_{\mathbf{xx}}(\mu, \tau; \mathbf{q})$ ³³ dont les deux plus connues demeurent la fonction d'Ambiguïté spatiale (pour les signaux déterministes) et la fonction d'Ambiguïté moyenne spatiale (pour les signaux aléatoires), notées respectivement dans la suite $\mathbf{FA}_{\mathbf{xx}}(\mu, \tau)$ et $\bar{\mathbf{A}}_{\mathbf{xx}}(\mu, \tau)$. On note encore que dans le cas de signaux cyclo-stationnaires, on a coutume de plutôt parler de matrice (ou de fonction dans le cas scalaire) de corrélation cyclique (μ) représente alors la fréquence cyclique ou l'un de ses multiples),

33. En utilisant les mêmes notations que dans [217] en ce qui concerne les noyaux des transformées bilinéaires ou quadratiques.

- “fréquence Doppler - fréquence” (μ, ν) , et l’on travaille alors avec des distributions spatiales quadratiques notées $\mathbf{D}_{\mathbf{x}\mathbf{x}}(\mu, \nu; \mathbf{k})$. A notre connaissance, aucune distribution de ce domaine n’a encore été utilisée en SAS. On note toutefois que dans le cas de signaux cyclo-stationnaires, on parle plutôt de matrice de densité (inter)spectrale cyclique (ou tout simplement de densité spectrale cyclique dans le cas scalaire).

Généralement, les matrices des ensembles construits à l’issue de la deuxième étape se décomposent toutes sous la forme donnée par l’Eq. (1) ou par l’Eq. (3) de l’Annexe 4 en fonction du type de mélange considéré (instantané ou convolutif). Dans une telle décomposition, les matrices des distributions spatiales quadratiques des sources, $\mathbf{D}_{\mathbf{ss}}(\alpha, \beta)$, sont alors soit des matrices diagonales (ou bloc diagonales en modèle convolutif), ou zéro-(bloc) diagonales (cf. Annexe 4). L’algorithme de décomposition matricielle conjointe utilisé au niveau de la troisième étape dépend donc de la structure des matrices de l’ensemble construit soit directement soit au moyen d’un détecteur automatique de points particuliers. Ainsi dans le cas où ces matrices sont (bloc) diagonales alors la décomposition matricielle adoptée n’est rien d’autre que la (B)DC, alors que dans le cas où elles sont zéro-(bloc-) diagonales, la décomposition matricielle à utiliser est la Z(B)DC, deux décompositions que nous avons introduites au chapitre précédent.

Les nouvelles méthodes de séparation de sources S – AIS que nous allons présenter sont basées sur des décompositions matricielles conjointes qui possèdent l’avantage d’être plus générales et donc de pouvoir s’étendre à d’autres applications dans lesquelles les sources pourront être : réelles ou complexes, déterministes ou aléatoires, statistiquement mutuellement indépendantes ou corrélées, stationnaires, cyclo-stationnaires ou non-stationnaires et le système de mélange ne sera pas obligatoirement réel et/ou unitaire. Par ailleurs, il faut rappeler que le problème de la séparation de mélanges linéaires de sources au moyen de RTFSQ a déjà fait l’objet de nombreux travaux parmi lesquels [2][92][115]. Toutes les méthodes de l’état de l’art reposent sur une étape préalable de sélection de points $\alpha - \beta$ particuliers permettant ainsi la construction des ensembles de matrices devant ensuite être (bloc) diagonalisées conjointement et/ou zéro-(bloc) diagonalisées conjointement (cf. Annexe 4 pour un rapide aperçu des principaux détecteurs existants).

3.2 Approches exploitant la diversité temps-fréquence des signaux S – AIS

3.2.1 Approches basées sur l’utilisation combinée de spectres temps-fréquence spatiaux quadratiques et d’algorithmes de diagonalisation conjointe

Dans cette section, nous nous focalisons sur des approches de séparation sources en modèle sur-déterminé à base des spectres temps-fréquence spatiaux quadratiques combinés à nos nouveaux algorithmes de diagonalisation conjointe. Nous appliquerons ces techniques au problème de la dé-collision des trames S – AIS.

Pour assurer l’identifiabilité du mélange et la séparabilité des signaux, il faut que les signatures des sources S – AIS apparaissent comme suffisamment différentes dans le plan $t-f$ pour permettre la sélection d’un nombre suffisant de points $t-f$, lesquels permettront la construction des ensembles de matrices à décomposer conjointement.

Concernant les trois approches que nous allons étudier, nous utilisons un seul détecteur automatique de points $t-f$ pour la construction de l’ensemble des matrices devant être diagonalisées conjointement [115] (cf. Annexe 4 équation 11). Les méthodes résultantes ont été baptisées $\text{DC}_{\text{LMEH-TF}}$, $\text{DC}_{\text{GCP}_{\text{EH-TF}}}$ et $\text{DC}_{\text{GC-TF}}$ respectivement. La première repose sur l’algorithme de Levenberg-Marquardt présenté dans la section §2.1.6.2 du Chapitre 2, la deuxième est basée sur

l'algorithme du Gradient conjugué préconditionné proposé dans §2.1.6.1, (tous deux requièrent l'utilisation d'un préconditionneur calculé au niveau de §2.1.5.2). La dernière est fondée sur l'algorithme du gradient conjugué présenté lui aussi dans le chapitre précédent.

En outre, les approches fondées sur l'algorithme du gradient conjugué et Levenberg-Marquardt ont fait l'objet d'une publication [61]. Enfin nous rappelons que ces approches de séparation de sources **S** – **AIS** basées sur l'utilisation de STFSQ sont adaptées au cas de mélanges linéaires instantanés (voire convolutifs) sur-déterminés c'est à dire que le nombre de capteurs est supposé être supérieur au nombre de sources **S** – **AIS** recherchées.

3.2.2 Énoncé du problème

Nous considérons le modèle de séparation (ou de dé-collision) défini au niveau de l'Eq. (1.51). Nos développements sont basés sur les hypothèses suivantes :

Hyp1. Les bruits $n_j(t)$ pour tout $j = 1, \dots, m$ sont des bruits blancs Gaussiens stationnaires, centrés, mutuellement indépendants et indépendants des sources.

Hyp2. Il existe des points t - f correspondant à des auto-termes sur une seule source.

Hyp3. Le système **H** mélangeant les sources est de rang plein (système surdéterminé avec $m > n$). Il s'agit d'un système linéaire instantané où chaque source $s'_j(t)$ est affectée par un retard τ_j et une fréquence Doppler Δf_j : $s'_j(t) = [\exp\{-i2\pi\Delta f_j t\} s_j(t - \tau_j)]^T$.

3.2.3 Principe des approches $\text{DC}_{\text{LM}_{\text{EH}}-\text{TF}}$, $\text{DC}_{\text{GCP}_{\text{EH}}-\text{TF}}$ et $\text{DC}_{\text{GC}}-\text{TF}$

Le spectre temps-fréquence spatial quadratique (STFSQ) des observations noté $\mathbf{D}_{\mathbf{xx}}(t, \nu)$ de dimension $m \times m$ est de la forme suivante :

$$\mathbf{D}_{\mathbf{xx}}(t, \nu) = \mathbf{H}\mathbf{D}_{\mathbf{ss}}(t, \nu)\mathbf{H}^H + \mathbf{D}_{\mathbf{nn}}(t, \nu) + \mathbf{H}\mathbf{D}_{\mathbf{sn}}(t, \nu) + \mathbf{D}_{\mathbf{ns}}(t, \nu)\mathbf{H}^H, \quad (3.1)$$

où $\mathbf{D}_{\mathbf{ss}}(t, \nu)$ est une matrice carrée de taille $n \times n$ représentant le STFSQ des sources définie comme :

$$\mathbf{D}_{\mathbf{ss}}(t, \nu) = \begin{pmatrix} \mathbf{D}_{\mathbf{s}_1\mathbf{s}_1}(t, \nu) & \mathbf{D}_{\mathbf{s}_1\mathbf{s}_2}(t, \nu) & \dots & \mathbf{D}_{\mathbf{s}_1\mathbf{s}_n}(t, \nu) \\ \mathbf{D}_{\mathbf{s}_2\mathbf{s}_1}(t, \nu) & \mathbf{D}_{\mathbf{s}_2\mathbf{s}_2}(t, \nu) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{D}_{\mathbf{s}_{n-1}\mathbf{s}_n}(t, \nu) \\ \mathbf{D}_{\mathbf{s}_n\mathbf{s}_1}(t, \nu) & \dots & \mathbf{D}_{\mathbf{s}_n\mathbf{s}_{n-1}}(t, \nu) & \mathbf{D}_{\mathbf{s}_n\mathbf{s}_n}(t, \nu) \end{pmatrix}, \quad (3.2)$$

$\mathbf{D}_{\mathbf{nn}}(t, \nu)$ représente le STFSQ des bruits, alors que $\mathbf{D}_{\mathbf{sn}}(t, \nu)$ et $\mathbf{D}_{\mathbf{ns}}(t, \nu)$ sont les STFSQ croisés entre les sources et bruits. D'après **Hyp1**, les bruits sont indépendants des sources, si bien que le STFSQ des observations vaut alors :

$$\mathbf{D}_{\mathbf{xx}}(t, \nu) = \mathbf{H}\mathbf{D}_{\mathbf{ss}}(t, \nu)\mathbf{H}^H + \mathbf{D}_{\mathbf{nn}}(t, \nu). \quad (3.3)$$

Nous commençons par une étape de débruitage, en choisissant de ne conserver que les points t - f (t_k, ν_k) d'énergie suffisante et en écartant les points de faible énergie. Cela se traduit à travers la condition suivante :

$$|D_{x_i x_j}(t_k, \nu_k)|^2 > \epsilon \quad i, j \in \{1, \dots, m\}, \quad (3.4)$$

où ϵ est un nombre réel de faible valeur (ϵ peut être supérieur à 0.01 lorsque les matrices sont normalisées) et $D_{x_i x_j}(t_k, \nu_k) = (\mathbf{D}_{\mathbf{xx}}(t_k, \nu_k))_{ij}$. Rappelons encore qu'il est possible d'estimer la

variance du bruit lorsque $m > n$ par la moyenne des $m - n$ plus petites valeurs propres de la matrice de corrélation des observations à l'origine.

Nous allons dorénavant considérer le modèle non bruité (ou débruité) suivant :

$$\mathbf{D}_{\mathbf{xx}}(t, \nu) \simeq \mathbf{H}\mathbf{D}_{\mathbf{ss}}(t, \nu)\mathbf{H}^H. \quad (3.5)$$

L'hypothèse **Hyp₂** assure l'identifiabilité du mélange. Nous rappelons que les auto-termes sont les termes quadratiques qui correspondent aux éléments diagonaux de la matrice $\mathbf{D}_{\mathbf{ss}}(t, \nu)$. En raison du principe des interférences, en un point (t, ν) où seraient présents des auto-termes sur plus d'une seule source, ces sources vont interférer pour donner également des cross-termes et la matrice $\mathbf{D}_{\mathbf{ss}}(t, \nu)$ ne sera alors plus diagonale (cf Annexe 4 et les 4 catégories de points temps-fréquence). Par conséquent, la matrice $\mathbf{D}_{\mathbf{ss}}(t, \nu)$ ne peut être diagonale qu'en des points où il n'y a qu'un auto-terme sur une seule source et une structure de matrice diagonale avec un seul terme non-nul sur la diagonale est la seule structure possible de matrice diagonale (la matrice est alors de rang 1).

3.2.4 Détecteur de “bons” points temps-fréquences

Dans le cadre de notre application, sur l'ensemble des détecteurs présentés au niveau de l'Annexe 4, nous avons opté pour celui par L. Giulieri *et al.* dans [115]. Il est noté $\mathbf{C}_{\mathbf{G}_1}$ et il tire avantage de la propriété de rang 1 *i.e.* les matrices de l'ensemble $\mathcal{M} = \{\mathbf{M}_i = \mathbf{D}_{\mathbf{xx}}(t_i, \nu_i) \in \mathcal{C}^{(m \times m)}\}, \forall i, \dots, N_m$ construit au moyen de ce détecteur admettent toutes une décomposition de la forme donnée par l'Eq. (3.5) (qui n'est autre que la décomposition présentée au niveau de l'Eq. (2.1) obtenue en considérant $L = 0$, $L' = 1$ et $\mathbf{A} = \mathbf{H}$). Dans cette décomposition, les matrices $\mathbf{D}_{\mathbf{ss}}(t_i, \nu_i) \forall i, \dots, N_m$ sont non seulement des matrices diagonales mais en plus elles ne possèdent qu'un seul terme non nul sur la diagonale. La DC de cet ensemble permettra alors d'identifier le système de séparation. Cette opération est réalisée au moyen d'un des trois algorithmes itératifs introduits au niveau du Chapitre 2.

3.2.5 Résumé des approches proposées

Les différentes approches utilisant des STFSQ ont été notées $\text{DC}_{\text{LM}_{\text{EH}}-\text{TF}}$, $\text{DC}_{\text{GCP}_{\text{EH}}-\text{TF}}$ et $\text{DC}_{\text{GC}}-\text{TF}$. Nous en résumons ci-dessous les principales étapes :

- 1) Application d'un opérateur quadratique et construction des matrices de STFSQ sur les observations non blanchies.
- 2) Débruitage des matrices t - f à l'aide de la règle donnée au niveau de l'Eq. (3.4).
- 3) Construction de l'ensemble $\mathcal{M}$ des matrices à diagonaliser conjointement en utilisant le critère $\mathbf{C}_{\mathbf{G}_1}$ (cf. Annexe 4 équation 11) pour la sélection automatique de points t - f utiles.
- 4) Diagonalisation conjointe de l'ensemble $\mathcal{M}$ en utilisant un algorithme de DC non-unitaire (Levenberg-Marquardt ou Gradient conjugué préconditionné ou Gradient conjugué).
- 5) Estimation de la matrice de mélange et restitution des sources S – AIS.

3.2.6 Simulations informatiques

Dans cette sous-section, nous illustrons au moyen de simulations informatiques le comportement des méthodes $\text{DC}_{\text{LM}_{\text{EH}}-\text{TF}}$, $\text{DC}_{\text{GCP}_{\text{EH}}-\text{TF}}$ et $\text{DC}_{\text{GC}}-\text{TF}$ et les comparons avec le même type

d'approches (à l'étape 4 précédente au lieu d'utiliser l'un de nos algorithmes de diagonalisation conjointe, nous utiliserons l'algorithme de D. Nion [184] qui minimise une autre fonction de coût que celle que nous utilisons : la méthode que nous construirons ainsi sera baptisée $DC_{GC_{NION-TF}}$). Ces méthodes reposent toutes sur l'utilisation du détecteur C_{G_1} .

Nous considérons $m = 3$ mélanges de $n = 2$ trames de 256 bits correspondant à deux navires avec des caractéristiques différentes. Les deux trames sont générées conformément à la recommandation S – AIS appelée dans [162]. Elles sont représentées au niveau de la Fig. 3.1. Ensuite, ces trames sont codées en NRZI et modulées en GMSK en respectant le produit temps bande spécifié dans le système S – AIS ($BT_s = 0.4$ avec $T_s = 1/9600$ s). L'ordre du filtre Gaussien est choisi égal à $OF = 21$. Nous fixons pour la première source (resp. la deuxième source) un retard causé par la distance entre le navire et le satellite de l'ordre de 5 ms (resp. 0 ms). Nous considérons également que l'effet Doppler dû à la vitesse du satellite par rapport au navire est de 10 Hz (resp. 0 Hz). Ces deux sources possèdent la même fréquence porteuse de 161.975 MHz. Ces sources correspondent à 1208 échantillons (points) en temps. La partie réelle de la représentation temporelle des sources est illustrée au niveau la Fig. 3.1. Nous utilisons ensuite comme distribution temps-fréquence spatiale quadratique le spectre de Pseudo Wigner-Ville (calculé en utilisant une fenêtre de Hamming de taille 151 points en temps et en prenant 604 points en fréquence). La partie réelle (resp. la partie imaginaire) du STFSQ des sources est donnée à gauche (resp. à droite) de la Fig. 3.2. La Fig. 3.3 illustre la partie réelle des signaux observés alors que la Fig. 3.4 fournit leur STFSQ.

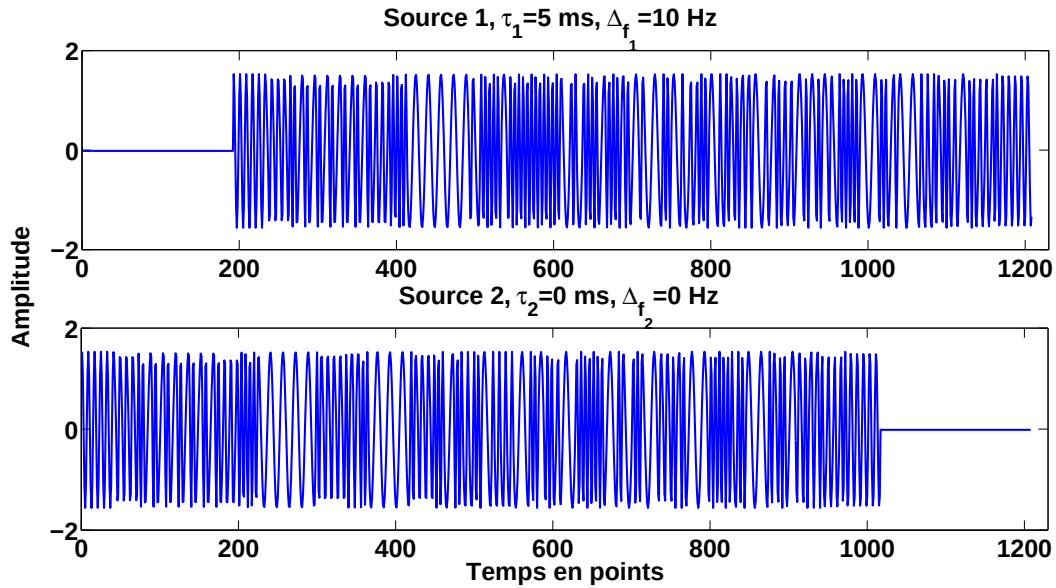


FIGURE 3.1 – Partie réelle des sources S – AIS après codage NRZI et GMSK. Haut : source 1 avec retard de 5 ms et Doppler de 10 Hz. Bas : source 2 avec retard de 0 ms et Doppler de 0 Hz.

Les sources sont mélangées au moyen de la matrice de mélange $\mathbf{H}$ dont les composantes sont données ci-dessous :

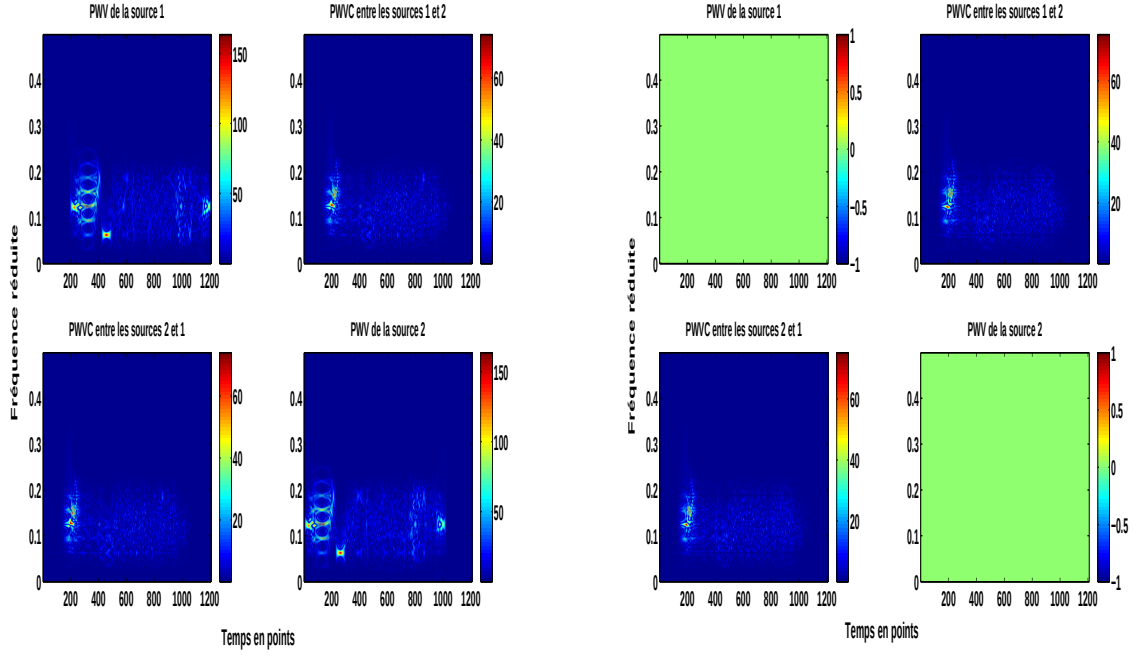


FIGURE 3.2 – À gauche : partie réelle de la distribution de Pseudo Wigner-Ville des sources S – AIS. À droite : partie imaginaire de la distribution de Pseudo Wigner-Ville des sources S – AIS.

$$\mathbf{H} = \begin{pmatrix} 0.2082 & 0.1450 \\ -0.8666 & -0.1220 \\ -0.7612 & 0.8398 \end{pmatrix}. \quad (3.6)$$

En ce qui concerne le détecteur de points $\mathbf{C}_{G_1}$ utilisé, nous avons fixé les seuils $\epsilon_1 = 6$ et $\epsilon_2 = 0.3$. Pour juger de la qualité des résultats obtenus, nous nous servons de l'indice d'erreur introduit dans [171][149] qui n'est rien d'autre qu'un cas particulier de celui décrit dans le paragraphe §2.1.6. Les points temps-fréquence retenus par le détecteur automatique sont représentés au niveau de la Fig. 3.5.

Enfin, nous appliquons à l'ensemble des matrices retenues les différents algorithmes de diagonalisation conjointes : $\mathbf{DC}_{GC-TF}$, $\mathbf{DC}_{GCP_{HE}-TF}$ et $\mathbf{DC}_{LM_{HE}-TF}$. A la convergence, l'indice d'erreur obtenu en contexte bruité ($RSB = 20$ dB) vaut alors -26.35 dB. Nous donnons également l'évolution de l'indice d'erreur en fonction du rapport signal à bruit (RSB) considéré au niveau de la Fig. 3.6. Ces résultats résultent d'un moyennage sur 30 tirages aléatoires (par méthode de Monte-Carlo).

Nous observons au moyen de ces courbes le relativement bon comportement des trois approches proposées $\mathbf{DC}_{GCP_{HE}-TF}$, $\mathbf{DC}_{LM_{HE}-TF}$ et $\mathbf{DC}_{GC-TF}$ dans la mesure où nous constatons un gain en dB par rapport à l'approche $\mathbf{DC}_{GC-TF}$ qui elle repose sur l'algorithme de diagonalisation conjointe $\mathbf{DC}_{GC_{NION}-TF}$. Les meilleurs performances sont généralement atteintes au moyen des méthodes du second ordre et plus précisément la méthode $\mathbf{DC}_{GCP_{HE}-TF}$ qui repose sur l'algorithme du gradient conjugué préconditionné. Nous vérifions enfin qu'en exploitant la diversité temps-fréquence même en contexte très difficile (faible RSB) nous semblons parvenir à séparer les messages AIS.

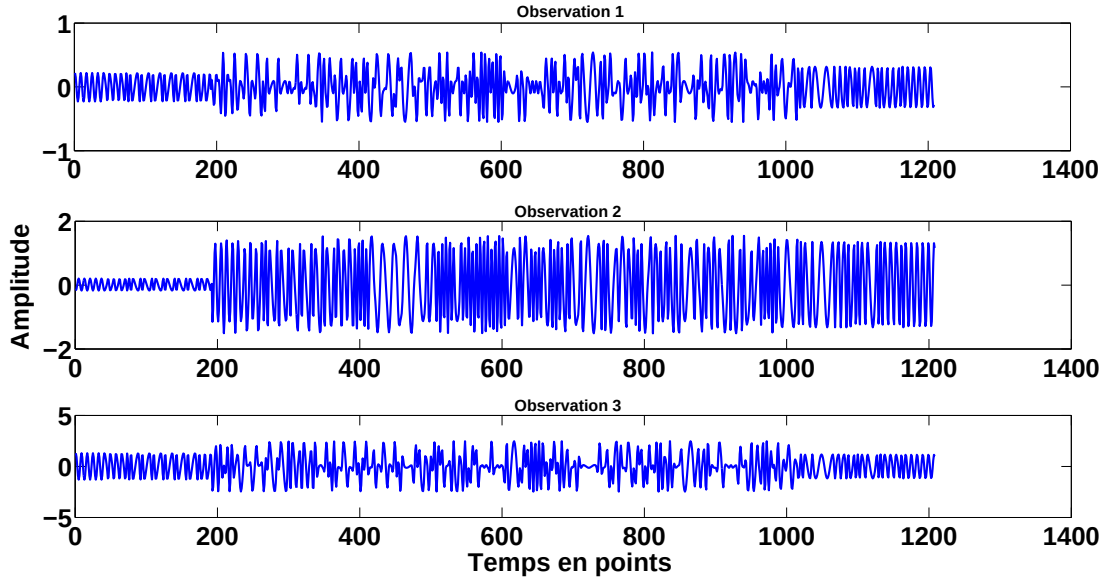


FIGURE 3.3 – Partie réelle des trois signaux supposés reçus au niveau des capteurs de l’antenne.

3.3 Approches exploitant la diversité Doppler-Retard des sources $S - AIS$

3.3.1 Approches exploitant le caractère cyclo-stationnaire à l’ordre deux des sources $S - AIS$

A l’origine, les méthodes de SAS ont été développées dans un contexte de signaux stationnaires (éventuellement bruités), en exploitant la propriété d’indépendance statistique des signaux pour parvenir à les séparer. Cette hypothèse relativement forte impliquait alors une limite quant aux domaines d’applications possibles. Par la suite, l’utilisation de connaissances a priori sur les signaux traités a permis le développement de nouvelles méthodes de séparation aveugle exploitant de nouvelles contraintes (parcimonie, non négativité, etc...). De même l’introduction de nouvelles hypothèses comme la non-stationnarité des signaux a permis de considérer de nouvelles diversités (passage au domaine temps-fréquence-capteur³⁴ par exemple, ou encore fréquence d’émission - fréquence d’excitation - concentrations en spectroscopie de fluorescence (application pour laquelle on utilise alors plutôt des décompositions tensorielles telles que la décomposition Canonique Polyadique (CP))). Cela a permis d’appliquer des méthodes de séparation aveugle à de nouvelles classes de signaux et donc d’étendre le champs de leurs applications possibles. Dans cette sous-section, nous allons donc nous appuyer sur une information a priori supplémentaire à savoir le fait que les sources $S - AIS$ sont des processus cyclo-stationnaires. La théorie de la cyclo-stationnarité a vu le jour dans les années 70 à travers les travaux de thèse de Hurd [134]. Ses travaux ont ensuite été repris dans les années 80 par W. Gardner [100][101][102]. Il a notamment montré leur utilité dans la modélisation des signaux de télécommunications numériques. De nombreuses méthodes, exploitant cette propriété ont alors vu le jour (cf. la dernière section de l’Annexe 4 pour un bref état de l’art).

34. Catégorie d’approches à laquelle nous nous intéressons dans cette thèse.

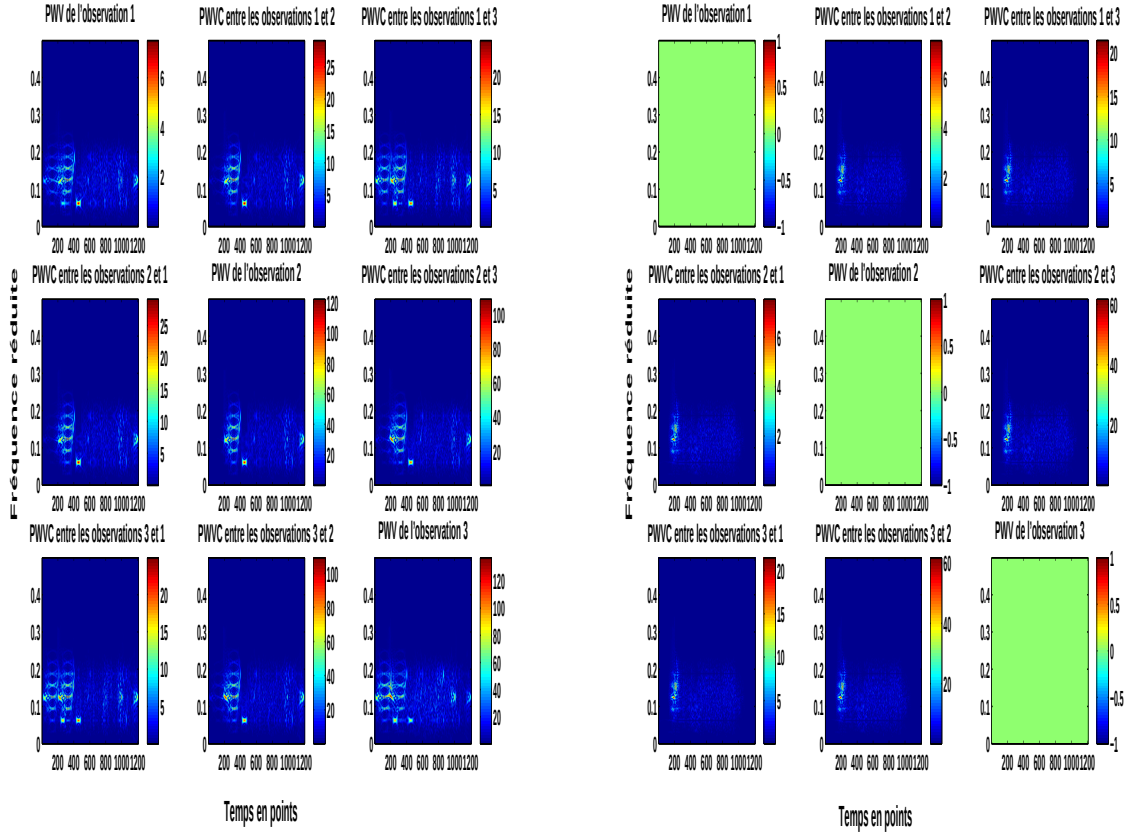


FIGURE 3.4 – À gauche : partie réelle de la distribution de Pseudo Wigner-Ville des observations. À droite : partie imaginaire de la distribution de Pseudo Wigner-Ville des observations.

3.3.2 Les outils de ce domaine : fonction d'ambiguïté et fonction d'ambiguïté moyenne spatiale

Dans la section précédente, les distributions (et les spectres) temps-fréquence spatiales (aux) quadratiques ont été introduits et représentés sous une forme matricielle. Des décompositions matricielles leur ont ensuite été appliquées³⁵. Il a été rappelé que la relation qui lie les distributions t - f spatiales quadratiques des observations aux distributions temps-fréquence spatiales quadratiques des sources étaient de la même nature que celle qui lie la matrice de covariance des observations à celle des sources. Cette relation clé est celle qui permet de résoudre le problème de l'estimation des directions d'arrivée en traitement d'antenne et celui de l'estimation du système mélangeant et de la réestimation des sources en SAS. En ce qui concerne la SAS, il a été montré que les distributions temps-fréquence spatiales étaient un outil efficace sous condition toutefois que les signatures des sources diffèrent en certains points du plan t - f ³⁶.

Dans le cas de sources cyclo-stationnaires, le domaine fréquence Doppler-retard semble toutefois être un domaine d'analyse plus naturel pour la ré-estimation des signaux sources que ne l'est le domaine temps-fréquence. Comme le préconise M. G. Amin dans [13], il convient alors d'utiliser

35. On note quand même que ces données auraient également pu être considérées sous une forme tensorielle et que d'autres types de décompositions que celles utilisées ici.

36. Des sources de signatures identiques sont bien évidemment non séparables.

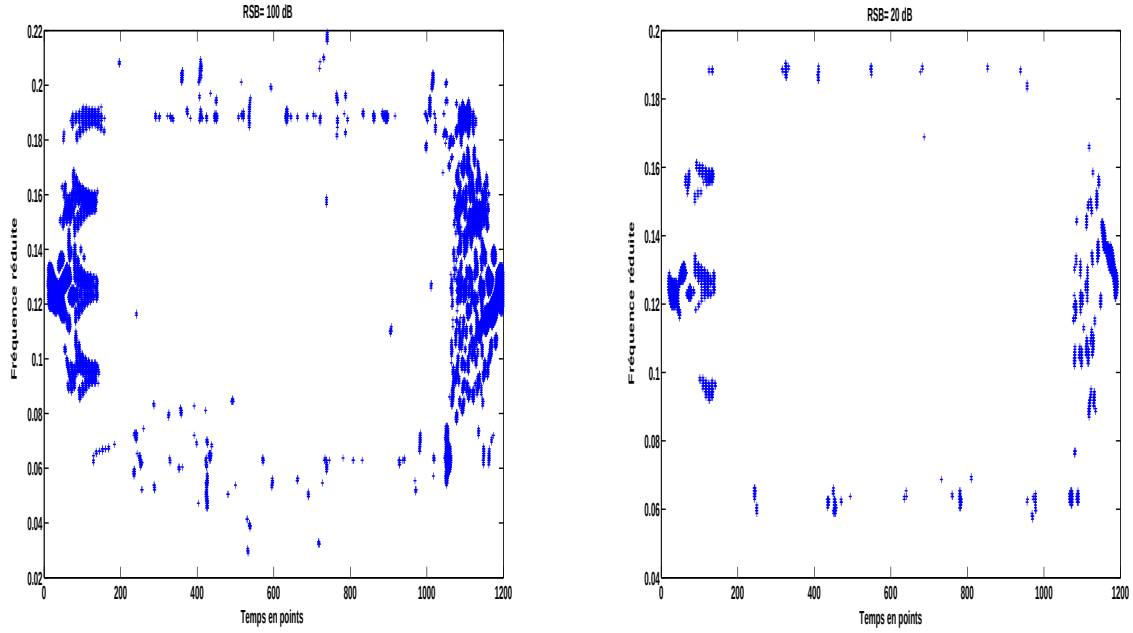


FIGURE 3.5 – Points temps-fréquence sélectionnés à l’aide du détecteur C_{G_1} . À gauche : $RSB = 100$ dB. À droite : $RSB = 20$ dB.

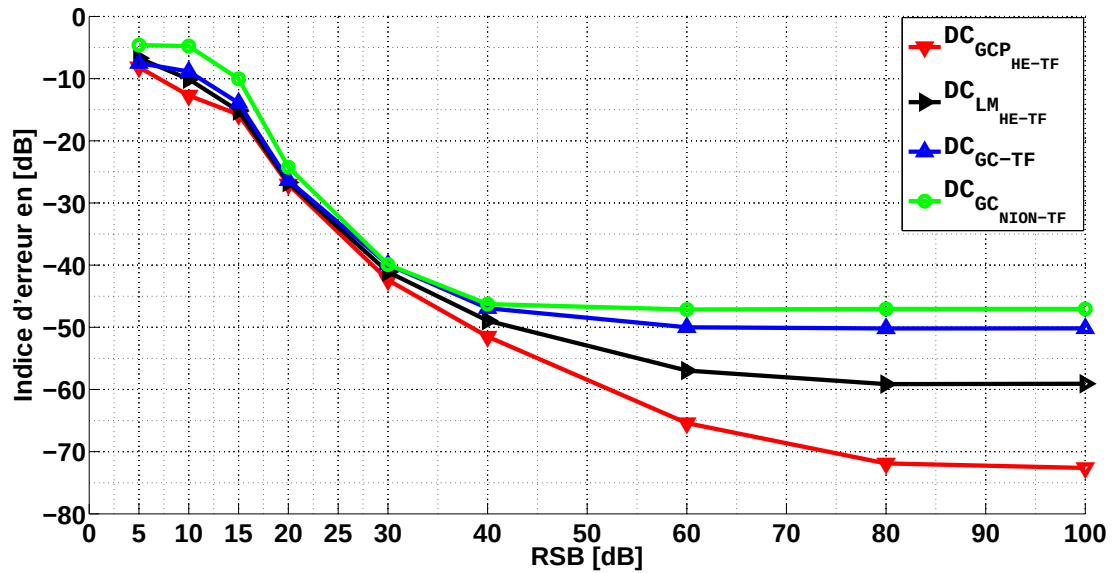


FIGURE 3.6 – I en dB en fonction du RSB. Comparaison des résultats obtenus avec les différentes méthodes proposées DC_{LM_HE-TF} , DC_{GCP_HE-TF} et DC_{GC-TF} à $DC_{GC_NION-TF}$.

plutôt des approches basées sur des informations tirées des fonctions d’Ambiguïtés spatiales (FAS) ou sur des fonctions d’Ambiguïtés moyennes spatiales (éventuellement généralisées que nous

noterons $\mathbf{F}\mathbf{A}_{\mathbf{S}\mathbf{S}}[130]$). On rappelle qu'avec les notations adoptées précédemment, la fonction d'Ambiguïté spatiale est définie de manière suivante :

$$\mathbf{F}\mathbf{A}_{\mathbf{S}\mathbf{S}}(\mu, \tau) = \int_{-\infty}^{\infty} \mathbf{x}\left(t + \frac{\tau}{2}\right) \mathbf{x}^H\left(t - \frac{\tau}{2}\right) e^{-2j\pi\mu\tau} dt, \quad (3.7)$$

où μ et τ désignent respectivement le décalage en fréquence ou Doppler et le décalage temporel ou retard.

Les sources $\mathbf{S} - \mathbf{AIS}$ sont des processus aléatoires, cyclo-stationnaires³⁷. Nous utiliserons donc non pas des fonctions d'Ambiguïté spatiales mais des fonctions d'Ambiguïté moyennes spatiales (éventuellement généralisées) ($\mathbf{F}\mathbf{A}_{\mathbf{S}\mathbf{S}}$). Pour tout signal vectoriel complexe $\mathbf{z}(t)$, la $\mathbf{F}\mathbf{A}_{\mathbf{S}\mathbf{S}}$ est donnée par la relation suivante [130][146][145] :

$$\bar{\mathbf{A}}_{\mathbf{z}}(\mu, \tau) = \int_{-\infty}^{\infty} \mathbf{r}_{\mathbf{z}}(t, \tau) e^{-j2\pi\mu t} dt = \mathbb{E} \{ \langle \mathbf{z}, (\mathbf{k}_{\mu, \tau} \mathbf{z}) \rangle \}, \quad (3.8)$$

avec $(\mathbf{k}_{\mu, \tau} \mathbf{z})$ est l'opérateur de translations élémentaires du retard-Doppler de $\mathbf{z}$ défini par $(\mathbf{k}_{\mu, \tau} \mathbf{z})(t) = \mathbf{z}(t - \tau) e^{j2\pi\mu(t - \frac{\tau}{2})}$ et

$$\mathbf{r}_{\mathbf{z}}(t, \tau) = \mathbf{R}_{\mathbf{z}}\left(t + \frac{\tau}{2}, t - \frac{\tau}{2}\right) = \mathbb{E} \left\{ \mathbf{z}\left(t + \frac{\tau}{2}\right) \mathbf{z}^H\left(t - \frac{\tau}{2}\right) \right\}, \quad (3.9)$$

où $\mathbf{R}_{\mathbf{z}}(t, \tau)$ représente la matrice de corrélation de $\mathbf{z}(t)$, $\mathbb{E} \{ \cdot \}$ désigne l'opérateur d'espérance mathématique et $(\cdot)^H$ est l'opérateur de transposition conjugaison. L'interprétation de la fonction $\bar{\mathbf{A}}_{\mathbf{z}}(\mu, \tau)$ est qu'elle caractérise la corrélation moyenne de toutes les paires de points Doppler-retard séparées par τ en temps et par μ en fréquence [146][145].

Nous rappelons que les termes diagonaux de la matrice $\bar{\mathbf{A}}_{\mathbf{z}}(\mu, \tau)$ sont des termes quadratiques que l'on appelle des auto-termes tandis que les termes hors diagonale sont des termes bilinéaires que l'on appelle termes-croisés ou cross-termes (voire inter-termes).

3.3.3 Sous l'effet d'un mélange linéaire instantané

Dans un contexte de mélange linéaire instantané non bruité des signaux pour lequel $\mathbf{x}(t) = \mathbf{H}\mathbf{s}(t)$, alors la fonction d'Ambiguïté spatiale des observations s'écrit :

$$\mathbf{F}\mathbf{A}_{\mathbf{S}\mathbf{S}}(\mu, \tau) = \mathbf{H}\mathbf{F}\mathbf{A}_{\mathbf{S}\mathbf{S}}(\mu, \tau)\mathbf{H}^H. \quad (3.10)$$

On constate que cette équation est très similaire à l'Eq. (3.5), les matrices impliquées étant les $\mathbf{F}\mathbf{A}_{\mathbf{S}\mathbf{S}}$ des observations et des sources et non plus les distributions (ou les spectres) temps-fréquence spatiaux quadratiques des observations et des sources. Nous rappelons que les termes diagonaux de la matrice $\mathbf{F}\mathbf{A}_{\mathbf{S}\mathbf{S}}(\mu, \tau)$ sont des termes quadratiques que l'on appelle des auto-termes tandis que les termes hors diagonale sont des termes bilinéaires que l'on appelle termes-croisés ou cross-termes (voire inter-termes).

Certains problèmes de séparation aveugle de sources peuvent donc être résolus au moyen des $\mathbf{F}\mathbf{A}_{\mathbf{S}\mathbf{S}}$. En remplaçant juste les distributions temps-fréquence spatiales $\mathbf{D}_{\mathbf{S}\mathbf{S}}(t, \nu)$ par les fonctions d'Ambiguïté spatiales $\mathbf{F}\mathbf{A}_{\mathbf{S}\mathbf{S}}(\mu, \tau)$, on peut facilement transposer les méthodes de $\mathbf{SAS}$ décrites précédemment au domaine Doppler-retard. Il suffit ensuite de suivre les mêmes étapes de traitement que celles décrites précédemment. Dans [13], les auteurs montrent que l'intérêt principal

37. L'une des difficultés majeures est liée au fait que seules deux fréquences cycliques très proches sont utilisées pour les transmettre.

de cette approche réside aussi dans le fait que la sélection des auto-termes sur les sources est simplifiée, ceux-ci étant positionnés près de l'origine.

Plus tard, le même auteur présente dans [14], les FAS et discute de leurs applications aux problèmes de la DOA et de la SAS en insistant sur deux particularités qui les distinguent des autres distributions spatiales quadratiques et qui font d'elles un outil important pour le traitement du signal :

- 1) Dans le domaine des Ambiguïtés, les auto-termes sont positionnés à l'origine et à proximité, ce qui facilite la construction de la matrice souhaitée. Toutefois, les termes croisés entre deux signaux sources sont éloignés de l'origine.
- 2) Les auto-termes de tous les signaux à bande étroite, indépendamment de leurs fréquences et de leurs phases, tombent sur l'axe de retard ($\nu = 0$), tandis que ceux des signaux à large bande tombent sur une région différente (μ, τ) où se répartissent dans tout le plan des ambiguïtés.

Ces deux particularités permettent à la FAS d'être un bon choix pour la récupération et la localisation de sources à bande étroite. Lorsque l'on travaille dans un contexte de signaux aléatoires avec des fonctions d'Ambiguïté moyennes et que l'on continue à s'intéresser au modèle de mélange linéaire instantané donné par l'Eq. (1.51), il est possible de calculer la FAMSG des signaux observés notée $\bar{\mathbf{A}}_{\mathbf{xx}}(\mu, \tau)$. En un point Doppler-retard donné (μ_k, τ_k) , il s'agit d'une matrice $m \times m$ admettant la décomposition suivante :

$$\begin{aligned}\bar{\mathbf{A}}_{\mathbf{xx}}(\mu, \tau) &= \mathbf{H}\bar{\mathbf{A}}_{\mathbf{ss}}(\mu, \tau)\mathbf{H}^H + \bar{\mathbf{A}}_{\mathbf{nn}}(\mu, \tau), \\ &= \mathbf{H}\bar{\mathbf{A}}_{\mathbf{ss}}(\mu, \tau)\mathbf{H}^H + \mathbf{R}_{\mathbf{nn}}(\tau)\alpha(\mu),\end{aligned}\quad (3.11)$$

avec $\bar{\mathbf{A}}_{\mathbf{ss}}(\mu, \tau)$ la matrice de dimensions $n \times n$ des FAMSG des sources dont la définition a été donnée au niveau de l'Eq.(3.8) et $\mathbf{R}_{\mathbf{nn}}(\tau) = \sigma_{\mathbf{n}}^2\alpha(\tau)\mathbf{I}_m$ où :

$$\alpha(\mu) = \int_{-\infty}^{\infty} e^{-j2\pi\mu t} dt. \quad (3.12)$$

Lorsque $\mu \neq 0$, alors l'Eq. (3.11) devient :

$$\bar{\mathbf{A}}_{\mathbf{xx}}(\mu, \tau) = \mathbf{H}\bar{\mathbf{A}}_{\mathbf{ss}}(\mu, \tau)\mathbf{H}^H. \quad (3.13)$$

Nous nous intéressons alors à construire un ensemble de matrices issues des FAMSG pour des points μ loin de l'origine. De ce fait, nous faisons appel aux algorithmes de ZDC (au lieu de DC puisque les auto-termes sont positionnés à l'origine).

3.3.4 Approches de séparation basées sur les fonctions d'ambiguïté moyennes spatiales combinées à des algorithmes de zéro-diagonalisation conjointe non unitaire

Nous supposons que les signatures des sources S-AIS sont suffisamment différentes dans le plan Doppler-Retard pour que l'on puisse sélectionner un nombre suffisant de points (μ, τ) pour lesquelles les matrices $\bar{\mathbf{A}}_{\mathbf{ss}}(\mu, \tau)$ possèdent une structure particulière. Elles seront diagonales en des points où ne sont présents que des auto-termes sur les sources et zéro-diagonales en des points où ne seront présents que des cross-termes (ou termes croisés) sur les sources. Notre objectif sera de tirer avantage de ces propriétés pour i) parvenir à sélectionner de tels points Doppler-Retard

afin de construire des ensembles particuliers de matrices ii) appliquer à ces ensembles de matrices les algorithmes de décompositions matricielles conjointes non unitaires que nous avons développés iii) et ainsi estimer la matrice de mélange (ou sa pseudo-inverse la matrice de séparation) et ré-estimer les signaux sources.

3.3.5 Construction de l'ensemble $\mathcal{X}$ de matrices à partir des points Doppler-Retard sélectionnés

Pour construire l'ensemble $\mathcal{X}$ de matrices à zéro-diagonaliser conjointement, il nous faut sélectionner des points Doppler-Retard aux propriétés particulières. Nous opérons sur les observations non blanchies et nous utilisons le détecteur $\mathbf{C}_{F-ZD}$ suggéré dans [92] (voir l'Annexe 4 pour plus de détails). Toutes les matrices appartenant à l'ensemble $\mathcal{X}$ (dont la taille est notée N_m avec $N_m \in N^*$) se décomposent en $\mathbf{H}\bar{\mathbf{A}}_{ss}(\mu, \tau)\mathbf{H}^H$. Les matrices $\mathbf{X}_k = \bar{\mathbf{A}}_{ss}(\mu_k, \tau_k) \quad \forall k = 1, \dots, N_m$ sélectionnées sont toutes des matrices zéro-diagonales. Une fois cet ensemble $\mathcal{X}$ construit nous lui appliquerons l'un de nos algorithmes de zéro-diagonalisation conjointe non-unitaire basé sur la minimisation de la fonction de coût donnée au niveau de l'Eq. (2.53) afin d'estimer la matrice de séparation $\mathbf{B}$.

Ainsi, trois nouvelles méthodes de séparation de sources $\mathbf{S} - \mathbf{AIS}$ peuvent être dérivées. Ces trois approches sont baptisées : $\mathbf{ZDC}_{GC-DR}$, $\mathbf{ZDC}_{LM_{EH}-DR}$ et $\mathbf{ZDC}_{GCP_{EH}-DR}$. Elles sont respectivement basées sur l'utilisation d'algorithme de zéro-diagonalisation conjointe non unitaire de type i) Gradient conjugué proposé au niveau §2.2.5.2, ii) Levenberg-Marquardt proposé au niveau de §2.2.5.3 et iii) Gradient conjugué pré-conditionné proposé au niveau de §2.2.5.2.

3.3.6 Un récapitulatif du principe des approches suggérées

Les méthodes suggérées combinent l'utilisation de la fonction d'ambiguïté moyenne spatiale généralisée avec des algorithmes de zéro-diagonalisation conjointe non unitaire. Elles sont notées $\mathbf{ZDC}_{GC-DR}$, $\mathbf{ZDC}_{LM_{EH}-DR}$ et $\mathbf{ZDC}_{GCP_{EH}-DR}$ et leur principe est le suivant :

- 1) Calcul des matrices de **FAMSG** des observations non blanchies,
- 2) Construction de l'ensemble $\mathcal{X}$ des matrices à zéro-diagonaliser conjointement en utilisant le critère $\mathbf{C}_{F-ZD}$ de sélection automatique de points Doppler-Retard utiles,
- 3) Zéro-diagonalisation conjointe non unitaire de l'ensemble $\mathcal{X}$ en utilisant l'un des algorithmes de **ZDC** non-unitaire suggérés au niveau du Chapitre 2,
- 4) Estimation de la matrice de séparation $\mathbf{B}$ et reconstruction des sources $\mathbf{S} - \mathbf{AIS}$.

3.3.7 Simulations informatiques

Dans ce paragraphe, nous effectuons des simulations informatiques afin de valider l'approche proposée sur les signaux particuliers que nous traitons dans le cadre de notre application en surveillance maritime. Nous évaluons aussi le comportement des méthodes proposées et les comparons avec le même type d'approche mais en remplaçant nos algorithmes de zéro-diagonalisation conjointe par celui proposé dans [49]. Cette approche sera alors notée $\mathbf{ZDC}_{Chabriel-DR}$. Toutes ces méthodes sont testées sur un même ensemble des matrices issues de la fonction d'ambiguïté moyenne spatiale généralisée. Ces matrices ont été sélectionnées au moyen du détecteur proposé dans [92].

Nous considérons $m = 3$ mélanges de $n = 2$ sources correspondant à deux navires de caractéristiques distinctes. Les trames sont générées conformément à la recommandation $\mathbf{S} - \mathbf{AIS}$ [162].

Elles sont représentées au niveau de la Fig. 1.3. Elles sont codées en NRZI et modulées en GMSK en considérant un produit temps bande de $BT_s = 0.4$ avec $T_s = 1/9600$ s imposé par la norme, et l'ordre du filtre Gaussien est choisi égal à $OF = 21$ [263][196]. Dans ces simulations, la fréquence porteuse des deux sources est supposée être différente bien que très proche : elle est de 161.975 MHz pour la première source et de 162.025 MHz pour la seconde source. Dans l'exemple considéré ici, la première source subit un retard de 10 ms et un décalage Doppler de 4 kHz, alors que la deuxième n'est affectée ni par un retard (0 ms) ni par un effet Doppler (0 Hz). Ces sources correspondent alors à 1400 échantillons de points temps qui sont mélangées en fonction d'une matrice de mélange $\mathbf{H}$ dont les composantes sont les suivantes :

$$\mathbf{H} = \begin{pmatrix} -1.1974 & 1.3646 \\ 0.8623 & 1.6107 \\ 0.1568 & -0.9674 \end{pmatrix}. \quad (3.14)$$

La représentation temporelle de la partie réelle des sources est donnée au niveau de la Fig. 3.7. Tandis que la partie réelle (resp. la partie imaginaire) de leur FAMSG est donnée à gauche (resp. à droite) de la Fig. 3.8. Le détecteur C_{F-ZD} est appliqué à l'ensemble des matrices issues de la FAMSG des observations en considérant les valeurs de seuils suivantes $\epsilon_1 = 0.07$ et $\epsilon_2 = 0.08$. Au niveau de la Fig. 3.9, nous avons représenté les FAMSG des signaux observés. Les points Doppler-retard sélectionnés sont affichés au niveau de la Fig. 3.10. Nous utilisons l'indice d'erreur donné dans [171][149] pour juger de la qualité de l'estimation.

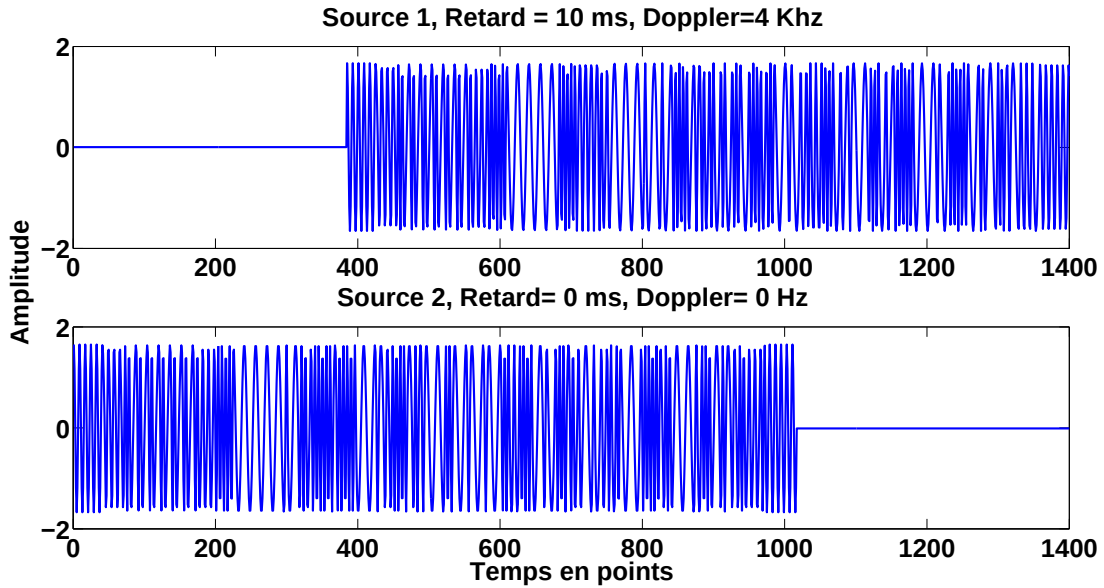


FIGURE 3.7 – Partie réelle des sources (trames S – AIS générées après codage NRZI et GMSK). Haut : première source avec retard de 10 ms et Doppler de 4 Khz. Bas : deuxième source avec retard de 0 ms et Doppler de 0 Hz.

Nous appliquons ensuite à l'ensemble des matrices Doppler-retard sélectionnées les différents algorithmes de ZDC sans contrainte d'unitarité sur la matrice recherchée, à savoir i) $ZDC_{LM_{HE-DR}}$, ii) $ZDC_{GCP_{HE-DR}}$, iii) ZDC_{GC-DR} et iv) $ZDC_{Chabriel-DR}$. En contexte relativement bruité ($RSB = 20$ dB), nous trouvons pour l'algorithme $ZDC_{GCP_{HE-DR}}$ (resp. ZDC_{GC-DR} , $ZDC_{LM_{HE-DR}}$ et $ZDC_{Chabriel-DR}$) un indice d'erreur égal à -44 dB (resp. - 39 db, -40 dB -10 dB). Enfin nous avons également tracé

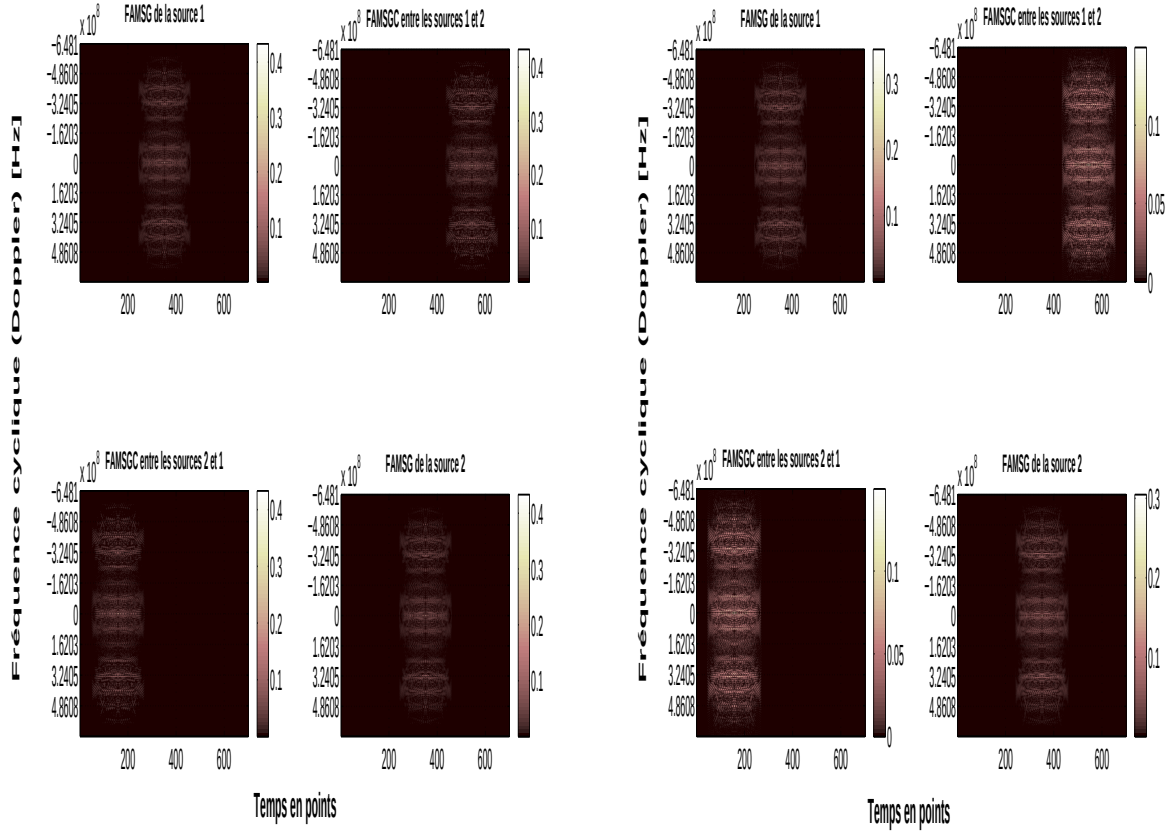


FIGURE 3.8 – À gauche : partie réelle de la FAMSG des sources $S - AIS$. À droite : partie imaginaire de la FAMSG des sources $S - AIS$.

l'évolution de l'indice d'erreur en fonction du rapport signal à bruit au niveau de la Fig. 3.11. Nous précisons encore que ces courbes résultent d'un moyennage sur 30 tirages par méthode de Monte-Carlo.

Au regard de cette figure, nous vérifions la validité de l'approche proposée et le comportement satisfaisant des approches $ZDC_{GCP_{HE-DR}}$, $ZDC_{LM_{HE-DR}}$ et ZDC_{GC-DR} par rapport à l'approche $ZDC_{Chabriel-DR}$. Nous observons aussi que les solutions $ZDC_{GCP_{HE-DR}}$, $ZDC_{LM_{HE-DR}}$ basées sur le calcul des matrices Hessiennes exactes sont les plus performants si on les compare à l'approche ZDC_{GC-DR} . Nous notons enfin que les meilleurs résultats sont généralement obtenus au moyen de la méthode $ZDC_{GCP_{HE-DR}}$ et ce même en contexte difficile *i.e* à faible RSB. Enfin, en comparant l'approche exploitant la diversité temps-fréquence présentée précédemment, à celle exploitant la diversité Doppler-retard des signaux $S - AIS$, nous remarquons que cette dernière approche semble assez prometteuse. De par sa relative robustesse au bruit, elle semble permettre de solutionner le problème de la SAS (et de la dé-collision des signaux) dans un contexte de surveillance maritime. Il reste quand même encore à étudier des cas plus réalistes et donc plus complexes dans lesquels le nombre de messages $S - AIS$ reçus au niveau de l'antenne du récepteur embarqué dans le satellite serait beaucoup plus élevé. On pourrait également considérer des modèles de mélange

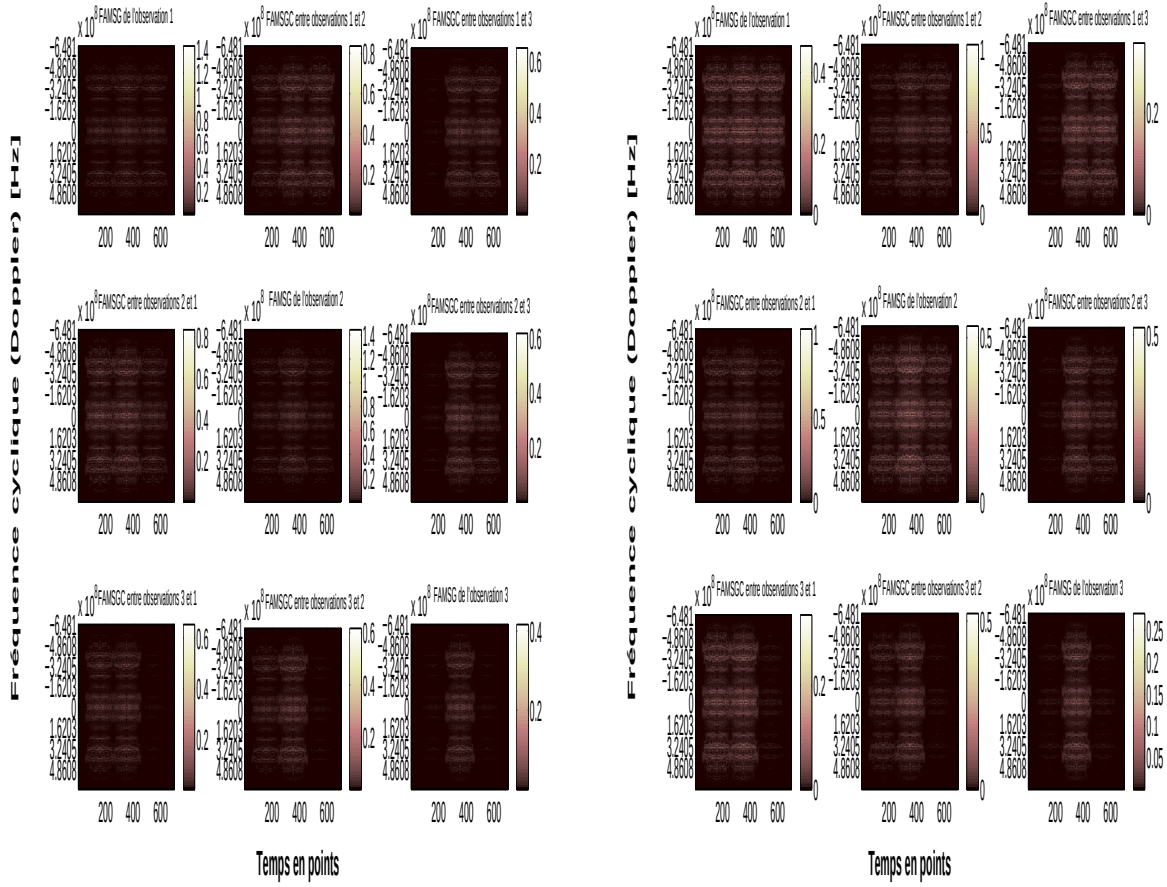


FIGURE 3.9 – À gauche : partie réelle de la FAMSG des observations. À droite : partie imaginaire de la FAMSG des observations.

plus compliqués.

3.4 Conclusion

Dans ce chapitre, nous avons suggéré de nouvelles approches destinées à la résolution du problème de la collision des messages $S - AIS$. Ayant - sous certaines hypothèses - reformulé ce problème comme un problème de séparation aveugle de mélanges linéaires instantanés sur-déterminé de sources, nous avons alors développé des méthodes inspirées des méthodes de séparation aveugle de sources fondées sur l'utilisation de distributions spatiales quadratiques. Ces approches combinent l'application d'un opérateur bilinéaire et le passage à un nouveau domaine $(\alpha, \beta, \text{capteurs})$ dans lequel les signaux se distinguent davantage grâce à l'ajout d'une diversité supplémentaire. Il est alors possible d'effectuer une sélection automatique de points utiles à la construction d'ensembles de matrices aux propriétés particulières. L'application d'algorithmes de décomposition matricielle conjointe particuliers (diagonalisation conjointe ou zéro-diagonalisation conjointe) permet alors l'estimation du système de séparation et la ré-estimation des signaux sources -

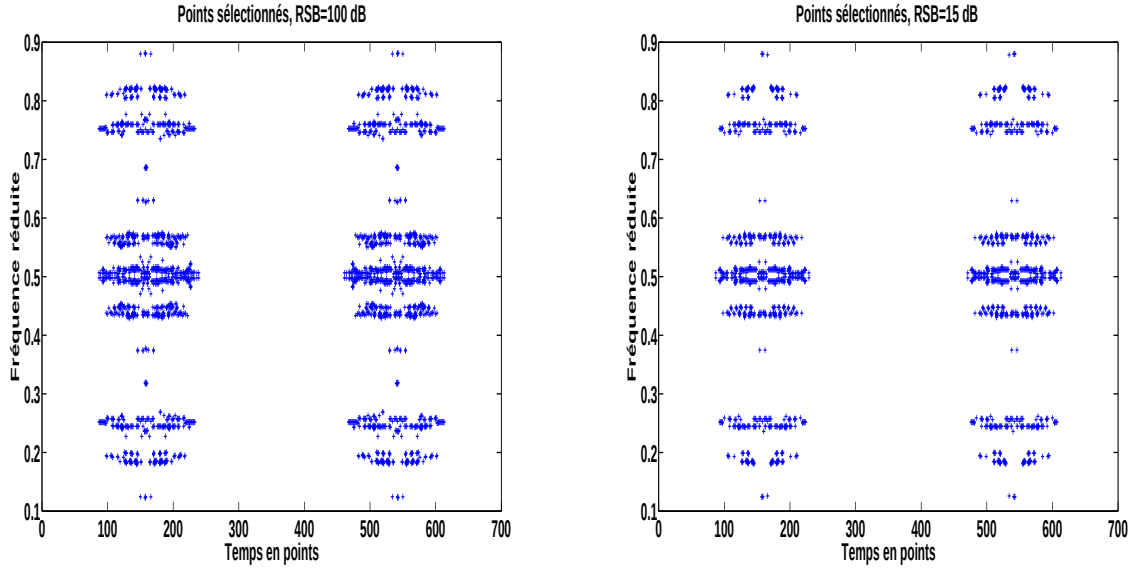


FIGURE 3.10 – Points Doppler-retard sélectionnés à l’aide du détecteur C_{F-ZD} . À gauche : $RSB = 100$ dB. À droite : $RSB = 15$ dB.

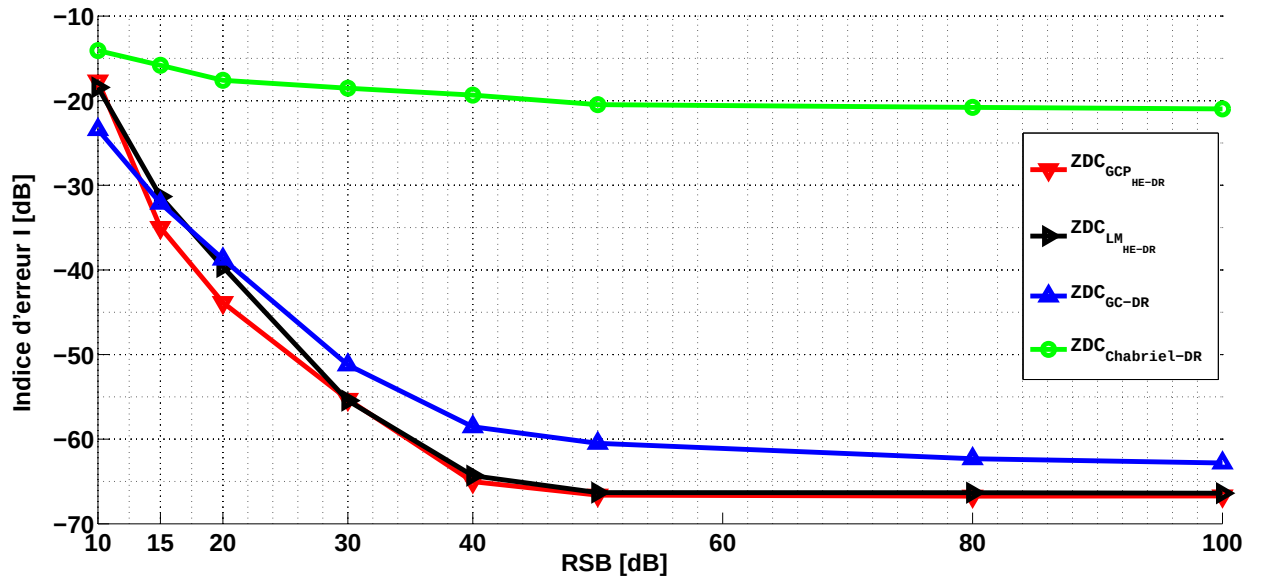


FIGURE 3.11 – I en dB en fonction du RSB. Comparaison des algorithmes proposés ZDC_{GC-DR} , $ZDC_{LM_{HE-DR}}$, $ZDC_{GCP_{HE-DR}}$ à $ZDC_{Chabriel-DR}$.

aux indéterminations classiques de la séparation de sources -. Les approches proposées possèdent l'avantage d'être assez générales en ce sens qu'elles peuvent s'étendre à d'autres applications pour lesquelles les sources peuvent aussi bien être réelles ou complexes, déterministes ou aléatoires, voire corrélées du moment qu'il s'agit de signaux non-stationnaires ou cyclo-stationnaires. Par ailleurs, les méthodes proposées ici pourraient s'étendre à des modèles de mélange convolutif.

Chapitre 4

Nouvelles méthodes de séparation en surveillance maritime par satellite : cas sous-déterminé

Dans ce chapitre, nous nous intéressons à des familles d’approches distinctes de celles présentées au niveau du Chapitre 3 afin de remédier au problème de collision des sources $S - AIS$. Les approches proposées dans ce chapitre ne reposent pas toutes sur les mêmes hypothèses en ce qui concerne les signaux reçus et ne s’appliquent donc pas toutes dans les mêmes contextes, ainsi le nombre de capteurs dont on dispose pourra varier, de même que le mode opératoire qui pourra être soit un mode semi-supervisé avec utilisation de trames d’apprentissage et information a priori soit un mode aveugle. On pourra également tenir compte ou non de problèmes liés à la synchronisation des signaux, etc Après avoir étudié des approches en contexte sur-déterminé, ce chapitre est quant à lui consacré au cas sous-déterminé dans lequel on considère qu’il y a davantage de sources que de capteurs. Nous exploitons donc d’autres techniques de traitement afin de résoudre le problème de collision des signaux. Nous considérons plus spécifiquement le cas où plusieurs sources $S - AIS$ arrivent sur un seul capteur ($m < n$ avec $m = 1$), il s’agit d’un modèle multi-entrées une seule-sortie (MESS³⁸). Pour traiter ce problème, nous ciblons plus particulièrement les approches basées à la fois sur la détection multi-utilisateurs conjointe (DMUC) et séquentielle (DMUS). Nous commençons par développer une méthode de type déflation connue sous le nom de suppression successive des interférences (SSI³⁹) que nous noterons DMUS_{SSI}. Les approches de type déflation ne sont plus conjointes, comme nous l’avons déjà mentionné au niveau du Chapitre 1, *i.e.* elles n’ont plus pour objectif de séparer tous les signaux simultanément mais elles les extraient les uns après les autres. Le principe de l’approche présentée ici sera de chercher à estimer le signal de plus forte énergie, de le soustraire ensuite des observations reçues au niveau du capteur, puis de continuer à estimer les différents signaux en fonction de ce schéma itératif répétitif. Dans un second temps, nous présentons une seconde méthode baptisée DMUC_{JMLSE} basée sur l’estimateur de séquence à maximum de vraisemblance conjoint (JMLSE⁴⁰) qui n’est rien d’autre qu’une variante de l’algorithme de VITERBI. Cette approche repose sur l’utilisation d’un algorithme adaptatif de moindres carrés récursifs. Enfin, après la présentation de chaque approche, nous effectuerons des simulations informatiques afin d’évaluer leur comportement toujours dans un cadre de surveillance maritime par satellite.

38. En anglais MISO pour Multiple Input Single Output.

39. En anglais SIC pour Successive Interference Cancellation.

40. En anglais Joint Maximum Likelihood Sequence Estimation.

4.1 Technique de dé-collision fondée sur la méthode de suppression successive des interférences SSI

4.1.1 Principe de l'approche DMUS_{SSI}

La première approche que nous allons considérer pour résoudre le problème de la collision des signaux **S – AIS** (utilisateurs) est fondée sur la méthode de la suppression successive des interférences (SSI). Ce type de méthodes requiert l'estimation de différents paramètres afin de procéder à l'atténuation des interférences. Elle fusionne l'algorithme SSI traditionnel [79] à des estimations de paramètres à savoir i) amplitude, ii) retard iii) et effet Doppler. Dans le schéma de détection DMUS_{SSI} présenté au niveau de la Fig. 4.1, nous commençons par analyser les signaux **S – AIS** (utilisateurs) par rapport à leur niveau de puissance. Le signal d'information de l'utilisateur (noté User dans la Fig. 4.1) dont la puissance est la plus élevée est estimé, démodulé et retranché des observations reçues au niveau d'un seul capteur. Nous appliquons la même procédure pour les $n - 1$ utilisateurs restants. Néanmoins, nous signalons que ce genre d'approche reste sous-optimal, dans la mesure où le dernier utilisateur (signal AIS) bénéficie de la réduction des interférences causées par les autres utilisateurs. Toutefois lorsque l'estimation de l'interférence provoquée par les autres utilisateurs n'est pas très précise, cela peut constituer une source d'erreurs qui risque de se propager tout le long de la boucle d'annulation d'interférences. Enfin, quand le nombre d'utilisateurs est élevé, le traitement doit être suffisamment rapide pour éviter toutes sortes de retards qui peuvent limiter l'approche en terme de performances. Finalement pour pouvoir réaliser la phase de démodulation, l'estimation des paramètres d'entrée propres à chaque signal est nécessaire. La technique d'estimation de chaque paramètre est fourni au niveau du tableau Tab. 4.1 suivant :

Paramètre	Technique d'estimation
Doppler Δf_j	Nous utilisons l'estimateur aveugle conjoint du décalage de fréquence de P. Bianchi et al. [29] interprété comme une généralisation de l'estimateur présenté dans [28].
Retard τ_j	– À base de la connaissance de la séquence d'apprentissage modulée en GMSK suivant la norme S – AIS. – Recherche du maximum de la fonction d'inter-corrélation entre la séquence d'apprentissage modulée et le signal reçu $x(t)$.
Amplitude A_j	– À base de la connaissance de la séquence d'apprentissage d'amplitude a_{app_j} – $\tilde{A}_j = \sum_{j=1}^J \frac{ y[k-\tau_j]a_{app_j} }{ a_{app_j} ^2}$

TABLE 4.1 – Estimation des paramètres d'entrée : effet Doppler Δf_j , retard τ_j et amplitude A_j

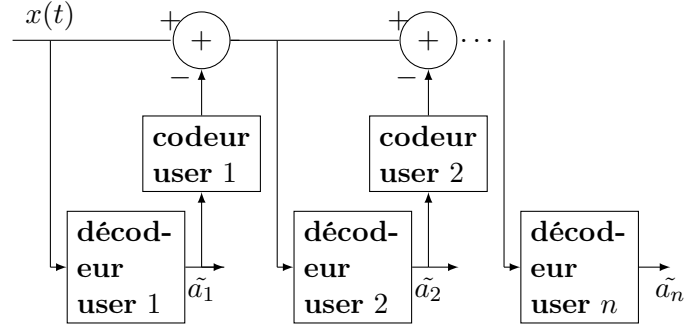


FIGURE 4.1 – Séquenceur SSI pour soustraction, démodulation et estimation successive des sources S – AIS.

4.1.2 Approche DMUS_{SSI} en résumé

Toutes les étapes de l'approche basée sur l'utilisation de SSI, que nous notons DMUS_{SSI}, sont récapitulées ci-dessous :

- 1) Estimation de l'utilisateur (source S – AIS) de la puissance la plus élevée,
- 2) Estimation des paramètres Δf_j , τ_j et A_j de l'utilisateur (source S – AIS) estimé suivant le Tab. 4.1,
- 3) Soustraction des observations de l'utilisateur (source S – AIS) de puissance la plus forte et construction de nouvelles observations,
- 4) Processus inverse notamment la démodulation GMSK classique adaptée à S – AIS de l'utilisateur (source S – AIS) considéré comme le signal le plus fort,
- 5) Estimation de la trame S – AIS de l'utilisateur de puissance la plus forte.

Nous rappelons qu'il s'agit d'un processus itératif qui permet de récupérer les trames S – AIS de tous les utilisateurs. En outre, cette approche basée sur l'utilisation de SSI a fait l'objet d'une publication [60].

4.1.3 Simulations informatiques

Nous présentons dans ce paragraphe des simulations numériques afin d'illustrer l'efficacité et les performances de la technique de dé-collision proposée, notée DMUS_{SSI}. Pour ce faire, nous rappelons tout d'abord que les trames sont générées conformément à la recommandation du système S – AIS : $BT_s = 0.4$ avec $T_s = 1/9600$ s et l'ordre du filtre Gaussien est choisi égal à $OF = 21$. Toutefois, le signal reçu au niveau du seul capteur dont nous supposons pouvoir disposer ($m = 1$) est ré-échantillonné d'un facteur $FSE = 4$ (facteur de sur-échantillonnage utilisé pour augmenter la probabilité d'estimation). Nous considérons en premier lieu, que deux sources arrivent sur un même capteur en créant une collision. En effet, la seconde source subit un retard de 0.0260 ms par rapport à la première source et la fréquence Doppler de la seconde source est de 100 Hz. Le rapport de puissance entre les deux sources est de 10%. Nous mentionnons que tous les résultats affichés résultent du moyennage sur 10 tirages par méthode de Monte-Carlo. Nous considérons ensuite le cas où trois sources S – AIS arrivent en collision sur le même capteur. La seconde source arrive avec un retard de 0.0026 ms et la troisième arrive avec un retard de

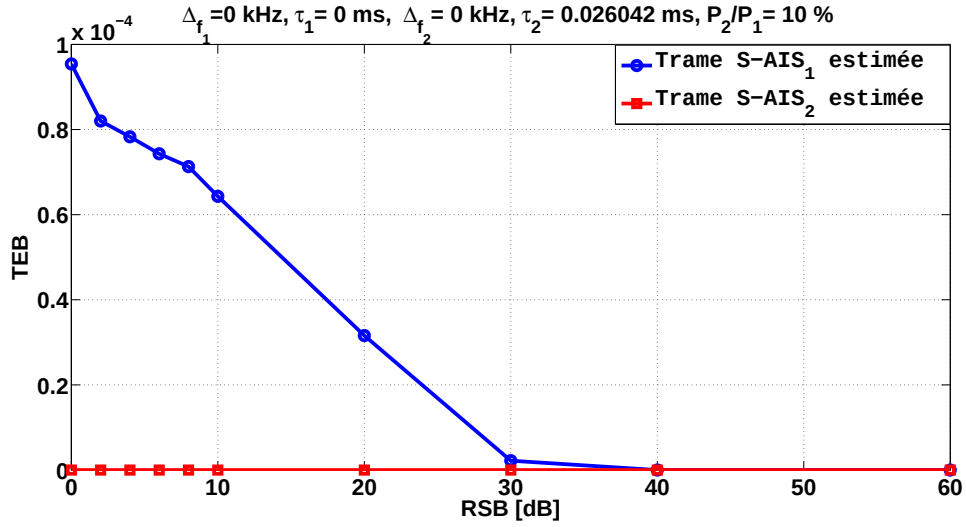


FIGURE 4.2 – Comportement de l’algorithme DMUS_{SSI} : taux d’erreur binaire (TEB) en fonction du RSB. Trame 1 avec retard de 0 ms et Doppler de 0 kHz et trame 2 avec retard de 0.0260 ms et Doppler de 100 Hz.

0.0521 ms par rapport à la première source. Par ailleurs, la fréquence Doppler de la seconde source est de 1 kHz et celle de la troisième source est de 3 kHz. Le rapport de puissance entre la troisième (resp. deuxième) et la première source est de 10% (resp. 20%).

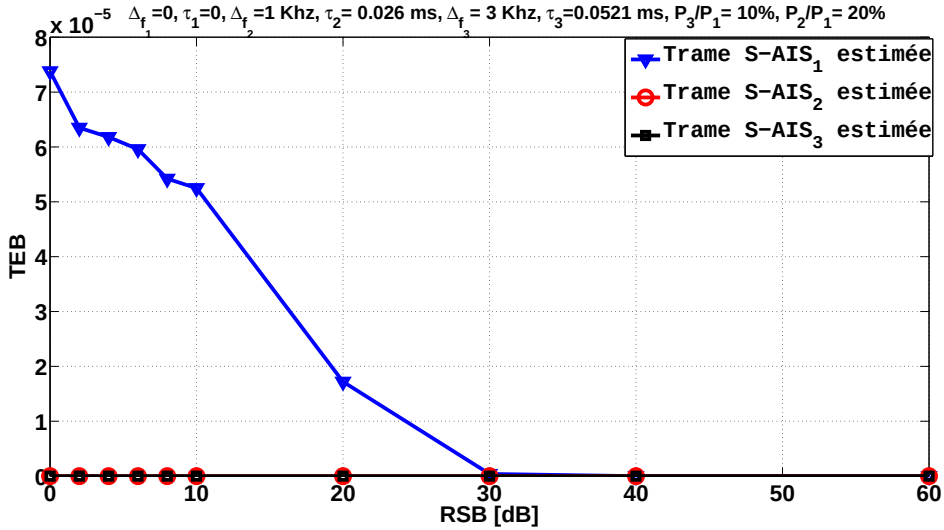


FIGURE 4.3 – Comportement de l’algorithme DMUS_{SSI} : TEB en fonction du RSB : la source 1 avec retard de 0 ms et Doppler de 0 kHz, la source 2 avec retard de 0.0260 ms et un Doppler de 1 kHz et la source 3 avec un retard de 0.0521 ms et un Doppler de 3 kHz.

Nous traçons, au niveau des deux Fig. 4.2 et 4.3, le taux d’erreur binaire (TEB) en fonction du rapport signal à bruit. Nous notons au niveau de la Fig. 4.2 que la deuxième trame S – AIS est estimée sans aucun bit erroné pour n’importe quelle valeur du RSB. Nous notons également qu’à

partir d'un RSB supérieur à 30 dB, la première trame/séquence est elle aussi estimée sans le moindre bit erroné. Toutefois, pour un RSB = 2 dB, nous avons un seul bit erroné sur 8.2×10^5 bits. En outre, nous constatons aussi que la dernière trame profite bel et bien de la réduction totale des interférences causée par la deuxième.

Passons maintenant à la Fig. 4.3. Nous observons que la deuxième et la troisième trame sont estimées sans le moindre bit erroné, tandis que la première source n'est estimée sans le moindre bit erroné qu'à partir d'un certain niveau de RSB (ici supérieur à 30 dB). Néanmoins, pour un RSB = 10 dB, nous n'obtenons qu'un seul bit erroné sur 5.25×10^5 bits.

Au regard des résultats obtenus sur ces deux simulations, nous pouvons conclure que l'approche DMUS_{SS1} peut être une solution pour la dé-collision (éventuellement la séparation) des trames (sources) S – AIS. Toutefois, cette technique atteint ses limites lorsque les sources possèdent la même puissance ou des puissances très proches. En pratique, il y a quand même assez peu de chances que cette problématique se pose, les différentes sources arrivant souvent avec des puissances distinctes.

4.2 Technique de dé-collision à base de l'algorithme de JMLSE

4.2.1 Principe de l'approche DMUC_{JMLSE}

Dans cette partie, nous proposons une autre approche de type DMUC baptisée DMUC_{JMLSE} afin de traiter le problème de la séparation (dé-collision) des sources S – AIS toujours dans le cas sous-déterminé où un seul capteur est utilisé. Cette méthode est fondée sur l'utilisation d'un estimateur de type JMLSE (Joint Maximum Likelihood Sequence Estimation) proposé en 1972 par G.D. Forney [99]. Le JMLSE repose sur le même principe que l'algorithme de VITERBI [234]. Il permet toutefois de simplifier la complexité de la recherche de la séquence d'origine la plus probable ayant produit la séquence traitée dans un treillis⁴¹. L'algorithme JMLSE est désormais la simple forme de l'égaliseur de VITERBI qui n'est rien d'autre qu'une version conjointe du MLSE. En outre, cet algorithme peut être le meilleur choix en présence d'interférences inter-symboles finies et d'un bruit blanc Gaussien (même dans un contexte asynchrone).

Pour assurer le bon fonctionnement de l'algorithme JMLSE, un certain nombre d'hypothèses sont nécessaires. En effet :

- Hyp₁** Le nombre de sources S – AIS doit être connu,
- Hyp₂** Les trames ne doivent pas être fortement corrélées,
- Hyp₃** Le bruit additif est un bruit blanc Gaussien centré,
- Hyp₄** Les sources S – AIS émises sont indépendantes du bruit,
- Hyp₅** Pour augmenter la probabilité d'estimation, nous avons choisi de ré-échantillonner les observations $x(t)$ d'un facteur FSE = 4,
- Hyp₆** Le nombre de coefficient du filtre FIR modélisant le canal est connu.

Pour faciliter le développement et les calculs, nous avons opté pour un modèle discret, $\tau_j = 0$, $\Delta f_j = 0$ et le signal est en bande de base. Toutefois, la problématique est reformulée et exprimée par l'Eq. 4.1 au lieu de l'Eq. 1.48. En effet :

$$x[k] = \sum_{j=1}^J h_j[k] a_j[k] + n[k], \forall k = 0, \dots, K, \quad (4.1)$$

41. Ensemble partiellement ordonné dans lequel chaque couple d'éléments admet une borne supérieure et une borne inférieure.

avec K est le nombre d'échantillons reçus au niveau du capteur, $h_j[k]$ est la réponse impulsionnelle du canal j et a_j est la séquence transmise en bande de base par le $j^{\text{ème}}$ utilisateur. L'objectif majeur est de retrouver un maximum de vraisemblance c'est à dire chercher le chemin optimal dans un treillis. Ce processus consiste alors à estimer la séquence $(\hat{\mathbf{a}}_1, \dots, \hat{\mathbf{a}}_n)$ de $(a_1, \dots, a_n)$ des J utilisateurs à partir des observations x . Cet estimateur maximise bien évidemment la fonction f ⁴² :

$$(\hat{\mathbf{a}}_1, \dots, \hat{\mathbf{a}}_n) = \arg \max_{(\tilde{\mathbf{a}}_1, \dots, \tilde{\mathbf{a}}_n) \in \mathcal{C}^J} \{f(x / (\tilde{\mathbf{a}}_1, \dots, \tilde{\mathbf{a}}_n), h_j)\}, \quad (4.2)$$

avec $\mathcal{C}^J$ représente toutes les possibilités des J séquences tirées d'un alphabet $\mathcal{C}$, $\hat{\mathbf{c}}_j = [\hat{c}_j[0], \dots, \hat{c}_j[K]]^T$ avec $j = 1, \dots, J$ est un vecteur regroupant les estimées $(\hat{\mathbf{a}}_1, \dots, \hat{\mathbf{a}}_n)$ de $(a_1, \dots, a_n)$ de la séquence j . Ensuite, en développant mathématiquement l'Eq. 4.1, le problème nous conduit alors à une minimisation au sens de moindres carrées définie comme suit :

$$(\hat{\mathbf{a}}_1, \dots, \hat{\mathbf{a}}_n) = \arg \max_{(\tilde{\mathbf{a}}_1, \dots, \tilde{\mathbf{a}}_n)} \left\{ \sum_{k=0}^K \gamma[k] \right\}, \quad (4.3)$$

avec $\gamma[k] = |x[k] - \sum_{j=1}^J h_j[k] \tilde{\mathbf{a}}_j|^2$ représente la métrique de branche.

Pour atteindre l'objectif de l'approche DMUC_{JMLSE}, nous devons chercher la séquence de métrique la plus faible. La métrique du chemin est définie comme :

$$\xi[k+1] = \xi[k] + \gamma[k+1], \quad (4.4)$$

où $\xi[k]$ est la métrique du chemin subsistant de l'étape précédente dans le treillis.

Nous signalons que cette méthode est très sensible aux erreurs d'estimation du canal de transmission. On trouve dans la littérature deux catégories d'algorithmes pour estimer le canal :

- Algorithmes utilisant une séquence d'apprentissage (préambule). Cette méthode s'avère moins performante lorsque l'on traite un contexte asynchrone dans lesquels les préambules des différents utilisateurs sont souvent identiques,
- Algorithmes traitant l'estimation en aveugle.

Nous proposons dans DMUC_{JMLSE} de traiter de manière semi-aveugle le problème de l'estimation du canal, c.à.d nous n'exploiterons la séquence d'apprentissage que pour détecter l'utilisateur de plus forte énergie. De la même manière, l'estimation du canal peut être réalisée éventuellement par la méthode des moindres carrées. Les canaux des autres utilisateurs sont initialisés selon l'expression suivante :

$$\hat{\mathbf{h}}_j = [0, 0, \dots, C, 0, \dots, 0], \text{ pour } 2 \leq j \leq J \quad (4.5)$$

où C est une constante réelle qui dépend du modèle du canal choisi.

Pour procéder à l'estimation des séquences émises, nous considérons par exemple, $a_{j,\sigma,\rho}[k+1]$, $\forall j = 1, \dots, J$. L'hypothèse sur la séquence émise à $k+1$ correspond à la transition de l'état σ à ρ pour l'utilisateur j . La fonction métrique notée par $\gamma_{\sigma,\rho}[k+1]$ permettant de passer de l'état σ à ρ est donnée par :

$$\gamma_{\sigma,\rho}[k+1] = \left| x[k] - \hat{\mathbf{h}}_j^T[k] \tilde{a}_{j,\sigma,\rho}[k+1] \right|^2, \forall k = 1, \dots, K. \quad (4.6)$$

42. La densité de probabilité du signal discret reçu en bande de base exprimée par $f(x / (\tilde{\mathbf{a}}_1, \dots, \tilde{\mathbf{a}}_n), h_j)$ avec $(\tilde{\mathbf{a}}_1, \dots, \tilde{\mathbf{a}}_n)$ sont les hypothèses possibles sur les symboles transmis.

Comme nous l'avons mentionné précédemment, nous avons intérêt à déterminer la métrique de branche $\xi_{\sigma,\rho}[k+1]$. En effet, la métrique des chemins est donnée par l'expression suivante :

$$\xi_{\sigma,\rho}[k+1] = \xi_{\sigma,\rho}[k] + \gamma_{\sigma,\rho}[k+1], \quad (4.7)$$

où $\xi_{\sigma,\rho}[k+1]$ est obtenue par la minimisation de $\xi_{\sigma,\rho}[k]$.

Par la suite, la séquence est toutefois estimée par $\tilde{a}_j[k-\delta]$ avec δ est une constante décisive pour les performances du système utilisé. Enfin, le vecteur $\hat{\mathbf{h}}_j$ de chaque canal j à l'instant $k+1$ est estimé au moyen de l'algorithme adaptatif LMS⁴³ [243] dont le schéma d'optimisation est donné ci-dessous :

$$\hat{\mathbf{h}}_j[k+1] = \hat{\mathbf{h}}_j[k] + \alpha[k]e[k]\tilde{a}_j[k-\delta], \quad (4.8)$$

où $\alpha[k]$ est le pas optimal d'adaptation de l'algorithme LMS et $e[k]$ qui représente l'erreur d'estimation est donnée par $e[k] = y[k] - \sum_{j=0}^J \hat{\mathbf{h}}_j^T[k]a_j[k-\delta]$.

4.2.2 Approche DMUC_{JMLSE} en résumé

Toutes les étapes de l'approche DMUC_{JMLSE} fondée sur l'algorithme de JMLSE sont récapitulées ci-dessous :

- 1) Considérer les échantillons du signal reçu $x[k]$,
- 2) Initialiser le canal $\hat{\mathbf{h}}_1$ de l'émetteur le plus fort par la méthode des moindres carrés et les autres canaux sont initialisés par l'Eq. (4.5),
- 3) Initialiser pour tous les états σ, ρ les hypothèses sur la séquence émise à $k=0$,
- 4) Pour $k=1, \dots, K$, et tous les états σ, ρ :
- 5) Calculer la métrique de branche à l'aide de l'Eq.(4.6),
- 6) Calculer la métrique des chemins à l'aide de l'Eq. (4.7),
- 7) Déterminer le meilleur chemin pour chaque état ρ par $\lambda_{\rho}(k+1) = \min_{\sigma} \{\lambda_{\sigma,\rho}(k+1)\}$,
- 8) Remonter pour obtenir les estimées $\tilde{\mathbf{a}}_j[k-\delta]$,
- 9) Mettre à jour les coefficients du canal selon le schéma décrit au niveau de l'Eq. (4.8).

Nous notons que cette approche basée fondée sur l'algorithme de JMLSE a fait l'objet de publications [58][59].

4.2.3 Simulations informatiques

4.2.3.1 Exemple didactique simple

Commençons par un exemple didactique simple afin de bien comprendre le principe de l'approche DMUC_{JMLSE} proposée. Pour ce faire, nous choisissons de considérer trois utilisateurs $J=3$ reliés à un canal de type instantané. La Fig. 4.4 présente les trois séquences émises qui subiront ensuite le traitement proposé au niveau de la méthode DMUC_{JMLSE}. Pour créer les observations, nous utilisons un canal généré d'une manière aléatoire selon une distribution Gaussienne. L'objectif de cet exemple est de montrer les différentes métriques de branches : ces dernières sont calculées et présentées dans le Tab. 4.2 tout en supposant que le canal utilisé est connu dans un premier temps.

Nous remarquons bien que le treillis rassemblé par les zéros obtenus de chaque échantillon et chaque état, comme illustré dans la Fig. 4.5, estime exactement les séquences émises présentées au niveau de la Fig. 4.4.

43. En anglais Least Mean Squares.

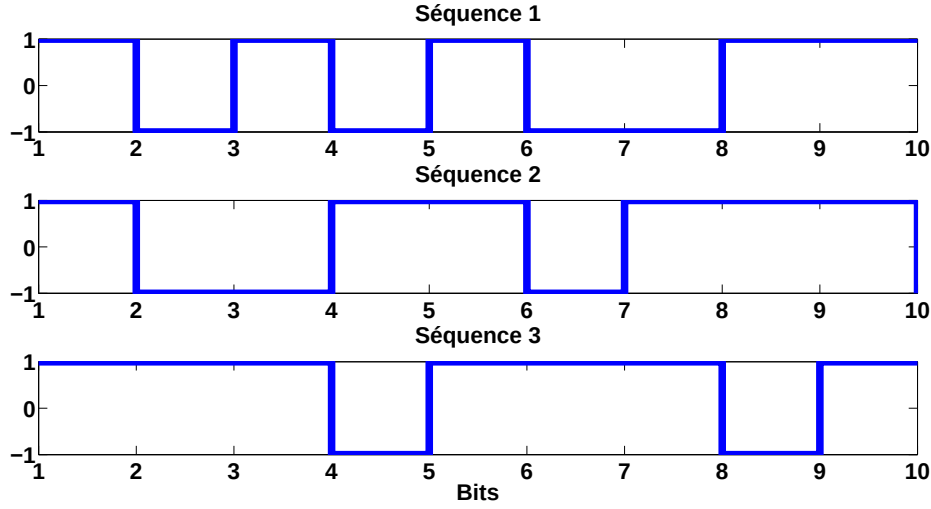
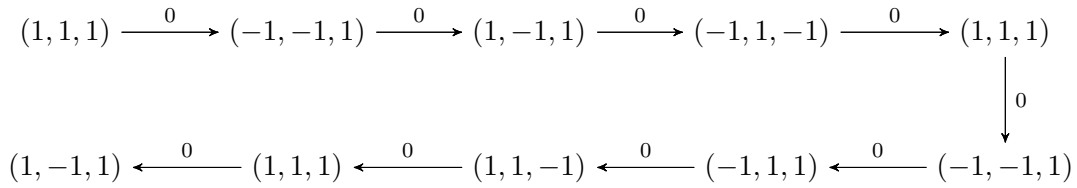


FIGURE 4.4 – Les trois séquences émises de la cadre de l'exemple didactique.

État	(1, 1, 1)	(1, -1, -1)	(-1, -1, 1)	(-1, 1, -1)	(-1, 1, 1)	(-1, -1, -1)	(1, -1, 1)	(1, 1, -1)
k=0	0	9.1845	16.3240	61.7856	19.6696	55.7358	0.1558	11.7330
k=1	7.3218	14.3754	0	14.7022	1.8889	6.0515	1.7729	26.6861
k=2	4.8218	0.7739	5.6798	0.4793	0.0351	2.2606	0	9.4591
k=3	0.7645	6.7236	1.3716	0	5.3789	1.3181	0.0749	2.0877
k=4	0	28.0853	0.0326	0.8481	9.1205	5.1963	10.2432	4.4061
k=5	8.3844	19.3185	0	0.0001	13.3077	13.2706	0.5661	0.5585
k=6	9.1304	7.3883	0.1312	0.0034	0	0.0921	7.0724	9.4888
k=7	0.0001	1.5925	0.0867	0.9512	0.9359	0.0822	1.6125	0
k=8	0	0.3455	0.1723	16.1786	4.3994	2.2790	6.3132	3.7048
k=9	0.0474	36.2064	2.8948	16.7935	3.6833	18.6259	0	33.6329

TABLE 4.2 – Métriques de branches calculées dans l'exemple didactique utilisant l'approche $DMUC_{JMLSE}$.FIGURE 4.5 – Approche $DMUC_{JMLSE}$: treillis construit de l'exemple explicatif.

4.2.3.2 Cas de signaux S – AIS en bande de base

Nous allons illustrer le comportement de la technique de $DMUC$, fondée sur l'algorithme $JMLSE$ en contexte instantané. Les sources S – AIS considérées en bande de base sont perturbées par un bruit blanc Gaussien. Elles sont supposées arriver en collision sur un seul capteur ($m = 1$) suivant l'Eq. (4.1) et nous les séparons à l'aide de l'approche $DMUC_{JMLSE}$. Tous les résultats

affichés proviennent du moyennage sur 100 tirages par méthode de Monte-Carlo.

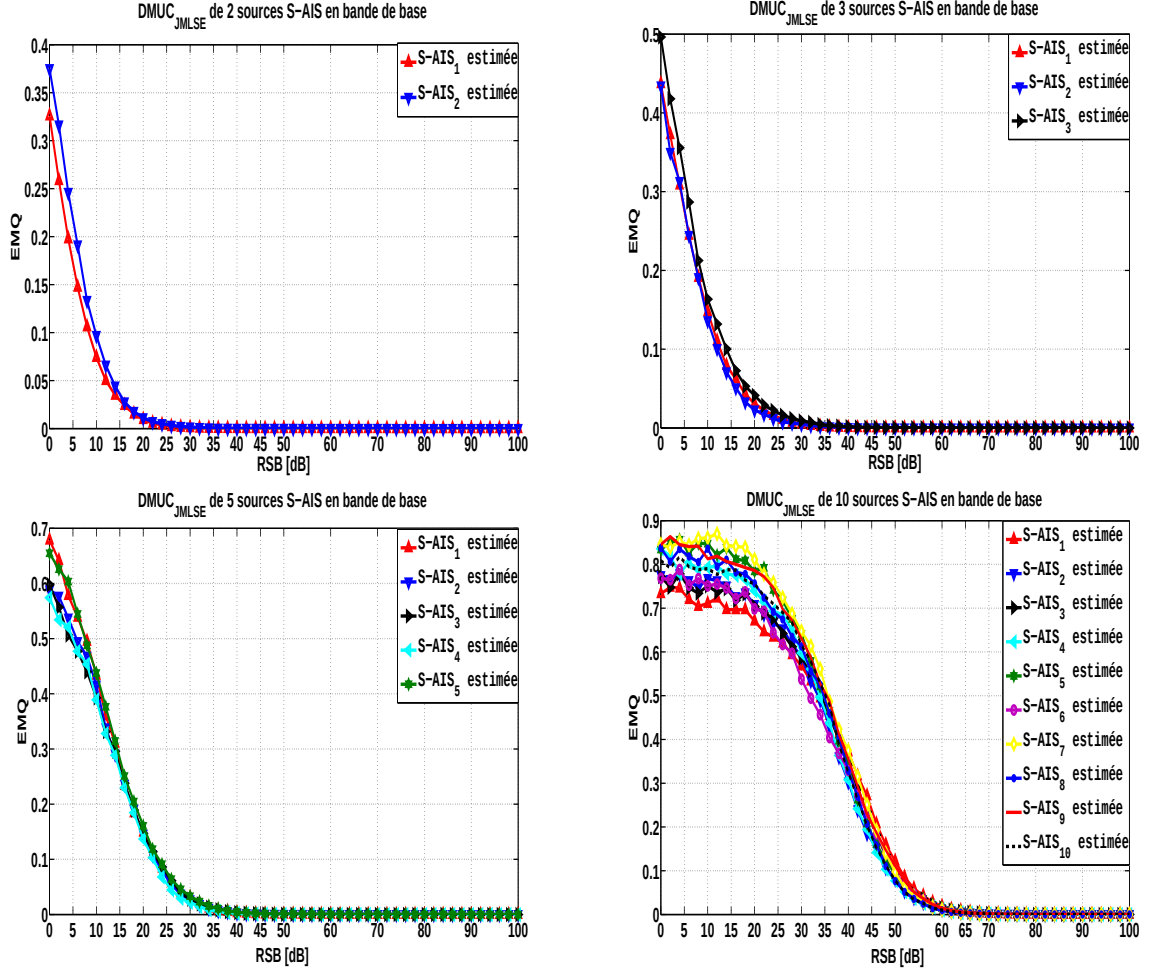


FIGURE 4.6 – Comportement de l'algorithme DMUC_{JMLSE} : EMQ en fonction du RSB. À droite : 3 sources S – AIS en bande de base en haut et 10 sources S – AIS en bande de base en bas. À gauche : 2 sources S – AIS en bande de base en haut et 5 sources S – AIS en bande de base en bas.

Nous traçons au niveau de la Fig. 4.6, l'erreur quadratique moyenne (EQM⁴⁴) des sources originales par rapport aux estimées en fonction du rapport signal à bruit (RSB). Nous constatons que pour des faibles valeurs du RSB, l'EQM est importante. Toutefois, quand le RSB augmente, les performances s'améliorent. La différence entre les quatre courbes de la Fig. 4.6 réside dans le nombre de sources S – AIS à estimer. Plus précisément, pour la courbe de gauche en haut, nous observons que l'EQM des 2 sources est nulle une fois que le RSB est supérieur 20 dB, alors que dans le cas où l'on estime 3 (resp. 5 et 10) sources S – AIS, l'EQM n'est nulle qu'au delà de 30 dB (resp. 35 dB et 60 dB). Les performances de l'approche DMUC_{JMLSE} dépendent considérablement du nombre de sources (utilisateurs) à estimer.

44. En anglais MSE pour Mean Square Error.

4.2.3.3 Cas des signaux S – AIS avec différents retards et Dopplers

Nous étudions maintenant le comportement de la technique de $DMUC_{JMLSE}$ en considérant le cas de deux et cinq sources S – AIS ($n = 2$ et 4) avec des retards et des Dopplers différents. Ces sources sont perturbées par un bruit blanc Gaussien suivant l'Eq. (1.48) et arrivent sur un seul capteur ($m = 1$). Les sources sont toujours générées conformément à la recommandation S – AIS [162] et suivant la Fig. 1.3, codées en NRZI et modulées en GMSK avec le produit temps bande spécifié de $BT_s = 0.4$ avec $T_s = 1/9600$ s. L'ordre du filtre Gaussien est choisi égal à $OF = 21$ [263][196]. En outre, la fréquence porteuse de 161,975 MHz des sources S – AIS est supposée être la même. La fréquence Doppler et le retard différent d'une source à l'autre. En effet, la première source subit un retard de 10 ms et un décalage Doppler de -4 kHz, la deuxième est affectée par un Doppler de 2 kHz et un retard de 5 ms, la troisième arrive avec un retard de 1 ms et un Doppler de 1 kHz, par contre la dernière est de 0 ms (retard) et 0 (Doppler) Hz. Tous les résultats sont moyennés sur 100 réalisations par méthode de Monte-Carlo.

Nous traçons au niveau de la Fig. 4.7, l'EQM des séquences émises par rapport aux séquences estimées en fonction du rapport signal à bruit (RSB). Pour les deux courbes de la Fig 4.7, nous constatons que pour les faibles valeurs du RSB, l'EQM est assez importante. Avec l'amélioration du RSB, les performances elles aussi s'améliorent. En revanche, pour la courbe de droite (resp. la courbe de gauche), nous constatons que l'EQM est nulle à partir de 90 dB (resp. 80 dB). En comparant ce point avec les résultats présentés en bande de base, nous constatons que la méthode $DMUC_{JMLSE}$ que nous avons proposée est très sensible au Doppler que l'on estime selon le Tab. 4.1 présenté précédemment. De manière générale, l'approche $DMUC_{JMLSE}$ se révèle relativement robuste par rapport au bruit et l'amplitude.

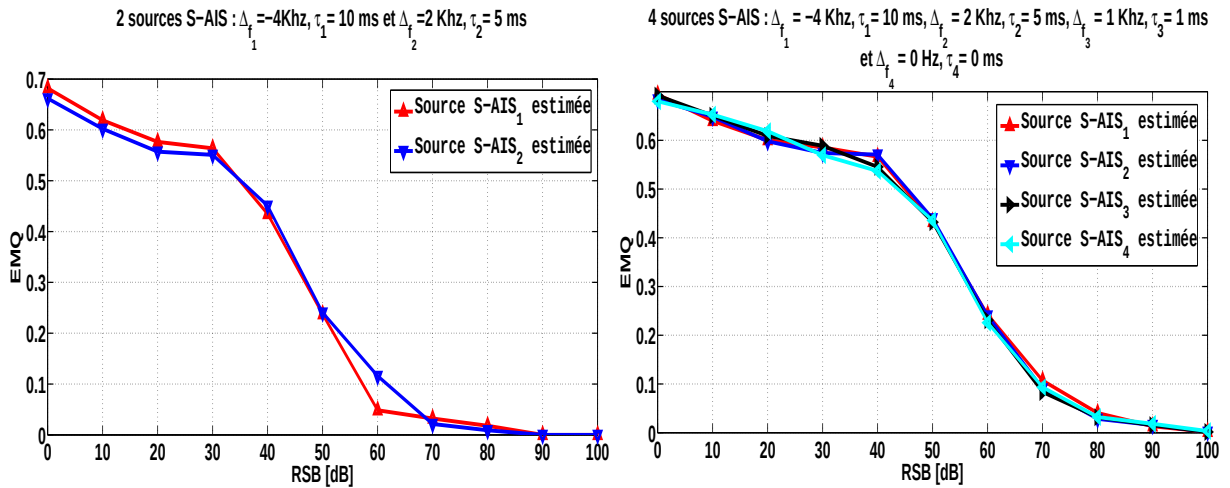


FIGURE 4.7 – Comportement de l'algorithme $DMUC_{JMLSE}$: EQM en fonction du RSB. À droite : 4 sources S – AIS. À gauche : 2 sources S – AIS.

4.3 Conclusion

Dans ce chapitre, nous avons considéré de nouvelles approches en réponse au problème de la collision de sources S – AIS. Nous nous sommes cette fois-ci placé en contexte sous-déterminé

($m < n$) et les signaux **S – AIS** sont même supposés arriver au niveau d'un seul capteur ($m = 1$). Les approches développées ici sont basées principalement sur des techniques de détection multi-utilisateurs conjointe et séquentielle. Dans un premier temps, nous avons développé une technique séquentielle fondée sur la méthode de suppression successive des interférences qui a été notée **DMUS_{SSI}**. Elle consiste à estimer le signal le plus fort dans le but de le retrancher ensuite des observations puis à continuer à estimer les différents signaux selon ce même schéma itératif. Dans le même registre d'idées, nous avons présenté une méthode basée sur l'estimateur de séquences à maximum de vraisemblance conjoint noté **DMUC_{JMLSE}**. Ce n'est rien d'autre qu'une variante de l'algorithme de VITERBI qui recherche le chemin le plus vraisemblable à travers un treillis. Enfin, des simulations informatiques ont été réalisées afin d'illustrer le comportement de ces approches dans un contexte de surveillance maritime par satellite. Les premiers résultats obtenus au moyen de ces deux méthodes sont encourageants en ce qui concerne l'élimination du problème de la collision des signaux **S – AIS** dans le cas extrême multi-entrée unique-sortie où l'on ne dispose plus que d'un seul capteur pour observer et traiter les signaux.

Conclusions et perspectives

Conclusions

Dans cette thèse, nous nous sommes intéressés au système d'identification automatique spatial dédié à la surveillance maritime par satellite. Nous nous sommes plus particulièrement focalisés sur le problème de la collision des signaux **S – AIS**. Après avoir prouvé qu'il était possible de modéliser ce problème comme un problème de **SAS** en prenant en compte des paramètres Doppler et retard, nous avons ensuite montré qu'un tel problème pouvait être résolu au moyen de plusieurs approches distinctes selon les hypothèses dans lesquelles nous nous plaçons.

Dans un premier temps, nous avons eu recours aux méthodes de séparation aveugle de sources. Du fait de la nature aléatoire et cyclo-stationnaire des signaux considérés, nous nous sommes alors plus particulièrement intéressés aux méthodes fondées sur l'utilisation de distributions spatiales quadratiques des domaine temps-fréquence ou Doppler-retard. Ces méthodes reposent ensuite pour la plupart sur l'utilisation de décompositions matricielles conjointes particulières. C'est ce qui nous a conduit à nous intéresser à de nouvelles décompositions matricielles (notamment la "zéro-bloc diagonalisation conjointe non unitaire"). Nous nous sommes également intéressés aux méthodes du second ordre c'est à dire utilisant un pré-conditionneur afin de mieux mesurer l'impact de ce dernier au niveau des méthodes d'optimisation itératives que nous avons utilisées. Cela nous a conduit à proposer plusieurs nouveaux algorithmes que nous avons ensuite testés sur des données entièrement simulées puis sur des mélanges simulés de trames **AIS**. Dans un second temps, nous nous sommes orientés vers les méthodes opérant en modèle sous-déterminé en nous focalisant même sur le cas le plus compliqué à savoir celui où l'on ne dispose plus que d'un seul capteur pour recevoir les signaux (modèle de type multi-entrées sortie unique). Deux méthodes basées respectivement sur la détection multi-utilisateurs conjointe et la détection séquentielle ont alors été développées. La méthode séquentielle proposée est fondée sur un processus de suppression successive des interférences (approche de type déflation) : elle consiste à estimer le signal de plus forte énergie, à le retrancher des signaux reçus au niveau du capteur unique dont on dispose puis à reconstruire de nouvelles observations, et à continuer à estimer les différents signaux selon ce schéma itératif. La seconde est basée sur l'estimateur de séquences à maximum de vraisemblance conjoint. Toutes les méthodes développées semblent donner des résultats encourageants et prometteurs par rapport à l'application visée.

Perspectives de recherche

Après avoir montré l'efficacité de nos nouveaux algorithmes de **ZBDC_{NU}** et **BDC_{NU}**, il conviendrait de les appliquer à d'autres domaines applicatifs lesquels nécessiteraient l'utilisation d'un modèle de mélange convolutif de signaux. Nous avons également proposé au Chapitre 2 des

algorithmes pour résoudre un nouveau problème que nous avons choisi d'appeler “zéro-bloc diagonalisation conjointe” en ce sens qu'il généralise le problème de la zéro-diagonalisation conjointe. Il a d'ailleurs démontré son intérêt et son efficacité dans la résolution du problème désormais classique de la zéro-diagonalisation conjointe non-unitaire. L'objectif est maintenant de montrer son intérêt sur des applications pratiques concrètes et non plus sur des ensembles entièrement simulés. Nous pourrions alors développer de nouvelles méthodes de déconvolution multidimensionnelle (cas des mélanges linéaires convolutifs). Sur des approches à base de distributions spatiales quadratiques, il nous faudrait également réfléchir à l'élaboration de nouveaux détecteurs de points “utiles” permettant la construction d'ensembles de matrices à zéro-bloc diagonaliser conjointement.

Dans le cadre de notre application en surveillance maritime par satellite de nombreuses questions restent ouvertes. Elles concernent :

- l'amélioration de nos hypothèses sur le modèle considéré (il faudrait d'ailleurs pour cela disposer de données réelles par rapport auxquelles comparer nos modèles),
- l'hypothèse de la connaissance a priori du nombre de signaux émis dont nous aimerions pouvoir nous affranchir en utilisant par exemple des fonctions de coût pénalisées et d'autres méthodes d'optimisation,
- l'étude de l'apport des techniques de décompositions tensorielles dans ce type de problème,
- l'amélioration des cadences de traitement en basculant tous nos algorithmes sur leurs versions adaptatives ou encore en parallélisant certains.

Bibliographie

- [1] A. A. Belouchrani, M. Amin, N. Thirion-Moreau, and Y. Zhang. Source separation and localization using time-frequency distributions. *IEEE Signal Processing Magazine*, 30(6) :97–107, November 2013.
- [2] A. Abed-Meraïm, K. Belouchrani and R. Leyman. *Time-frequency signal analysis and processing : a comprehensive reference*, chapter Blind source separation using time-frequency distributions. Prentice-Hall, Editor : Boashash B., Oxford, UK, January 2003.
- [3] K. Abed-Meraïm, J.-F. Cardoso, A.Y. Gorokhov, P. Loubaton, and E. Moulines. On subspace methods for blind identification of single-input multiple-output fir systems. *Signal Processing, IEEE Transactions on*, 45(1) :42–55, Jan 1997.
- [4] K. Abed-Meraïm and Y. Hua. Combining cyclostationarity and spatial diversity for blind system identification. In *Signals, Systems and Computers, 1996. Conference Record of the Thirtieth Asilomar Conference on*, volume 1, pages 325–328, Nov 1996.
- [5] K. Abed-Meraïm, P. Loubaton, and E. Moulines. A subspace algorithm for certain blind identification problems. *Information Theory, IEEE Transactions on*, 43(2) :499–511, Mar 1997.
- [6] K. Abed-Meraïm, Y. Xiang, J.H. Manton, and Y. Hua. Blind source-separation using second-order cyclostationary statistics. *Signal Processing, IEEE Transactions on*, 49(4) :694–701, Apr 2001.
- [7] S. Achard, D.-T. Pham, and C. Jutten. Mesure de dépendance quadratique pour la séparation aveugle de source dans les mélanges post non linéaires. In *19ème Colloque du GRETSI, Traitement du Signal et des Images*, volume 1, pages 117–120, Paris, France, septembre 2003.
- [8] S. Achard, D.-T. Pham, and C. Jutten. Criteria based on mutual information minimization for blind source separation in post nonlinear mixtures. *Signal Process.*, 85(5) :965–974, May 2005.
- [9] L. Albera, A. Ferreol, P. Chevalier, and P. Comon. Icar, un algorithme d’ica à convergence rapide, robuste au bruit. In *in GRETSI 03, XIXeme colloque sur le Traitement du Signal et des Images*, volume 1, pages 193–196, 2003.
- [10] L. Albera, A. Ferreol, P. Chevalier, and P. Comon. Icar : Independent component analysis using redundancies,. In *ISCAS 04, IEEE International Symposium on Circuits and Systems, Vancouver*, 2004.
- [11] L. Albera, A. Ferreol, P. Chevalier, and P. Comon. Icar, a tool for blind source separation using fourth order statistics only. *IEEE Transactions On Signal Processing*, 53(10) :3633–3643, October 2005.

- [12] L. Albera, A. Ferréol, P. Comon, and P. Chevalier. Sixth order blind identification of underdetermined mixtures (birth) of sources. In *in Proc. Fourth Int. Symp. Independent Component Analysis Blind Signal Separation*, pages 909–914, 2003.
- [13] M. G. Amin. Blind time-frequency analysis for source discrimination in multisensor array processing. Technical report, Villanova University, Department of Electrical and Computer Engineering, October 1998.
- [14] M.G. Amin, A. Belouchrani, and Yimin Zhang. The spatial ambiguity function and its applications. *Signal Processing Letters, IEEE*, 7(6) :138–140, 2000.
- [15] B. Ans, J. Herault, and C. Jutten. Adaptative neural architectures : Detection of primitives. In *Proceedings of COGNITIVA 85*, pages 593–597, 1985.
- [16] J. Antoni. Cyclic spectral analysis of rolling-element bearing signals : Facts and fictions. *Journal of Sound and Vibration*, 304(3-5) :497–529, 2007.
- [17] J. Antoni, F. Bonnardot, A. Raad, and M. El Badaoui. Cyclostationary modelling of rotating machine vibration signals. *Mechanical Systems and Signal Processing*, 18(6) :1285–1314, 2004.
- [18] S. Araki, R. Mukai, S. Makino, T. Nishikawa, and H. Saruwatari. The fundamental limitation of frequency domain blind source separation for convolutive mixtures of speech. *Speech and Audio Processing, IEEE Transactions on*, 11(2) :109–116, Mar 2003.
- [19] O. Axelsson. A survey of preconditioned iterative methods for linear systems of algebraic equations. *BIT*, 25(1) :166–187, 1985.
- [20] M. Basseville, P. Flandrin, and N. Martin. Signaux nonstationnaires, analyse temps-fréquence et segmentation. *Fiches descriptives d’algorithmes Traitement du signal, Suppl.*, 9(1), 1992.
- [21] A. Belouchrani, K. Abed-Meraïm, M.G. Amin, and A.M. Zoubir. Joint anti-diagonalization for blind source separation. In *Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP ’01). 2001 IEEE International Conference on*, volume 5, pages 2789–2792, 2001.
- [22] A. Belouchrani, K. Abed-Meraïm, J.-F. Cardoso, and E. Moulines. A blind source separation technique using second-order statistics. *Signal Processing, IEEE Transactions on*, 45(2) :434–444, Feb 1997.
- [23] A. Belouchrani, M. G. Amin, and K. Abed-Meraïm. Direction finding in correlated noise fields based on joint block-diagonalization of spatio-temporal correlation matrices. *IEEE Signal Process. Lett.*, 4(9) :266–268, Sept. 1997.
- [24] A. Belouchrani, M. G. Amin, N. Thirion-Moreau, and Y. Zang. Source separation and localization using time-frequency distributions. *IEEE Signal Process. Mag.*, pages 97–107, Nov. 2013.
- [25] A. Belouchrani and M.G. Amin. Blind source separation based on time-frequency signal representations. *Signal Processing, IEEE Transactions on*, 46(11) :2888–2897, Nov 1998.
- [26] A. Belouchrani and A. Cichocki. Robust whitening procedure in blind source separation context. *IEEE Signal Processing Letter*, 23(24) :2050–2051, November 2000.
- [27] M. Benzi. Preconditioning techniques for large linear systems : A survey. *J. COMPUT. PHYS.*, 182 :418–477, 2002.
- [28] P. Bianchi, P. Loubaton, and F. Sirven. Non data aided estimation of the modulation index of continuous phase modulations. *IEEE Trans. Signal Processing*, 52(10) :2847–2861, 2000.

-
- [29] P. Bianchi, P. Loubaton, and F. Sirven. On the blind estimation of the parameters of continuous phase modulated signals. *Selected Areas in Communications, IEEE Journal on*, 23(5) :944–962, 2005.
 - [30] W. Bobillet, E. Grivel, R. Guidorzi, and M. Najim. Cancelling convolutive and additive noises for speech enhancement. In *EUROCON 2003. Computer as a Tool. The IEEE Region 8*, volume 2, pages 407–410, September 2003.
 - [31] F. Bonnardot. *Comparaison entre les analyses angulaire et temporelle des signaux vibratoires de machines tournantes. Etude du concept de cyclostationnarite floue*. PhD thesis, Institut National Polytechnique de Grenoble, Décembre 2004.
 - [32] A. Boonpoonga, S. Buritramart, P. Sirisuk, M. Krairiksh, and T.K. Sarkar. Adaptive array processing unit based on direct data domain least square approach using conjugate gradient method. In *Microwave Conference, 2007. APMC 2007. Asia-Pacific*, pages 1–4, Dec 2007.
 - [33] H. Bousbiah-Salah, A. Belouchrani, and Abed-Meraïm. Blind separation of non stationary sources using joint block diagonalization. In *Proc. Workshop on Statistical Signal Processing*, pages 448–451, 2001.
 - [34] H. Bousbiah-Salah, A. Belouchrani, and Abed-Meraïm. Jacobi-like algorithm for blind signal separation of convolutive mixtures. *IEEE Signal Processing Letter*, 37(16) :1049–1050, August 2001.
 - [35] R. Boustany and J. Antoni. A subspace method for the blind extraction of a cyclostationary source. In *Signal Processing Conference, 2005 13th European*, pages 1–4, Sept 2005.
 - [36] R. Boustany and J. Antoni. Blind extraction of a cyclostationary signal using reduced-rank cyclic regression : A unifying approach. *Mechanical Systems and Signal Processing*, 22(3) :520–541, April 2008.
 - [37] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, March 2004.
 - [38] J. Brewer. Kronecker products and matrix calculus in system theory. *IEEE Trans. on Circuits and Systems*, Cas-25(9) :772–781, September 1978.
 - [39] P. Burzigotti, A. Ginesi, and G. Colavolpe. Advanced receiver design for satellite-based ais signal detection. In *Advanced satellite multimedia systems conference (asma) and the 11th signal processing for space communications workshop (spsc), 2010 5th*, pages 1–8, Sept 2010.
 - [40] Dirk Callaerts. *Signal Separation Methods based on Singular Value Decomposition and their Application to the Real-Time Extraction of the Fetal Electrocardiogram from Cutaneous Recordings*. PhD thesis, K.U.Leuven - E.E. Dept., Dec. 1989.
 - [41] V. Capdevielle, Ch. Servière, and J.-L. Lacoume. Blind separation of wide-band sources : application to rotating machine signals. In *Proceeding of EUSIPCO 3*, pages 2085–2088, 1996.
 - [42] J.-F. Cardoso. On the performance of orthogonal sources separation algorithms. In *Proc. European Signal Processing Conference (EUSIPCO'94)*, pages 776–779, Edinburgh, Scotland, UK, September 1994.
 - [43] J.-F. Cardoso and A. Souloumiac. Blind beamforming for non-gaussian signals. *IEE Proceedings-F*, 40(6) :362–370, 1993.
 - [44] J.-F. Cardoso and A. Souloumiac. Jacobi angles for simultaneous diagonalization. *SIAM journal on Matrix Analysis and Applications*, 17(1) :161–164, January 1996.

- [45] M. Castella, A. Chevreuil, and J.-C. Pesquat. Séparation aveugle d'un mélange convolutif de sources non linéaires par une approche hiérarchique. In *GRETSI*, Paris, France, September 2003.
- [46] Marc Castella, S. Rhioui, E. Moreau, and J. Pesquet. Source separation by quadratic contrast functions : a blind approach based on any higher-order statistics. In *Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). IEEE International Conference on*, volume 3, pages iii/569–iii/572 Vol. 3, March 2005.
- [47] G. Chabriel and J. Barrère. Non-symmetrical joint zero-diagonalization and mimo zero-division multiple access. *IEEE Trans. Signal Proc.*, 59(5) :2296–2307, 2011.
- [48] G. Chabriel and J. Barrère. A direct algorithm for nonorthogonal approximate joint diagonalization. *IEEE Trans. Signal Proc.*, 60(1) :39–47, 2012.
- [49] G. Chabriel, J. Barrère, N. Thirion-Moreau, and E. Moreau. Algebraic joint zero-diagonalization and blind source separation. *IEEE Trans. Signal Proc.*, 56(3) :980–989, March 2008.
- [50] G Chabriel, M Kleinstember, E Moreau, H Shen, P Tichavsky, and A Yeredor. Joint matrices decompositions and blind source separation : A survey of methods, identification, and applications. *IEEE Signal Process. Mag.*, 31(3) :34–43, 2014.
- [51] M. Chabriel, G. and Kleinstember, E. Moreau, Hao Shen, P. Tichavsky, and A. Yeredor. Joint matrices decompositions and blind source separation : A survey of methods, identification, and applications. *IEEE Signal Processing Magazine*, 31(3) :34–43, May 2014.
- [52] N. Charkani. *Séparation auto-adaptative de sources pour les mélanges convolutifs. Application à la téléphonie mains-libres dans les voitures*. PhD thesis, INP Grenoble, Novembre 1996.
- [53] E. Chaumette, P. Comon, and D. Muller. Application of ica to airport surveillance. In *Higher-Order Statistics, 1993., IEEE Signal Processing Workshop on*, pages 210–214, 1993.
- [54] K. Chen. *Matrix Preconditioning Techniques and Applications (Cambridge Monographs on Applied and Computational Mathematics)*. 2005.
- [55] O. Cherrak, H. Ghennioui, E.-H. Abarkan, and N. Thirion-Moreau. Non-unitary joint block diagonalization of matrices using a levenberg-marquardt algorithm. In *Signal Processing Conference (EUSIPCO), 2013 Proceedings of the 21st European*, pages 1–5, Sept 2013.
- [56] O. Cherrak, H. Ghennioui, E.-H. Abarkan, and N. Thirion-Moreau. Nouvel algorithme de zéro-bloc diagonalisation conjointe par approche de levenberg-marquardt. In *GRETSI 2015*, Lyon, France, 08 september 2015.
- [57] O. Cherrak, H. Ghennioui, E.-H. Abarkan, N. Thirion-Moreau, and E. Moreau. Non-unitary joint zero-block diagonalization of matrices using a conjugate gradient algorithm. In *Signal Processing Conference (EUSIPCO), 2015 23rd European*, pages 1880–1884, Aug 2015.
- [58] O. Cherrak, H. Ghennioui, N. Thirion-Moreau, and E.-H. Abarkan. Joint maximum likelihood sequence estimation algorithm for decollision of ais signals in maritime surveillance context. *Mediterranean Telecommunication Journal*, 4(2) :14–19, 2014.
- [59] O. Cherrak, H. Ghennioui, N. Thirion-Moreau, and E.-H. Abarkan. Joint maximum likelihood sequence estimation algorithm for decollision of ais signals in maritime surveillance context. In *CMT 2014*, Mohammedia, Maroc, 22 Mai 2014.
- [60] O. Cherrak, H. Ghennioui, N. Thirion-Moreau, and E.-H. Abarkan. Successive interference cancellation technique for decollision of ais signals in maritime surveillance context by a leo satellite. In *WITS 2014*, Meknès, Maroc, 09 Avril 2014.

-
- [61] O. Cherrak, H. Ghennioui, N. Thirion-Moreau, and E.-H. Abarkan. Séparation aveugle de sources basée sur la bloc diagonalisation conjointe non-unitaire de matrices issues de transformées spatiales quadratiques : application a la surveillance maritime. In *TELECOM 2015 & 9ème JFMMA*, Meknès, Maroc, 13 Mai 2015.
 - [62] O. Cherrak, H. Ghennioui, N. Thirion-Moreau, and E.-H. Abarkan. Preconditioned optimization algorithms solving the problem of the non unitary joint block diagonalization : Application to blind separation of convolutive mixtures (en soumission). *Signal Process.*, 2016.
 - [63] S. Choi, A. Cichocki, and A. Belouchrani. Blind separation of second-order nonstationary and temporally colored sources. In *Statistical Signal Processing, 2001. Proceedings of the 11th IEEE Signal Processing Workshop on*, pages 444–447, 2001.
 - [64] A. Cichocki, R. Zdunek, A. H. Phan, and S. I. Amari. *Non negative matrix and tensor factorizations : Application to exploratory multi-way data analysis and blind separation*. Wiley, 2009.
 - [65] L. A. Cirillo, A. M. Zoubir, N. Ma, and M.G. Amin. Automatic classification of auto-and cross-terms of time-frequency distributions in antenna arrays. In *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on*, volume 2, pages 1445–1448, May 2002.
 - [66] L.A. Cirillo and M.G. Amin. Auto-term detection using time-frequency array processing. In *Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). 2003 IEEE International Conference on*, volume 6, pages 465–468, April 2003.
 - [67] Leon Cohen. *Time-frequency Analysis : Theory and Applications*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1995.
 - [68] P. Comon. Independant component analysis, a new concept ? *Signal Processing*, 36 :287–314, 1994.
 - [69] P Comon. Contrasts for multichannel blind deconvolution. *IEEE Signal Processing Letter*, 3(7) :209–211, July 1996.
 - [70] P. Comon. Tensors, a brief introduction. *IEEE Signal Processing Magazine*, 31(3) :44–53, May 2014.
 - [71] P. Comon and C. Jutten, editors. *Handbook of Blind Source Separation, Independent Component Analysis and Applications*. Academic Press, Oxford UK, Burlington USA, 2010.
 - [72] P. Comon and E. Moreau. Blind mimo equalization and joint diagonalization criteria. In *Proc. International Conference on Acoustic Speech and Signal Processing (ICASSP'01)*, volume 5, Salt Lake City, USA, May 2001.
 - [73] P. Comon and L. Rota. Blind separation of independent sources from convolutive mixtures. In *IEICE Trans. on Fundamentals of Electronics, Communications, and Computer Sciences*, pages 542–549, March 2003.
 - [74] Y. H. Dai and Y. Yuan. A nonlinear conjugate gradient method with a strong global convergence property. *SIAM J. Optim*, 10 :177–182, 1999.
 - [75] A. Dapena and L. Castedo. Blind source separation in the frequency domain : A novel solution to the amplitude and the permutation indeterminacies. In *Bio-inspired Applications of Connectionism, 6th International Work-Conference on Artificial and Natural Neural Networks, IWANN 2001 Granada, Spain, June 13-15, 2001, Proceedings, Part II*, pages 603–610, 2001.

- [76] A. Dapena and L. Castedo. A novel frequency domain approach for separating convolutive mixtures of temporally-white signals. *Digital Signal Processing*, 13(2) :301–316, 2003.
- [77] A. Dapena, D. Iglesia, and L. Castedo. Blind separation of cyclostationary signals. In *Signal Processing Conference (EUSIPCO 1998), 9th European*, pages 1–4, Sept 1998.
- [78] A. Dapena and C. Servière. Inversion of the sliding fourier transform using only two frequency bins and its application to source separation. *Signal Processing*, 83(2) :453–457, 2003.
- [79] R.-C. de Lamare, T.-T.-V. Vinhoza, and R. Sampaio-Neto. Blind adaptive and iterative interference cancellation receiver structures based on the constant modulus criterion in multipath channels. In *International Telecommunications Symposium (ITS'06)*, Fortaleza-CE, Brazil, September 2006.
- [80] L. De Lathauwer. *Signal processing based on multilinear algebra*. PhD thesis, Université catholique de Leuven, Belgique, September 1997.
- [81] L. De Lathauwer, Callaerts D., B. De Moor, and J. Vandewalle. Fetal electrocardiogram extraction by source subspace separation. In *Proc. IEEE SP / Athos workshop*, pages 134–138, Girona, Spain, June 1995.
- [82] L. De Lathauwer, C. Févotte, B. De Moor, and J. Vandewalle. Jacobi algorithm for block diagonalization in blind identification. In *Proc. 23rd Symposium on Information Theory in the Benelux*, pages 155–162, Louvain-la-Neuve, Belgique, May 2002.
- [83] S. Dégerine and E. Kane. A comparative study of approximate joint diagonalization algorithms for blind source separation in presence of additive noise. *IEEE Trans. Signal Proc.*, 55(6) :3022–3031, June 2007.
- [84] N. Delfosse and P. Loubaton. Adaptive separation of independent sources : a deflation approach. In *Acoustics, Speech, and Signal Processing, 1994. ICASSP-94., 1994 IEEE International Conference on*, volume iv, pages IV/41–IV/44 vol.4, Apr 1994.
- [85] L. Di Persia, D. Milone, and M. Yanagida. Indeterminacy free frequency-domain blind separation of reverberant audio sources. *Audio, Speech, and Language Processing, IEEE Transactions on*, 17(2) :299–311, 2009.
- [86] D. Donoho. *On minimum entropy deconvolution*. Academic Press, New York, London, 1981.
- [87] T. Eriksen, G. Høye, B. Narheim, and B.J. Meland. Maritime traffic monitoring using a space-based ais receiver. *Acta Astronautica*, 58(10) :537–549, May 2006.
- [88] D. J. Evans. The use of pre-conditioning in iterative methods for solving linear systems with symmetric positive definite matrices. *J. Inst. Math. Its Appl.*, 4 :295–314, 1968.
- [89] P. Fabry. *Techniques de séparation de sources appliquées à des signaux vibratoires de machines tournantes*. PhD thesis, Grenoble, INPG, Grenoble, 2001.
- [90] E.-M. Fadaili, N. Thirion-Moreau, and E. Moreau. Algorithme de zéro-diagonalisation conjointe pour la séparation de sources déterministes. In *20ème Colloque GRETSI*, September 2005.
- [91] E. M. Fadaili, N. Thirion-Moreau, and E. Moreau. Combined non-orthogonal joint zero-diagonalization and joint diagonalization for source separation. In *IEEE Workshop on Statistical Signal Processing (SSP'2005)*, Bordeaux, France, Juillet 2005.
- [92] E. M. Fadaili, N. Thirion-Moreau, and E. Moreau. Non-orthogonal joint diagonalization/zero-diagonalization for source separation based on time-frequency distributions. *IEEE Trans. Signal Proc.*, 55(5) :1673–1687, May 2007.

-
- [93] A. Ferréol and P. Chevalier. Higher order blind source separation using cyclostationarity property of the signals. In *Proc. International Conference on Acoustic Speech and Signal Processing (ICASSP'97)*, Munich, Germany, April 1997.
 - [94] A. Ferreol and P. Chevalier. On the behavior of current second and higher order blind source separation methods for cyclostationary sources. *Signal Processing, IEEE Transactions on*, 48(6) :1712–1725, Jun 2000.
 - [95] A. Ferréol, P. Chevalier, and L. Albera. Fourth order blind identification of underdetermined mixtures of sources (fobium). *IEEE Transactions on Signal Processing*, 53(5) :1640–1653, May 2005.
 - [96] C. Févotte. *Approche temps-fréquence pour la séparation aveugle de sources non-stationnaires*. PhD thesis, Université de Nantes, France, Octobre 2003.
 - [97] C. Févotte and C. Doncarli. A unified presentation of blind separation methods for convolutive mixtures using block-diagonalization. In *Proc. Independent Component Analysis (ICA'03)*, pages 349–354, Nara, Japan, April 2003.
 - [98] C. Févotte and C. Doncarli. Two contributions to blind source separation using time-frequency distributions. *IEEE Signal Processing Letter*, 11(3) :386–389, March 2004.
 - [99] G.D. Forney. Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference. *Information Theory, IEEE Transactions on*, 18(3) :363–378, May 1972.
 - [100] W. A. Gardner. Measurement of spectral correlation. *IEEE Trans. Acoust.*, 34(5) :1111–1123, Oct. 1986.
 - [101] W. A. Gardner. The spectral correlation theory of cyclo-stationary time-series. *IEEE Trans. Signal Proc.*, 11 :13–36, 1986.
 - [102] W.A. Gardner. Exploitation of spectral redundancy in cyclostationary signals. *Signal Processing Magazine, IEEE*, 8(2) :14–36, April 1991.
 - [103] E. Gassiat. *Déconvolution aveugle*. PhD thesis, Université de Paris-Sud. Faculté des Sciences d'Orsay, 1989.
 - [104] H. Gavin. The levenberg-marquardt method for nonlinear least squares curve-fitting problems. *Department of Civil and Environmental Engineering, Duke University*, pages 1–15, 2011.
 - [105] G. GELLE, M. COLAS, and C. SERVIERE. blind source separation : a tool for rotating machine monitoring by vibrations analysis. *Journal of Sound and Vibration*, 248(5) :865–885, 2001.
 - [106] G. Gelle, M. Colas, and C. Servière. Blind source separation : a new pre-processing tool for rotating machines monitoring? *IEEE trans. Instrumentation and Mesurement*, 53(3) :790–795, June 2003.
 - [107] H. Ghennioui. *Séparation de mélanges convolutifs de sources par l'utilisation de représentations temps-fréquence spatiales quadratiques. Application au traitement de la parole et aux télécommunications numériques*. co-tutelle, Université de Toulon & Université Mohammed V-Agdal de Rabat, Juillet 2008.
 - [108] H. Ghennioui, E.-M. Fadaili, N. Thirion-Moreau, A. Adib, and E. Moreau. A non-unitary joint block diagonalization algorithm for blind separation of convolutive mixtures of sources. *IEEE Signal Processing Letter*, 14(11) :860–863, November 2007.

- [109] H. Ghennioui, E.-M. Fadaili, N. Thirion-Moreau, A. Adib, and E. Moreau. Une nouvelle méthode sans blanchiment pour la séparation aveugle de mélanges convolutifs de sources basée sur l'utilisation de distributions temps-fréquence spatiales quadratiques. In *Colloque du GRETSI*, pages 465–468, Troyes, septembre 2007.
- [110] H. Ghennioui, N. Thirion-Moreau, E. Moreau, and A. Aboutajdine. Gradient-based joint block diagonalization algorithms : Application to blind separation of FIR convolutive mixtures. *Signal Processing*, 90(6) :1836–1849, 2010.
- [111] H. Ghennioui, N. Thirion-Moreau, E. Moreau, D. Aboutajdine, and A. Adib. A novel approach based on non-unitary joint block-diagonalization for the blind mimo equalization of cyclo-stationary signals. In *Signal Processing Conference, 2008 16th European*, pages 1–5, Aug 2008.
- [112] H. Ghennioui, N. Thirion-Moreau, E. Moreau, D. Aboutajdine, and A. Adib. Two new gradient based non-unitary joint block-diagonalization algorithms. In *Signal Processing Conference, 2008 16th European*, pages 1–5, Aug 2008.
- [113] H. Ghennioui, N. Thirion-Moreau, E. Moreau, A. Adib, and D. Aboutajdine. Non unitary joint-block diagonalization of complex matrices using a gradient approach. In Mike E. Davies et al., editor, *7th International Conference on Independent Component Analysis and Signal Separation*, pages 201–208, London, UK, Sept. 2007. Springer-Verlag Berlin Heidelberg, LNCS 4666.
- [114] Hicham Ghennioui. *Blind separation of linear convolutive mixtures of correlated sources*. PhD thesis, Université du Sud Toulon Var, Jul 2008.
- [115] L. Giulieri, H. Ghennioui, N. Thirion-Moreau, and E. Moreau. Non orthogonal joint diagonalization of spatial quadratic time-frequency matrices for source separation. *IEEE Signal Processing Letter*, 12(5) :415–418, May 2005.
- [116] L. Giulieri, N. Thirion-Moreau, and P.-Y. Arquès. Blind sources separation based on bilinear time-frequency representations : A performance analysis. In *ICASSP*, pages 1649–1652, Orlando, May 2002.
- [117] L. Giulieri, N. Thirion-Moreau, and P.-Y. Arquès. Blind sources separation based on bilinear time-frequency representations : a performance analysis. In *Proc. IEEE ICASSP*, pages 1649–1652, Orlando, FL, May 2002.
- [118] L. Giulieri, N. Thirion-Moreau, and P.-Y. Arquès. Blind sources separation based on quadratic time-frequency representations : a method without pre-whitening. In *Proc. International Conference on Acoustic Speech and Signal Processing (ICASSP'03)*, pages 289–292, Honk-Kong, China, April 2003.
- [119] G.-H. Golub and C.-F. Van Loan. *Matrix computation*. Univ. Press Baltimore, 1989.
- [120] N. Gong, W. Shao, and H. Xu. The conjugate gradient method with neural network control. In *Intelligent Systems and Knowledge Engineering (ISKE), 2010 International Conference on*, pages 82–84, Nov 2010.
- [121] A. Gorokhov and P. Loubaton. Subspace based techniques for second order blind separation of convolutive mixtures with temporally correlated sources. *IEEE Transactions on Circuit and Systems*, 44(9) :813–820, September 1997.
- [122] W. Hager and H. Zhang. A survey of nonlinear conjugate gradient methods. *Pacific J. Optim.*, 2 :35–58, 2006.

-
- [123] R. A. Harshman. Foundation of the Parafac procedure : models and conditions for an explanatory multimodal factor analysis. *UCLA Working papers in phonetics*, 16 :1–84, 1970.
 - [124] J. Héroult and B. Ans. Circuits neuronaux à synapses modifiables : décodage de messages composites par apprentissage non supervisé. In *C.R. de l'Académie des sciences*, pages 525–528, 1984.
 - [125] J. Héroult and B. Ans. Réseaux de neurones à synapses modifiables : décodage de messages sensoriels composite par apprentissage non supervisé et permanent. *Compte-rendu de l'Académie des Sciences, Paris*, t. 299, série 3(13) :525–528, 1984.
 - [126] J. Héroult, C. Jutten, and B. Ans. Détection de grandeurs primitives dans un message composite par une architecture de calcul neuromimétique en apprentissage non supervisé. In *GRETSI*, pages 1017–1022, 1985.
 - [127] J. Héroult, C. Jutten, and B. Ans. Détection de grandeurs primitives dans un message composite par une architecture de calcul neuromimétique en apprentissage non supervisé. In *Colloque GRETSI*, pages 1017–1022, Nice, France, Mai 1985.
 - [128] F. L. Hitchcock. The expression of a tensor or a polyadic as a sum of products. *J. Math. and Phys.*, 6 :165–189, 1927.
 - [129] A. Hjørunngnes. *Complex-Valued Matrix Derivatives With Applications in Signal Processing and Communications*. Cambridge, University Press, Feb. 2011.
 - [130] F. Hlawatsch and F. Auger, editors. *Temps-fréquence et traitement statistique*. Hermès Sciences Publications, 2005.
 - [131] F. Hlawatsch and G.F. Boudreaux-Bartles. Linear and quadratic time-frequency signal representations. *IEEE Signal Processing Magazine*, pages 21–67, April 1992.
 - [132] A. Holobar, C. Févotte, C. Doncarli, and D. Zazula. Single autotermes separation based on bilinear time-frequency representations. In *Proc. European Signal Processing Conference (EUSIPCO'02)*, pages 565–568, Toulouse-France, September 2002.
 - [133] G. Høye, B. Narheim, T. Eriksen, and B. J. Meland. *Euclid JP 9.16, Space-Based AIS Reception for Ship Identification*. FFI, 2004.
 - [134] H. L. Hurd. *An investigation of periodicaly correlated processes*. PhD thesis, Duke University, 1970.
 - [135] S. Ikeda and N. Murata. A method of blind separation based on temporal structure of signals. In *Ds Actes Intl. Conf. on Neural Information Processing (ICONIP'98)*, pages 737–742, Kitakyushu, Japan, Oct. 1998.
 - [136] Aissa Ikhlef. *Blind source separation in MIMO communication systems*. PhD thesis, Université Rennes 1, Sep 2008.
 - [137] M.Z. Ikram and D.R. Morgan. A beamforming approach to permutation alignment for multichannel frequency-domain blind speech separation. In *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on*, volume 1, pages 881–884, 2002.
 - [138] Y. Inouye and T. Sato. Iterative algorithms based on multistage criteria for multichannel blind deconvolution. *Signal Processing, IEEE Transactions on*, 47(6) :1759–1764, Jun 1999.
 - [139] International Maritime Organization. *International Convention for the Safety Of Life at Sea, Safety of Navigation, SOLAS chapter V*. IMO, 1974.

- [140] M.G. Jafari, S.R. Alty, and J.A. Chambers. New natural gradient algorithm for cyclostationary sources. *Vision, Image and Signal Processing, IEE Proceedings*, 151(1) :62–68, Feb 2004.
- [141] P. Jallon, A. Chevreuil, P. Loubaton, and P. Chevalier. Separation of convolutive mixtures of cyclostationary sources : a contrast function based approach. In *proc. ICA'04*, Grenade, Spain, September 2004.
- [142] C. Jutten. *Calcul neuro-mimétique et traitement du signal : Analyse en composantes indépendantes*. PhD thesis, UJF-INP Grenoble, 1987.
- [143] C. Jutten and J. Héroult. Une solution neuromimétique au problème de séparation de sources. *Traitement du signal*, 5(6) :389–403, 1988.
- [144] C. Jutten and J. Héroult. Blind separation of sources, part i : An adaptive algorithm based on neuromimetic architecture. *Signal Processing*, 24(1) :1–10, 1991.
- [145] W. Kozek. *Matched Weyl-Heisenberg Expansions of Nonstationary Environments*. PhD thesis, Vienna University of Technology, 1997.
- [146] W. Kozek. On the transfer function calculus for underspread ltv channels. *IEEE Trans. Signal Proc.*, 45(1), January 1997.
- [147] S. Kurita, H. Saruwatari, S. Kajita, K. Takeda, and F. Itakura. Evaluation of blind signal separation method using directivity pattern under reverberant conditions. In *Acoustics, Speech, and Signal Processing, 2000. ICASSP '00. Proceedings. 2000 IEEE International Conference on*, volume 5, pages 3140–3143, 2000.
- [148] D. Lahat, J.-F. Cardoso, and H. Messer. Joint block diagonalization algorithms for optimal separation of multidimensional components. In F. Theis, A. Cichocki, A. Yeredor, and M. Eds. Zibulevsky, editors, *J. In Latent Variable Analysis and Signal Separation, ser. LNCS*, volume 7191 of *Heidelberg : Springer*, pages 155–162, 2012.
- [149] Giulieri Laurent. *Séparation aveugle de sources basée sur l'utilisation des transformées spatiales quadratiques*. PhD thesis, Université de Toulon et du Var, Décembre 2003.
- [150] D. Lee and H. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755) :788–791, 1999.
- [151] D. D. Lee and H. S. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755) :788–791, 1999.
- [152] D. D. Lee and H. S. Seung. Algorithms for non-negative matrix factorization. In *In NIPS*, pages 556–562. MIT Press, 2000.
- [153] H. Lev-Ari. Efficient solution of linear matrix equations with application to multistatic antenna array processing. *Communications in Information and Systems*, 5(1) :123–130, 2005.
- [154] X.-L. Li and X. D. Zhang. Nonorthogonal joint diagonalization free of degenerate solution. *IEEE Trans. Signal Proc.*, 55(5) :1803–1814, May 2007.
- [155] Y.-C. Liang, A.R. Leyman, and B.-H. Soong. Blind source separation using second-order cyclic-statistics. In *Signal Processing Advances in Wireless Communications, First IEEE Signal Processing Workshop on*, pages 57–60, April 1997.
- [156] Y.-C. Liang, A.R. Leyman, and B.-H. Soong. A network structure approach to blind source separation using second order cyclic statistics. In *Circuits and Systems, 1997. ISCAS '97., Proceedings of 1997 IEEE International Symposium on*, volume 4, pages 2549–2551, Jun 1997.

-
- [157] N. Linh-Trung, A. Belouchrani, K. Abed-Meraim, and B. Boashash. *Time-frequency signal analysis and processing : a comprehensive reference*, chapter Undetermined blind sources separation for FM-like signals, pages 3357–3366. Prentice-Hall, Editor : B. Boashash, Oxford, UK, January 2003.
 - [158] R. Liu and Y. Inouye. Blind equalization of mimo-fir channels driven by white but higher order colored source signals. *Information Theory, IEEE Transactions on*, 48(5) :1206–1214, May 2002.
 - [159] P. Loubaton and P. Regalia. Blind deconvolution of multivariate signals : a deflation approach. In *Proc. ICC 93*, pages 1160–1164, Geneva, Switzerland, June 1993.
 - [160] D. G. Luenberger. *Optimization by vector space methods*. Wiley, 1969.
 - [161] Recommendation ITU-R M.1371-5. *Technical characteristics for an automatic identification system using time division multiple access in the VHF maritime mobile frequency bandy*. International Telecommunication Union, Radiocommunication Sector, 2014.
 - [162] International Telecommunications Union Recommendation ITU-R M.2169. *Improved Satellite Detection of AIS*. December 2009.
 - [163] J. R. Magnus and H. Neudecker. *Matrix differential calculus with applications in statistics and econometrics*. John Wiley and sons, Inc, Third Edition, 2007.
 - [164] A. Mansour. *Contribution à la séparation aveugle de sources*. PhD thesis, INP de Grenoble, France, Janvier 1997.
 - [165] A. Mansour, C. Jutten, and P. Loubaton. Subspace method for blind separation of sources and for a convolutive mixture model. In *Signal Processing Conference (EUSIPCO)*, pages 2081–2084, Trieste, Italy, Sep. 1996.
 - [166] A. Mansour, C. Jutten, and P. Loubaton. An adaptive subspace algorithm for blind separation of independent sources in convolutive mixture. *IEEE Transactions on Signal Processing*, 48 :583–586, March 2000.
 - [167] D. W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal on Applied Math.*, 11 :431–441, 1963.
 - [168] K. Matsuoka. Elimination of filtering indeterminacy in blind source separation. *Neurocomputing*, 71(10-12) :2113–2126, 2008.
 - [169] V. Maurandi, C. De Luigi, and E. Moreau. Fast jacobi like algorithms for joint diagonalization of complex symmetric matrices. In *Signal Processing Conference (EUSIPCO), 2013 Proceedings of the 21st European*, pages 1–5, Sept 2013.
 - [170] N. Mitianoudis and M. Davies. Audio source separation of convolutive mixtures. *IEEE Transactions on Speech and Audio Processing*, 11(5) :489–497, September 2003.
 - [171] E. Moreau. A generalization of joint diagonalization criteria for source separation. *IEEE Transactions on Signal Processing*, 49(3) :530–541, March 2001.
 - [172] E. Moreau and T. Adali. *Blind identification and separation of complex valued signals*. Focus Digital Signal and Image Processing Series. ISTE Ltd and Wiley & Sons, Inc, Sept. 2013.
 - [173] E. Moreau, J. Pesquet, and N. Thirion-Moreau. Convolutive blind signal separation based on asymmetrical contrast functions. *Signal Processing, IEEE Transactions on*, 55(1) :356–371, Jan 2007.
 - [174] E. Moreau and J.-C. Pesquet. Generalized contrasts for multichannel blind deconvolution of linear systems. *IEEE Signal Processing Letter*, 4(6) :182–183, June 1997.

- [175] E. Moreau, J.-C. Pesquet, and N. Thirion-Moreau. An equivalence between non symmetrical contrasts and cumulant matching for blind signal separation. In *Proc. First Int. Workshop on Independent Component Analysis and Signal Separation (ICA'99)*, pages 301–306, January 1999.
- [176] E. Moreau and N. Thirion-Moreau. Nonsymmetrical contrasts for sources separation. *Signal Processing, IEEE Transactions on*, 47(8) :2241–2252, Aug 1999.
- [177] D. Morteza and E.A. Reza. A novel initialization approach for solving permutation ambiguity of frequency domain blind source separation. In *Computers Informatics (ISCI), 2012 IEEE Symposium on*, pages 189–192, March 2012.
- [178] Fadaili El Mostafa. *Décompositions matricielles conjointes et séparation aveugle de sources*. PhD thesis, Université de Toulon et du Var, Juillet 2006.
- [179] E. Moulines, P. Duhamel, J. Cardoso, and S. Mayrargue. Subspace methods for the blind identification of multichannel fir filters. *Signal Processing, IEEE Transactions on*, 43(2) :516–525, Feb 1995.
- [180] N. Murata, S. Ikeda, and A. Ziehe. An approach to blind source separation based on temporal structure of speech signals. Technical report 98-2, Riken Brain Science Institute, April 1998.
- [181] N. Murata, S. Ikeda, and A. Ziehe. An approach to blind source separation based on temporal structure of speech signals. Technical report 98-2, Riken Brain Science Institute, April 1998.
- [182] K. Murota and K. Hirade. Gmsk modulation for digital mobile radio telephony. *IEEE Transactions on Communications*, COM-29(7) :1044–1050, July 1981.
- [183] H. B. Nielson. Damping parameter in marquardt's method. Technical report imm-rep-1999-05, dept. of mathematical modeling, Technical University Denmark., 1998.
- [184] D. Nion. A tensor framework for non-unitary joint block diagonalization. *IEEE Trans. Signal Proc.*, 59(10) :4585–4594, Oct. 2011.
- [185] T. Nishikawa, H. Saruwatari, K. Shikano, and S. Makino. Multistage ica for blind source separation of real acoustic convolutive mixture. In *Proc. Independent Component Analysis (ICA'03)*, pages 523–528, Nara, Japan, April 2003.
- [186] P. Paatero. The multilinear engine—a table-driven, least-squares program for solving multilinear problems, including the n-way parallel factor analysis mode. *J. Comput. Graph. Stat.*, 8(4) :854–888, Dec. 1999.
- [187] L. Parra and C. Alvino. Geometric source separation : Merging convolutive source separation with geometric beamforming. *IEEE Trans. on Speech Audio Processing*, 10(6) :352–362, Sept. 2002.
- [188] L. Parra and C. Spence. Convolutive blind source separation of non-stationary sources. *IEEE Transactions on Speech and Audio Processing*, 8(3) :320–327, May 2000.
- [189] L. Parra and C. Spence. Convolutive blind source separation of non-stationary sources. *IEEE Transactions on Speech and Audio Processing*, 8(3) :320–327, May 2000.
- [190] D. T. Pham. Joint approximate diagonalization of positive definite matrices. *SIAM J. on Matrix Anal. and Appl.*, 22(4) :1136–1152, 2001.
- [191] D.-T. Pham. Blind separation of cyclostationary sources using joint block approximate diagonalization. In *Proc. Independent Component Analysis (ICA'07)*, pages 244–251, 2007.

-
- [192] D.-T. Pham and J.-F. Cardoso. Blind separation of instantaneous mixtures of non-stationary sources. *IEEE Transactions Signal Processing*, 49(9) :1837–1848, September 2001.
 - [193] D. T. Pham, C. Servière, and H. Boumaraf. Blind separation of convolutive audio mixtures using nonstationarity. In *Independent Component Analysis and Blind Signal Separation (ICA'03)*, pages 975–980, Nara, Japan, April 2003.
 - [194] D. T. Pham, C. Servière, and H. Boumaraf. Blind separation of speech mixtures based on nonstationarity. In *Intl. Symp. on Signal Processing and its Applications (ISSPA'03)*, pages 73–76, Paris, France, Jul. 2003.
 - [195] E. Polak. *Optimization algorithms and consistent approximations*. Springer, 1997.
 - [196] R. Prevost, M. Coulon, D. Bonacci, J. LeMaitre, J.-P. Millerioux, and J.-Y. Tournet. Interference mitigation and error correction method for ais signals received by satellite. In *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*, pages 46–50, 2012.
 - [197] K. Rahbar and J. Reilly. Blind source separation algorithm for mimo convolutive mixtures. In *Independent Component Analysis and Signal Separation (ICA'01)*, pages 242–247, San Diego, Californie, Dec. 2001.
 - [198] K. Rahbar and J.-P. Reilly. A frequency domain method for blind source separation of convolutive audio mixtures. *IEEE Transactions on Speech and Audio Processing*, 13(5) :832–844, September 2005.
 - [199] K. Rahbar, J.P. Reilly, and J.H. Manton. Blind identification of mimo fir systems driven by quasistationary sources using second-order statistics : a frequency domain approach. *Signal Processing, IEEE Transactions on*, 52(2) :406–417, Feb 2004.
 - [200] R. B. Randall, J. Antoni, and S. Chobsaard. The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals. *Mechanical Systems and Signal Processing*, 15(5) :945–962, 2001.
 - [201] S. Rhioui, N. Thirion-Moreau, and E. Moreau. Under-determined blind identification of cyclo-stationary signals with unknown cyclic frequencies. In *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, volume 4, May 2006.
 - [202] J. Roussel. *Modélisation cyclo-stationnaire et séparation de sources des signaux électromyographiques*. PhD thesis, Université d'Orléans, Décembre 2014.
 - [203] H. Sawada, S. Araki, and S. Makino. Measuring dependence of bin-wise separated signals for permutation alignment in frequency-domain bss. In *IEEE International Symposium on Circuits and Systems*, pages 3247–3250, New Orleans, USA, 2007.
 - [204] H. Sawada, Ryo Mukai, S. Araki, and S. Makino. A robust and precise method for solving the permutation problem of frequency-domain blind source separation. *Speech and Audio Processing, IEEE Transactions on*, 12(5) :530–538, Sept 2004.
 - [205] C. Servière. Separation of speech signals under reverberant conditions. In *Signal Processing Conference, 2004 12th European*, pages 1693–1696, Sept 2004.
 - [206] C. Servière and D.T. Pham. A novel method for permutation correction in frequency-domain in blind separation of speech mixtures. In *the 5th International Conference on Independent Component Analysis and Blind Signal Separation (ICA'04)*, pages 807–815, Granada, Spain, September 2004.

- [207] C. Servière and D.T. Pham. Permutation correction in the frequency domain in blind separation of speech mixtures. *EURASIP Journal on Advances in Signal Processing*, 2006(1) :1–16, 2006.
- [208] H. Shen and M. Kleinstaubler. Non-unitary matrix joint diagonalization for complex independent vector analysis. *EURASIP Journal on Advances in Signal Processing*, 2012(1), 2012.
- [209] J. R. Shewchuk. Introduction to the conjugate gradient method without the agonizing pain. Technical report, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA, 1994.
- [210] C. Simon, P. Loubaton, and C. Jutten. Separation of a class of convolutive mixtures : a contrast function approach. *Signal Processing*, 81(4) :883–887, 2001.
- [211] P. Smaragdis. Blind separation of convolved mixtures in the frequency domain. *Neurocomputing*, 22 :21–34, 1998.
- [212] A. Souloumiac. Nonorthogonal joint diagonalization by combining givens and hyperbolic rotations. *IEEE Trans. Signal Proc.*, 57(6) :2222–2231, June 2009.
- [213] P. Sudhakar and R. Gribonval. A sparsity-based method to solve the permutation indeterminacy in frequency domain convolutive blind source separation. In *ICA 2009, 8th International Conference on Independent Component Analysis and Signal Separation*, LNCS, Paraty, Brazil, March 2009.
- [214] A. Taleb and Jutten C. Source separation in post-nonlinear mixtures. *IEEE transactions on signal processing*, 47(10) :2807–2820, October 1999.
- [215] N. Thirion, J. Mars, and C. Lacoume. Séparation aveugle de signaux large bande : Un nouveau challenge en prospection sismique. In *15ème Colloque du GRETSI, Groupe d’Etudes du Traitement du Signal et des Images*, pages 1335–1338, Juan-les-Pins, France, septembre 1995.
- [216] N. Thirion-Moreau and M. Amin. *Handbook of Blind Source Separation, Independent Component Analysis and Applications*, Comon, P. and Jutten, C. Eds, chapter 11, Spatial quadratic time-frequency domain methods, pages 421–466. Academic Press, Oxford, Burlington, UK, 2010.
- [217] N. Thirion-Moreau and P.-Y. Arquès. Analyses temps-fréquence linéaires et quadratiques. *Techniques de l’Ingénieur, Traité Télécoms*, TE 5 240 :1–26, November 2002.
- [218] N. Thirion-Moreau and P.-Y. Arquès. Analyses temps-fréquence linéaires et quadratiques. *Techniques de l’ingénieur. Télécoms*, pages TE5240–1, 2002.
- [219] N. Thirion-Moreau and E. Moreau. Generalized criteria for blind multivariate signal equalization. *IEEE Signal Processing Letter*, 9(2) :72–74, February 2002.
- [220] P. Tichavsky and Z. Koldovsky. Algorithms for nonorthogonal approximate joint block-diagonalization. In *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*, pages 2094–2098, Aug 2012.
- [221] P. Tichavsky and A. Yeredor. Fast approximate joint diagonalization incorporating weight matrices. *IEEE Trans. Signal Proc.*, 57(3) :878–891, March 2009.
- [222] L. Tomazeli Duarte, S. Moussaoui, and C. Jutten. Source separation in chemical analysis : Recent achievements and perspectives. *IEEE Signal Processing Magazine*, 31(3) :135–146, May 2014.
- [223] L. Tong, R. Liu, V.C. Soon, and Y. F. Huang. Indeterminacy and identifiability of blind identification. *Circuits and Systems, IEEE Transactions on*, 38(5) :499–509, May 1991.

-
- [224] L. Tong, V. C. Soon, Y. F. Huang, and R. Liu. Amuse : a new blind identification algorithm. In *Proc. IEEE ISCAS*, pages 1784–1787, New Orleans, USA, May 1990.
 - [225] A. Touzni, I. Fijalkow, M. Larimore, and J.R. Treichler. A globally convergent approach for blind mimo adaptive deconvolution. In *Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on*, volume 4, pages 2385–2388 vol.4, May 1998.
 - [226] A. Touzni, I. Fijalkow, M. Larimore, and J.R. Treichler. A globally convergent approach for blind mimo adaptive deconvolution. *Signal Processing, IEEE Transactions on*, 49(6) :1166–1178, Jun 2001.
 - [227] T. Trainini and E. Moreau. A coordinate descent algorithm for complex joint diagonalization under hermitian and transpose congruences. *IEEE Trans. Signal Proc.*, 62(19) :4974–4983, Oct. 2014.
 - [228] M. Tria, M. Benidir, and E. Chaumette. blind separation of secondary radar signals using time-frequency analysis. In *16th European Signal Processing Conference (EUSIPCO 2008)*, pages 25–29, Lausanne, Switzerland, August 2008.
 - [229] J.K. Tugnait. Identification and deconvolution of multichannel linear non-gaussian processes using higher order statistics and inverse filter criteria. *Signal Processing, IEEE Transactions on*, 45(3) :658–672, Mar 1997.
 - [230] J.K. Tugnait. Adaptive blind separation of convolutive mixtures of independent linear signals. In *Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on*, volume 4, pages 2097–2100 vol.4, May 1998.
 - [231] J.K.E. Tunaley. An analysis of ais signal collisions. Technical report, August 2005.
 - [232] A. M. Turing. Rounding-off errors in matrix processes. *Quarterly J. Mech. Appl. Math.*, 1 :287–308, 1948.
 - [233] A.-J. van der Veen and A. Paulraj. An analytical constant modulus algorithm. *Signal Processing, IEEE Transactions on*, 44(5) :1136–1155, May 1996.
 - [234] A.J. Viterbi. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *Information Theory, IEEE Transactions on*, 13(2) :260–269, April 1967.
 - [235] T. Wahl and G. Høye. New possible roles of small satellites in maritime surveillance. *Acta Astronautica*, 56(1–2) :273–277, 2005.
 - [236] J. Walgate, A. J. Short, L. Hardy, and V. Vedral. Local distinguishability of multipartite orthogonal quantum states. *Physical Review Letter*, 85 :4972–4975, 2000.
 - [237] L. Wang, H. Ding, and F. Yin. An improved method of permutation correction in convolutive blind source separation. *Archives of Acoustics*, 35(4) :493–504, 2010.
 - [238] W. Wang, J.-A. Chambers, and S. Sanei. A joint diagonalization method for convolutive blind separation of nonstationary sources in the frequency domain. In *Proc. Independent Component Analysis and Blind Sources Separation (ICA'03)*, pages 939–944, Nara, Japan, April 2003.
 - [239] W. Wang, M.G. Jafari, S. Sanei, and J.A. Chambers. Blind separation of convolutive mixtures of cyclostationary sources using an extended natural gradient method. In *Signal Processing and Its Applications, 2003. Proceedings. Seventh International Symposium on*, volume 2, pages 93–96, July 2003.
 - [240] R. Weber and C. Faye. Real time detector for cyclostationary rfi in radio astronomy. In *EUSIPCO*, 1988.

- [241] J. R. Westlake. *A Handbook of Numerical Matrix Inversion and Solution of Linear Equations*. Wiley, New York, 1968.
- [242] A. Westner and V.-M. Bove. Applying blind source separation and deconvolution to real-world acoustic environments. In *Proc. 106th of Audio Engineering Society*, Munich, Germany, 1999.
- [243] B. Widrow, J. McCool, and M. Ball. The complex lms algorithm. *Proceedings of the IEEE*, 63(4) :719–720, April 1975.
- [244] S. Winter, H. Sawada, and S. Makino. Geometrical understanding of the pca subspace method for overdetermined blind source separation. In *Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). 2003 IEEE International Conference on*, volume 2, pages II–769–72 vol.2, April 2003.
- [245] H.-C. Wu and J.-C. Principe. Simultaneous diagonalization in the frequency domain (sdif) for source separation. In *Proc. Independent Component Analysis and Blind Sources Separation (ICA '99)*, pages 245–250, Aussois, France, 1999.
- [246] J. M. F. Xavier, Victor A. N. Barroso, and J. M. F. Moura. Closed-form blind channel identification and source separation in sdma systems through correlative coding. *IEEE Journal on Selected Areas in Communications*, 16 :1506–1517, 1998.
- [247] Y. Xiang, W. Yu, and J. Zhang. Blind separation of cyclostationary sources based on phase-frequency redundancy. In *Signal Processing, 2004. Proceedings. ICSP '04. 2004 7th International Conference on*, volume 1, pages 403–406, Aug 2004.
- [248] Y. Xiang, W. Yu, H. Zheng, and S. Nahavandi. Blind separation of cyclostationary signals from instantaneous mixtures. In *Intelligent Control and Automation, 2004. WCICA 2004. Fifth World Congress on*, volume 1, pages 309–312, June 2004.
- [249] X.-F. Xu, D.-Z. Feng, W. X. Zheng, and H. Zhang. Convolutional blind source separation based on joint block toeplitzization and block-inner diagonalization. *Signal Processing*, 90(1) :119–133, 2010.
- [250] L. Yang. *Storage ring lattice modeling and its applications*. PhD thesis, the faculty of the University Graduate School, Indiana University, September 2007.
- [251] M. Yang, Y. Zou, and L. Fang. Collision and detection performance with three overlap signal collisions in space-based ais reception. In *Trust, Security and Privacy in Computing and Communications (TrustCom), 2012 IEEE 11th International Conference on*, pages 1641–1648, June 2012.
- [252] A. Yeredor. Non-orthogonal joint-diagonalization in the least squares sense with application in blind sources separation. *IEEE Trans. Signal Proc.*, 50(7) :1545–1553, July 2002.
- [253] A. Yeredor, A. Ziehe, and K.-R. Müller. Approximate joint diagonalization using natural gradient approach. In *Lecture Notes in Computer Science, in Proceedings ICA*, pages 89–96, Granada, Espagne, May 2004.
- [254] A. Ypma, A. Leshem, and R. P. W. Duina. Blind separation of rotating machine sources : bilinear forms and convolutional mixtures. *Signal Processing*, 49 :349–368, 2002.
- [255] W.-J. Zeng, X.-L. Li, and X. Zhang, X.-D. and Jiang. An improved signal-selective direction finding algorithm using second-order cyclic statistics. In *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, pages 2141–2144, Taipei, Taiwan, April 2009.
- [256] X. Zeng, S. Li, G. Li, Y. Zhou, and D. Mo. Fetal ECG extraction by combining single-channel svd and cyclostationarity-based blind source separation. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 6(4), August 2013.

-
- [257] L. Zhang. Conjugate gradient approach to blind separation of temporally correlated signals. In *Communications, Circuits and Systems, 2004. ICCCAS 2004. 2004 International Conference on*, volume 2, pages 1008–1012, June 2004.
 - [258] W.-T. Zhang and S.-T. Lou. Iterative algorithm for joint zero diagonalization with application in blind source separation. *Neural Networks, IEEE Transactions on*, 22(7) :1107–1118, July 2011.
 - [259] W. T. Zhang and S. T. Lou. A low complexity iterative algorithm for joint zero-diagonalization. *IEEE Signal Process. Lett.*, 19(2) :115–118, Jan. 2012.
 - [260] Y. Zhang and M. Amin. Subspace analysis of spatial time-frequency distribution matrices. *IEEE Trans. on Signal Processing*, 47(4) :747–759, April 2001.
 - [261] Y. Zhang and M. G. Amin. Blind separation of nonstationary sources based on spatial time-frequency distributions. *EURASIP J. Appl. Signal Process.*, 2006 :1–13, January 2006.
 - [262] M. Zhou and A.-J. van der Veen. Stable subspace tracking algorithm based on a signed urv decomposition. *Signal Processing, IEEE Transactions on*, 60(6) :3036–3051, June 2012.
 - [263] Mu Zhou, A.-J. van der Veen, and R. Van Leuken. Multi-user leo-satellite receiver for robust space detection of ais messages. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 2529–2532, March 2012.
 - [264] A. Ziehe and K.-R. Müller. Tdesp- an efficient algorithm for blind separation using time structure. In *Proc. International Conference on Artificial Neural Networks*, pages 675–680, 1998.

Annexe 1 : Quelques propriétés algébriques importantes

Pour établir les résultats présentés dans les annexes qui suivent et qui seront ensuite utilisés au niveau des Chapitres 2 et 3 pour développer de nouveaux algorithmes de décompositions matricielles conjointes, nous allons nous appuyer sur un certain nombre de propriétés algébriques que nous rappelons ici.

Nous considérons les matrices suivantes :

- trois matrices carrées, $\mathbf{D}_1$, $\mathbf{D}_2$ et $\mathbf{D}_3$ dans $\mathbb{C}^{M \times M}$,
- cinq matrices rectangulaires dont trois $\mathbf{D}_4$, $\mathbf{D}_5$, $\mathbf{D}_6$ dans $\mathbb{C}^{N \times M}$, l'une $\mathbf{D}_7$ dans $\mathbb{C}^{M \times N}$ et la dernière $\mathbf{D}_8$ dans $\mathbb{C}^{F \times G}$.

Soit également les opérateurs suivants :

- $\text{tr}\{\cdot\}$ désignant l'opérateur trace appliqué à une matrice carrée,
- $d(\cdot)$ (ou parfois d) désignant la différentielle d'une fonction,
- $\text{vec}(\cdot)$ désignant l'opérateur de vectorisation qui permet de concaténer les colonnes d'une matrice $N \times M$ afin d'en faire un vecteur de taille $(N \times M) \times 1$,
- $\text{OffBdiag}_{(\mathbf{n})}\{\cdot\}$ désignant l'opérateur défini au niveau de l'Eq. (2.5) lequel permet d'extraire la partie zéro bloc-diagonale de la matrice bloc passée en argument. Il intervient au niveau de la fonction de coût $\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)$,
- $\text{Bdiag}_{(\mathbf{n})}\{\cdot\}$ désignant l'opérateur défini au niveau de l'Eq. (2.4) lequel permet d'extraire la partie bloc-diagonale de la matrice bloc passée en argument. Il intervient au niveau de la fonction de coût $\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)$,
- $\mathbf{T}_{\text{Boff}}$ désigne la matrice $N^2 \times N^2$ de "transformation" définie dans §2.1.5.2. Elle est utilisée avec la fonction de coût $\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)$,
- $\mathbf{T}_{\text{Bdiag}} = \text{diag}\{\text{vec}\{\text{Bdiag}\{\mathbf{1}_N\}\}\}$ définie dans §2.2.5.1.

Nos développements mathématiques sont fondés sur les propriétés suivantes (ils peuvent être trouvés dans [38][110][119][129][153] ou [163]) :

P₀. Avec la fonction de coût $\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)$, nous utilisons le fait que : $\|\text{OffBdiag}_{(\mathbf{n})}\{\mathbf{D}_1\}\|_F^2 = \text{tr}\left\{\mathbf{D}_1^H \text{OffBdiag}_{(\mathbf{n})}\{\mathbf{D}_1\}\right\}$,
avec la fonction de coût $\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)$, nous utilisons le fait que : $\|\text{Bdiag}_{(\mathbf{n})}\{\mathbf{D}_1\}\|_F^2 = \text{tr}\left\{\mathbf{D}_1^H \text{Bdiag}_{(\mathbf{n})}\{\mathbf{D}_1\}\right\}$.

P₁. $(\mathbf{D}_4 + \mathbf{D}_5)^T = \mathbf{D}_4^T + \mathbf{D}_5^T$.

P₂. $(\mathbf{D}_4 \mathbf{D}_7)^T = \mathbf{D}_7^T \mathbf{D}_4^T$.

P₃. $(\mathbf{D}_4 \mathbf{D}_7)^* = \mathbf{D}_4^* \mathbf{D}_7^*$.

P₄. $\text{tr}\{\mathbf{D}_1 + \mathbf{D}_2\} = \text{tr}\{\mathbf{D}_1\} + \text{tr}\{\mathbf{D}_2\}$.

P₅. $\text{tr}\{\mathbf{D}_4^H \mathbf{D}_5\} = (\text{vec}(\mathbf{D}_4))^H \text{vec}(\mathbf{D}_5)$.

- P₆.** $d(\text{tr}\{\mathbf{D}_1\}) = \text{tr}\{d(\mathbf{D}_1)\}$.
- P₇.** $\text{vec}(\mathbf{D}_1\mathbf{D}_2\mathbf{D}_3) = (\mathbf{D}_3^T \otimes \mathbf{D}_1)\text{vec}(\mathbf{D}_2) \Rightarrow \text{vec}(\mathbf{D}_1\mathbf{D}_2) = (\mathbf{I}_N \otimes \mathbf{D}_1)\text{vec}(\mathbf{D}_2) = (\mathbf{D}_2^T \otimes \mathbf{I}_N)\text{vec}\{\mathbf{D}_1\}$
- P_{7b}.** $\text{vec}(\mathbf{D}_4\mathbf{D}_1\mathbf{D}_7) = (\mathbf{D}_7^T \otimes \mathbf{D}_4)\text{vec}(\mathbf{D}_1) \Rightarrow \text{vec}(\mathbf{D}_4\mathbf{D}_1) = (\mathbf{I}_M \otimes \mathbf{D}_4)\text{vec}(\mathbf{D}_1) = (\mathbf{D}_1^T \otimes \mathbf{I}_N)\text{vec}\{\mathbf{D}_4\}$
- P₈.** $\text{vec}(\mathbf{D}_4^T) = \mathbf{K}_{N,M}\text{vec}(\mathbf{D}_4)$ avec $\mathbf{K}_{N,M}$ la $NM \times NM$ matrice carrée de commutation unique définie par : $\mathbf{K}_{N,M} = \sum_n^N \sum_m^M \mathbf{E}_{nm} \otimes \mathbf{E}_{nm}^T$ où $\mathbf{E}_{nm}$ sont les $N \times M$ matrices avec 1 dans la (n, m) position et zéros ailleurs : $(\mathbf{E}_{nm}) = (\delta_{nm})$ (δ_{nm} représente Kronecker de Delta). Par conséquent, la matrice de permutation $\mathbf{K}_{N,M}$ a un seul "1" à la fois dans chaque ligne et chaque colonne.
- P_{8b}.** $\mathbf{K}_{N,M}\mathbf{K}_{M,N} = \mathbf{I}_{MN}$ et $\mathbf{K}_{N,M} = \mathbf{K}_{M,N}^{-1} = \mathbf{K}_{M,N}^T$.
- P₉.** Pour $\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*) : \text{vec}(\text{OffBdiag}\{\mathbf{D}_1\}) = \mathbf{T}_{\text{Boff}}\text{vec}(\mathbf{D}_1)$,
pour $\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*) : \text{vec}(\text{Bdiag}\{\mathbf{D}_1\}) = \mathbf{T}_{\text{Bdiag}}\text{vec}(\mathbf{D}_1)$.
- P₁₀.** $d(\text{vec}(\mathbf{D}_4)) = \text{vec}(d(\mathbf{D}_4))$.
- P₁₁.** $d(\mathbf{D}_4 + \mathbf{D}_5) = d(\mathbf{D}_4) + d(\mathbf{D}_5)$.
- P₁₂.** $d(\mathbf{D}_4\mathbf{D}_7) = d(\mathbf{D}_4)\mathbf{D}_7 + \mathbf{D}_4d(\mathbf{D}_7)$.
- P₁₃.** $d(\mathbf{D}_4^H) = d(\mathbf{D}_4)^H$.
- P₁₄.** $\text{tr}\{\mathbf{D}_1\} = \text{tr}\{\mathbf{D}_1^T\}$.
- P₁₅.** $(\mathbf{D}_4 \otimes \mathbf{D}_8)^T = \mathbf{D}_4^T \otimes \mathbf{D}_8^T$.
- P₁₆.** $(\mathbf{D}_4 \otimes \mathbf{D}_8)^H = \mathbf{D}_4^H \otimes \mathbf{D}_8^H$.
- P₁₇.** $\text{vec}(\mathbf{D}_4 + \mathbf{D}_5) = \text{vec}(\mathbf{D}_4) + \text{vec}(\mathbf{D}_5)$.
- P₁₈.** $\mathbf{D}_4^H = (\mathbf{D}_4^*)^T$.
- P₁₉.** $(\mathbf{D}_4\mathbf{D}_7)^H = \mathbf{D}_7^H\mathbf{D}_4^H$.
- P₂₀.** $(\mathbf{D}_1 \otimes \mathbf{D}_2)(\mathbf{D}_3 \otimes \mathbf{D}_7) = (\mathbf{D}_1\mathbf{D}_3 \otimes \mathbf{D}_2\mathbf{D}_7)$.
- P₂₁.** $(\mathbf{D}_4 + \mathbf{D}_5)^H = (\mathbf{D}_4^H + \mathbf{D}_5^H)$.
- P₂₂.** Avec la fonction de coût $\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*) : \text{OffBdiag}_{(\mathbf{n})}\{\mathbf{D}_1 + \mathbf{D}_2\} = \text{OffBdiag}_{(\mathbf{n})}\{\mathbf{D}_1\} + \text{OffBdiag}_{(\mathbf{n})}\{\mathbf{D}_2\}$,
avec la fonction de coût $\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*) : \text{Bdiag}_{(\mathbf{n})}\{\mathbf{D}_1 + \mathbf{D}_2\} = \text{Bdiag}_{(\mathbf{n})}\{\mathbf{D}_1\} + \text{Bdiag}_{(\mathbf{n})}\{\mathbf{D}_2\}$.
- P₂₃.** $d\mathbf{D}_4^* = (d\mathbf{D}_4)^*$ ([129] p. 53).
- P₂₄.** $\text{tr}\{\mathbf{D}_1\mathbf{D}_6\mathbf{D}_5\} = \text{tr}\{\mathbf{D}_5\mathbf{D}_1\mathbf{D}_6\} = \text{tr}\{\mathbf{D}_6\mathbf{D}_5\mathbf{D}_1\} \Rightarrow \text{tr}\{\mathbf{D}_1\mathbf{D}_6\} = \text{tr}\{\mathbf{D}_6\mathbf{D}_1\}$.
- P_{24'}.** $\text{tr}\{\mathbf{D}_4\mathbf{D}_5\} = \text{tr}\{\mathbf{D}_5\mathbf{D}_4\}$.
- P₂₅.** $d(f(\mathbf{Z}, \mathbf{Z}^*)) = \text{tr}\{\mathbf{D}_1^T\mathbf{Z} + \mathbf{Z}^H\mathbf{D}_2\} = \text{tr}\{\mathbf{D}_1^T d\mathbf{Z} + \mathbf{D}_2^T d\mathbf{Z}^*\} \Rightarrow \frac{\partial f}{\partial \mathbf{Z}} = \mathbf{D}_1$ et $\frac{\partial f}{\partial \mathbf{Z}^*} = \mathbf{D}_2$.

Annexe 2 : Fonction de coût $\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)$ et bloc diagonalisation conjointe

1 Calcul exact des matrices Hessiennes complexes de la fonction de coût $\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)$

Selon [129], la différentielle au premier ordre de la fonction de coût $\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)$ définie au niveau l'Eq. (2.8) est donnée par :

$$d\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*) = \mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)) d\text{vec}(\mathbf{B}) + \mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)) d\text{vec}(\mathbf{B}^*), \quad (1)$$

si les vecteurs complexes $\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*))$ et $\mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*))$ de dimension $1 \times NM$ sont définis de la manière suivante :

$$\begin{aligned} \mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)) &= \text{vec}^T \left(\frac{\partial \mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}} \right), \\ \mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)) &= \text{vec}^T \left(\frac{\partial \mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}^*} \right). \end{aligned} \quad (2)$$

Nous rappelons également que les deux dérivées partielles qui sont impliquées (*i.e.* $\frac{\partial \mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}}$ et $\frac{\partial \mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}^*}$) ont été préalablement calculées dans [110] en utilisant certaines des propriétés précédentes $\mathbf{P}_0$ - $\mathbf{P}_{20}$. Il a été établi qu'elles valent :

$$\frac{\partial \mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}} = \sum_{i=1}^{N_m} \left[\left(\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{M}_i \mathbf{B}^H \} \right)^T \mathbf{B}^* \mathbf{M}_i^* + \left(\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{M}_i \mathbf{B}^H \} \right)^* \mathbf{B}^* \mathbf{M}_i^T \right] \quad (3)$$

$$\begin{aligned} \frac{\partial \mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}^*} &= \sum_{i=1}^{N_m} \left[\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{M}_i \mathbf{B}^H \} \mathbf{B} \mathbf{M}_i^H + \left(\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{M}_i \mathbf{B}^H \} \right)^H \mathbf{B} \mathbf{M}_i \right] \\ &= \left(\frac{\partial \mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}} \right)^*. \end{aligned} \quad (4)$$

$$\mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)) = (\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)))^*. \quad (5)$$

Comme indiqué dans [129], la différentielle à l'ordre 2 est alors donnée par :

$$d^2 \mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*) = d\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)) d\text{vec}(\mathbf{B}) + d\mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)) d\text{vec}(\mathbf{B}^*). \quad (6)$$

Ainsi, nous devons évaluer la différentielle des deux dérivées $\mathcal{D}_{\mathbf{B}}$ et $\mathcal{D}_{\mathbf{B}^*}$ (*i.e.* $d\mathcal{D}_{\mathbf{B}}$ et $d\mathcal{D}_{\mathbf{B}^*}$).

Nous commençons par $d\mathcal{D}_{\mathbf{B}}$. En utilisant les propriétés $\mathbf{P}_2$, $\mathbf{P}_3$, $\mathbf{P}_{10}$ - $\mathbf{P}_{14}$, $\mathbf{P}_{17}$ - $\mathbf{P}_{19}$ dans l'expression (3), nous obtenons alors :

$$\begin{aligned} d\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)) &= \sum_{i=1}^{N_m} \left[\text{vec}^T \left(\left(\text{OffBdiag}_{(\mathbf{n})} \{ d\mathbf{B}^* \mathbf{M}_i^T \mathbf{B}^T \} \right) \mathbf{B}^* \mathbf{M}_i^* \right. \right. \\ &\quad + \left(\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B}^* \mathbf{M}_i^T d\mathbf{B}^T \} \right) \mathbf{B}^* \mathbf{M}_i^* \\ &\quad + \text{vec}^T \left(\left(\text{OffBdiag}_{(\mathbf{n})} \{ d\mathbf{B}^* \mathbf{M}_i^* \mathbf{B}^T \} \right) \mathbf{B}^* \mathbf{M}_i^T \right. \\ &\quad + \left(\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B}^* \mathbf{M}_i^* d\mathbf{B}^T \} \right) \mathbf{B}^* \mathbf{M}_i^T \\ &\quad + \text{vec}^T \left(\left(\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{M}_i \mathbf{B}^H \} \right)^T d\mathbf{B}^* \mathbf{M}_i^* \right. \\ &\quad \left. \left. + \left(\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{M}_i^H \mathbf{B}^H \} \right)^T d\mathbf{B}^* \mathbf{M}_i^T \right) \right]. \end{aligned}$$

Après, en utilisant les propriétés $\mathbf{P}_1$, $\mathbf{P}_7$ et $\mathbf{P}_{17}$ nous trouvons que :

$$\begin{aligned} d\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)) &= \sum_{i=1}^{N_m} \left[\left((\mathbf{B}^* \mathbf{M}_i^*)^T \otimes \mathbf{I}_N \right) \text{vec} \left(\text{OffBdiag}_{(\mathbf{n})} \{ d\mathbf{B}^* \mathbf{M}_i^T \mathbf{B}^T \} \right) \right. \\ &\quad + \left((\mathbf{B}^* \mathbf{M}_i^*)^T \otimes \mathbf{I}_N \right) \text{vec} \left(\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B}^* \mathbf{M}_i^T d\mathbf{B}^T \} \right) \\ &\quad + \left((\mathbf{B}^* \mathbf{M}_i^T)^T \otimes \mathbf{I}_N \right) \text{vec} \left(\text{OffBdiag}_{(\mathbf{n})} \{ d\mathbf{B}^* \mathbf{M}_i^* \mathbf{B}^T \} \right) \\ &\quad + \left((\mathbf{B}^* \mathbf{M}_i^T)^T \otimes \mathbf{I}_N \right) \text{vec} \left(\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B}^* \mathbf{M}_i^* d\mathbf{B}^T \} \right) \\ &\quad + \left(\mathbf{I}_N \otimes \left(\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{M}_i \mathbf{B}^H \} \right)^T \right) (\mathbf{M}_i^H \otimes \mathbf{I}_N) \text{vec}(d\mathbf{B}^*) \\ &\quad \left. + \left(\mathbf{I}_N \otimes \left(\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{M}_i^H \mathbf{B}^H \} \right)^T \right) (\mathbf{M}_i \otimes \mathbf{I}_N) \text{vec}(d\mathbf{B}^*) \right]^T. \end{aligned}$$

Les propriétés $\mathbf{P}_2$, $\mathbf{P}_3$, $\mathbf{P}_7$, $\mathbf{P}_8$, $\mathbf{P}_9$ et $\mathbf{P}_{18}$ nous permettent d'établir que :

$$\begin{aligned} d\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)) &= \sum_{i=1}^{N_m} \left[(\mathbf{M}_i^H \mathbf{B}^H \otimes \mathbf{I}_N) \mathbf{T}_{\text{Boff}}(\mathbf{B} \mathbf{M}_i \otimes \mathbf{I}_N) \text{vec}(d\mathbf{B}^*) \right. \\ &\quad + (\mathbf{M}_i^H \mathbf{B}^H \otimes \mathbf{I}_N) \mathbf{T}_{\text{Boff}}(\mathbf{I}_N \otimes \mathbf{B}^* \mathbf{M}_i^T) \mathbf{K}_{N,M} \text{vec}(d\mathbf{B}) \\ &\quad + (\mathbf{M}_i \mathbf{B}^H \otimes \mathbf{I}_N) \mathbf{T}_{\text{Boff}}(\mathbf{B} \mathbf{M}_i^H \otimes \mathbf{I}_N) \text{vec}(d\mathbf{B}^*) \\ &\quad + (\mathbf{M}_i \mathbf{B}^H \otimes \mathbf{I}_N) \mathbf{T}_{\text{Boff}}(\mathbf{I}_N \otimes \mathbf{B}^* \mathbf{M}_i^*) \mathbf{K}_{N,M} \text{vec}(d\mathbf{B}) \\ &\quad + \left(\mathbf{I}_N \otimes \left(\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{M}_i \mathbf{B}^H \} \right)^T \right) (\mathbf{M}_i^H \otimes \mathbf{I}_N) \text{vec}(d\mathbf{B}^*) \\ &\quad \left. + \left(\mathbf{I}_N \otimes \left(\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{M}_i^H \mathbf{B}^H \} \right)^T \right) (\mathbf{M}_i \otimes \mathbf{I}_N) \text{vec}(d\mathbf{B}^*) \right]^T. \end{aligned}$$

Les propriétés $\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3, \mathbf{P}_{15}, \mathbf{P}_{16}, \mathbf{P}_{17}, \mathbf{P}_{18}, \mathbf{P}_{19}$ et $\mathbf{P}_{20}$ conduisent ensuite au résultat suivant :

$$\begin{aligned}
 d\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)) &= \sum_{i=1}^{N_m} \text{vec}^T(d\mathbf{B}) \left[(\mathbf{K}_{N,M})^T (\mathbf{I}_N \otimes \mathbf{M}_i \mathbf{B}^H) \mathbf{T}_{\text{Boff}}^T(\mathbf{B}^* \mathbf{M}_i^* \otimes \mathbf{I}_N) \right. \\
 &\quad \left. + (\mathbf{K}_{N,M})^T (\mathbf{I}_N \otimes \mathbf{M}_i^H \mathbf{B}^H) \mathbf{T}_{\text{Boff}}^T(\mathbf{B}^* \mathbf{M}_i^T \otimes \mathbf{I}_N) \right] \\
 &\quad + \text{vec}^T(d\mathbf{B}^*) \left[(\mathbf{M}_i^T \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{\text{Boff}}^T(\mathbf{B}^* \mathbf{M}_i^* \otimes \mathbf{I}_N) \right. \\
 &\quad \left. + (\mathbf{M}_i^* \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{\text{Boff}}^T(\mathbf{B}^* \mathbf{M}_i^T \otimes \mathbf{I}_N) \right. \\
 &\quad \left. + \left(\mathbf{M}_i^* \otimes \left(\text{OffBdiag}_{(n)}\{\mathbf{B} \mathbf{M}_i \mathbf{B}^H\} \right) \right) \right. \\
 &\quad \left. + \left(\mathbf{M}_i^T \otimes \left(\text{OffBdiag}_{(n)}\{\mathbf{B} \mathbf{M}_i^H \mathbf{B}^H\} \right) \right) \right]. \tag{7}
 \end{aligned}$$

Finalement, ce $(1 \times NM)$ vecteur peut être réécrit de la façon suivante :

$$d\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)) = \begin{bmatrix} d\text{vec}^T(\mathbf{B}^*) & d\text{vec}^T(\mathbf{B}) \end{bmatrix} \begin{bmatrix} \mathbf{A}_{00} \\ \mathbf{A}_{10} \end{bmatrix}, \tag{8}$$

ce qui conduit aux expressions de $\mathbf{A}_{00}$ et $\mathbf{A}_{10}$ données au niveau des équations (2.18)-(2.19).

En ce qui concerne maintenant la seconde différentielle $d\mathcal{D}_{\mathbf{B}^*}$, nous tirons avantage de l'Eq. (5). En effet, en utilisant $\mathbf{P}_{23}$, nous avons $d\mathcal{D}_{\mathbf{B}^*} = (d\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)))^*$ ce qui conduit finalement à :

$$d\mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)) = \begin{bmatrix} d\text{vec}^T(\mathbf{B}) & d\text{vec}^T(\mathbf{B}^*) \end{bmatrix} \begin{bmatrix} \mathbf{A}_{00}^* \\ \mathbf{A}_{10}^* \end{bmatrix}. \tag{9}$$

Cependant ce vecteur de dimension $(1 \times NM)$ est censé être écrit sous la forme suivante :

$$d\mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{BDC}}(\mathbf{B}, \mathbf{B}^*)) = \begin{bmatrix} d\text{vec}^T(\mathbf{B}^*) & d\text{vec}^T(\mathbf{B}) \end{bmatrix} \begin{bmatrix} \mathbf{A}_{01} \\ \mathbf{A}_{11} \end{bmatrix}, \tag{10}$$

ce qui implique finalement que : $\mathbf{A}_{11} = \mathbf{A}_{00}^*$ et $\mathbf{A}_{10}^* = \mathbf{A}_{01}$.

En incorporant les Eqs. (8) et (10) dans (6), les résultats indiqués par l'Eq. (2.14), peuvent être trouvés, où les matrices $\mathbf{A}_{ij}$ pour $i, j = 1, \dots, 2$ sont donnés par les expressions (2.18)-(2.19). L'expression exacte des quatre matrices Hessiennes complexes données au niveau des équations (2.15), (2.16) peut être obtenue.

2 Calcul du pas optimal

Pour déterminer à chaque itération le pas optimal, nous devons évaluer $\mathcal{C}_{\text{BDC}}(\mathbf{B}^{(m+1)}) = \mathcal{C}_{\text{BDC}}(\mathbf{B}^{(m)} - \mu \mathbf{D}^{(m)})$. Dans la suite, pour simplifier les différentes expressions, la dépendance vis à vis des itérations m sera omise. En utilisant les propriétés $\mathbf{P}_{19}$ et $\mathbf{P}_{21}$, la fonction de coût

$\mathcal{C}_{\text{BDC}}(\mathbf{B} - \mu\mathbf{D})$ peut être exprimée comme suit :

$$\begin{aligned}
\mathcal{C}_{\text{BDC}}(\mathbf{B} - \mu\mathbf{D}) &= \sum_{i=1}^{N_m} \left\| \text{OffBdiag}_{(\mathbf{n})} \left\{ \mu^2 \mathbf{D} \mathbf{M}_i \mathbf{D}^H - \mu (\mathbf{B} \mathbf{M}_i \mathbf{D}^H + \mathbf{D} \mathbf{M}_i \mathbf{B}^H) + \mathbf{B} \mathbf{M}_i \mathbf{B}^H \right\} \right\|_F^2 \\
&= \sum_{i=1}^{N_m} \left\| \text{OffBdiag}_{(\mathbf{n})} \left\{ \mu^2 \mathbf{C}_2 - \mu \mathbf{C}_1 + \mathbf{C}_0 \right\} \right\|_F^2 \\
&= \sum_{i=1}^{N_m} \text{tr} \left\{ (\mu^2 \mathbf{C}_2 - \mu \mathbf{C}_1 + \mathbf{C}_0)^H \text{OffBdiag}_{(\mathbf{n})} \left\{ \mu^2 \mathbf{C}_2 - \mu \mathbf{C}_1 + \mathbf{C}_0 \right\} \right\}
\end{aligned} \tag{11}$$

avec

$$\begin{aligned}
\mathbf{C}_2 &= \mathbf{D} \mathbf{M}_i \mathbf{D}^H \\
\mathbf{C}_1 &= (\mathbf{B} \mathbf{M}_i \mathbf{D}^H + \mathbf{D} \mathbf{M}_i \mathbf{B}^H), \\
\mathbf{C}_0 &= \mathbf{B} \mathbf{M}_i \mathbf{B}^H.
\end{aligned} \tag{12}$$

En utilisant ensuite les propriétés $\mathbf{P}_0, \mathbf{P}_4, \mathbf{P}_{19}, \mathbf{P}_{21}$ et $\mathbf{P}_{22}$, nous trouvons que :

$$\begin{aligned}
\mathcal{C}_{\text{BDC}}(\mathbf{B} - \mu\mathbf{D}) &= \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_0^H \text{OffBdiag}_{(\mathbf{n})} \left\{ \mathbf{C}_0 \right\} \right\} \\
&\quad - \mu \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_1^H \text{OffBdiag}_{(\mathbf{n})} \left\{ \mathbf{C}_0 \right\} + \mathbf{C}_0^H \text{OffBdiag}_{(\mathbf{n})} \left\{ \mathbf{C}_1 \right\} \right\} \\
&\quad + \mu^2 \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_2^H \text{OffBdiag}_{(\mathbf{n})} \left\{ \mathbf{C}_0 \right\} + \mathbf{C}_1^H \text{OffBdiag}_{(\mathbf{n})} \left\{ \mathbf{C}_1 \right\} + \mathbf{C}_0^H \text{OffBdiag}_{(\mathbf{n})} \left\{ \mathbf{C}_2 \right\} \right\} \\
&\quad - \mu^3 \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_2^H \text{OffBdiag}_{(\mathbf{n})} \left\{ \mathbf{C}_1 \right\} + \mathbf{C}_1^H \text{OffBdiag}_{(\mathbf{n})} \left\{ \mathbf{C}_2 \right\} \right\} \\
&\quad + \mu^4 \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_2^H \text{OffBdiag}_{(\mathbf{n})} \left\{ \mathbf{C}_2 \right\} \right\} = a_0 + a_1 \mu + a_2 \mu^2 + a_3 \mu^3 + a_4 \mu^4.
\end{aligned} \tag{13}$$

Cela conduit finalement aux résultats donnés au niveau des équations (2.35)-(2.39).

Annexe 3 : Fonction de coût $\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)$ et zéro-bloc diagonalisation conjointe

Nous signalons que les calculs développés dans cette annexe sont également basés sur les propriétés rappelées au niveau de l'Annexe 1, et que la démarche suivie est la même que dans le cas de la bloc diagonalisation conjointe. Nous commençons donc par déterminer les gradients matriciels complexes de la fonction de coût $\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)$, nous passons ensuite aux matrices Hessiennes complexes et terminons par calcul du pas optimal qui passe par le calcul des coefficients d'un polynôme d'ordre 4.

1 Calcul exact des gradients matriciels de la fonction de coût $\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)$

En utilisant la propriété $\mathbf{P}_0$, la fonction de coût $\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)$ peut être écrite sous la forme suivante :

$$\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*) = \sum_{i=1}^{N_m} \text{tr} \left\{ (\mathbf{B}\mathbf{X}_i\mathbf{B}^H)^H \mathbf{B} \text{diag}_{(\mathbf{n})} \{ \mathbf{B}\mathbf{X}_i\mathbf{B}^H \} \right\}.$$

D'après les propriétés $\mathbf{P}_4$, $\mathbf{P}_{12}$, $\mathbf{P}_{11}$ et $\mathbf{P}_6$, la différentielle de la fonction de coût s'écrit :

$$d\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*) = \sum_{i=1}^{N_m} \text{tr} \left\{ d \left((\mathbf{B}\mathbf{X}_i\mathbf{B}^H)^H \mathbf{B} \text{diag}_{(\mathbf{n})} \{ \mathbf{B}\mathbf{X}_i\mathbf{B}^H \} \right) \right\} \stackrel{\text{def}}{=} \mathcal{O}(\mathbf{B}, \mathbf{B}^*) + \mathcal{E}(\mathbf{B}, \mathbf{B}^*), \quad (1)$$

où :

$$\begin{aligned} \mathcal{O}(\mathbf{B}, \mathbf{B}^*) &= \sum_{i=1}^{N_m} \text{tr} \left\{ d \left((\mathbf{B}\mathbf{X}_i\mathbf{B}^H)^H \right) \mathbf{B} \text{diag}_{(\mathbf{n})} \{ \mathbf{B}\mathbf{X}_i\mathbf{B}^H \} \right\}, \\ \mathcal{E}(\mathbf{B}, \mathbf{B}^*) &= \sum_{i=1}^{N_m} \text{tr} \left\{ (\mathbf{B}\mathbf{X}_i\mathbf{B}^H)^H d \left(\mathbf{B} \text{diag}_{(\mathbf{n})} \{ \mathbf{B}\mathbf{X}_i\mathbf{B}^H \} \right) \right\}. \end{aligned}$$

Les propriétés $\mathbf{P}_{14}$, $\mathbf{P}_4$, $\mathbf{P}_{24}$, $\mathbf{P}'_{24}$, $\mathbf{P}_{13}$, $\mathbf{P}_{12}$, $\mathbf{P}_{13}$ et $\mathbf{P}_{10}$ nous permettent d'établir que :

$$\begin{aligned}\mathcal{O}(\mathbf{B}, \mathbf{B}^*) &= \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{B} \mathbf{X}_i^H d\mathbf{B}^H \text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i \mathbf{B}^H \} \right\} + \sum_{i=1}^{N_m} \text{tr} \left\{ d\mathbf{B} \mathbf{X}_i^H \mathbf{B}^H \text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i \mathbf{B}^H \} \right\} \\ &= \sum_{i=1}^{N_m} \text{tr} \left\{ \left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i \mathbf{B}^H \} \mathbf{B} \mathbf{X}_i^H \right)^T d\mathbf{B}^* \right\} \\ &\quad + \sum_{i=1}^{N_m} \text{tr} \left\{ \left(\left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i \mathbf{B}^H \} \right)^T \mathbf{B}^* \mathbf{X}_i^* \right)^T d\mathbf{B} \right\}.\end{aligned}$$

En utilisant les propriétés $\mathbf{P}_{14}$, $\mathbf{P}_4$, $\mathbf{P}_{24}$, $\mathbf{P}'_{24}$, $\mathbf{P}_5$, $\mathbf{P}_9$, $\mathbf{P}_{13}$ et $\mathbf{P}_{12}$, nous trouvons que :

$$\begin{aligned}\mathcal{E}(\mathbf{B}, \mathbf{B}^*) &= \sum_{i=1}^{N_m} (\text{vec}(\mathbf{B} \mathbf{X}_i \mathbf{B}^H))^H \text{vec} \left(d \left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i \mathbf{B}^H \} \right) \right) \\ &= \sum_{i=1}^{N_m} (\mathbf{T}_{\text{Bdiag}} \text{vec}(\mathbf{B} \mathbf{X}_i \mathbf{B}^H))^H \text{vec} (d(\mathbf{B} \mathbf{X}_i \mathbf{B}^H)) \\ &= \sum_{i=1}^{N_m} \left(\text{vec} \left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i \mathbf{B}^H \} \right) \right)^H \text{vec} (d(\mathbf{B} \mathbf{X}_i \mathbf{B}^H)) \\ &= \sum_{i=1}^{N_m} \text{tr} \left\{ \left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i \mathbf{B}^H \} \right)^H d(\mathbf{B} \mathbf{X}_i \mathbf{B}^H) \right\}.\end{aligned}$$

D'où :

$$\mathcal{E}(\mathbf{B}, \mathbf{B}^*) = \sum_{i=1}^{N_m} \text{tr} \left\{ \left(\left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i \mathbf{B}^H \} \right)^* \mathbf{B}^* \mathbf{X}_i^T \right)^T d\mathbf{B} \right\} + \text{tr} \left\{ \left(\left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i \mathbf{B}^H \} \right)^H \mathbf{B} \mathbf{X}_i \right)^T d\mathbf{B}^* \right\}.$$

En remplaçant $\mathcal{O}(\mathbf{B}, \mathbf{B}^*)$ et $\mathcal{E}(\mathbf{B}, \mathbf{B}^*)$ dans l'Eq. (1) par les expressions trouvées, nous obtenons :

$$\begin{aligned}d\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*) &= \sum_{i=1}^{N_m} \text{tr} \left\{ \left(\left(\left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i \mathbf{B}^H \} \right)^T \mathbf{B}^* \mathbf{X}_i^* \right)^T + \left(\left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i \mathbf{B}^H \} \right)^* \mathbf{B}^* \mathbf{X}_i^T \right)^T \right) d\mathbf{B} \right\} \\ &\quad + \text{tr} \left\{ \left(\left(\left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i \mathbf{B}^H \} \right) \mathbf{B} \mathbf{X}_i^H \right)^T + \left(\left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i \mathbf{B}^H \} \right)^H \mathbf{B} \mathbf{X}_i \right)^T \right) d\mathbf{B}^* \right\}.\end{aligned}$$

Nous exploitons maintenant la propriété $\mathbf{P}_{25}$ pour établir que :

$$\begin{aligned}\frac{\partial \mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}} &= \sum_{i=1}^{N_m} \left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i \mathbf{B}^H \} \right)^T \mathbf{B}^* \mathbf{X}_i^* + \sum_{i=1}^{N_m} \left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i \mathbf{B}^H \} \right)^* \mathbf{B}^* \mathbf{X}_i^T. \\ \frac{\partial \mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}^*} &= \sum_{i=1}^{N_m} \left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i \mathbf{B}^H \} \right) \mathbf{B} \mathbf{X}_i^H + \sum_{i=1}^{N_m} \left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i \mathbf{B}^H \} \right)^H \mathbf{B} \mathbf{X}_i.\end{aligned}$$

Nous en tirons le résultat sur le gradient matriciel donné au niveau de l'Eq. (2.55).

2 Calcul exact des matrices Hessiennes complexes de la fonction de coût $\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)$

D'après [129], la différentielle d'ordre 1 de la fonction de coût $\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)$ définie au niveau de l'Eq. (2.47) est donnée par :

$$d\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*) = \mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)) d\text{vec}(\mathbf{B}) + \mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)) d\text{vec}(\mathbf{B}^*). \quad (2)$$

Les $(1 \times NM)$ vecteurs complexes $\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*))$ et $\mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*))$ sont définis comme suit :

$$\begin{aligned} \mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)) &= \text{vec}^T \left(\frac{\partial \mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}} \right), \\ \mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)) &= \text{vec}^T \left(\frac{\partial \mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}^*} \right). \end{aligned} \quad (3)$$

Nous rappelons également que les deux dérivées partielles qui sont impliquées (*i.e.* $\frac{\partial \mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}}$ et $\frac{\partial \mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}^*}$) ont été calculées au niveau de la section précédente en utilisant certaines des propriétés précédemment données $\mathbf{P}_0$ - $\mathbf{P}_{20}$. Nous avons démontré qu'elles valent :

$$\frac{\partial \mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}} = \sum_{i=1}^{N_m} \left[\left(\text{Bdiag}_{(n)}\{\mathbf{B}\mathbf{X}_i\mathbf{B}^H\} \right)^T \mathbf{B}^* \mathbf{X}_i^* + \left(\text{Bdiag}_{(n)}\{\mathbf{B}\mathbf{X}_i\mathbf{B}^H\} \right)^* \mathbf{B}^* \mathbf{X}_i^T \right], \quad (4)$$

$$\begin{aligned} \frac{\partial \mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}^*} &= \sum_{i=1}^{N_m} \left[\text{Bdiag}_{(n)}\{\mathbf{B}\mathbf{X}_i\mathbf{B}^H\} \mathbf{B}\mathbf{X}_i^H + \left(\text{Bdiag}_{(n)}\{\mathbf{B}\mathbf{X}_i\mathbf{B}^H\} \right)^H \mathbf{B}\mathbf{X}_i \right] \\ &= \left(\frac{\partial \mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}} \right)^*. \end{aligned} \quad (5)$$

$$\mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)) = (\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)))^*. \quad (6)$$

Comme indiqué dans [129], la différentielle au second ordre est alors donnée par :

$$d^2\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*) = d\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)) d\text{vec}(\mathbf{B}) + d\mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)) d\text{vec}(\mathbf{B}^*). \quad (7)$$

Nous devons donc calculer la différentielle des deux dérivées $\mathcal{D}_{\mathbf{B}}$ et $\mathcal{D}_{\mathbf{B}^*}$ (*i.e.* $d\mathcal{D}_{\mathbf{B}}$ et $d\mathcal{D}_{\mathbf{B}^*}$).

Dans un premier lieu, nous commençons par $d\mathcal{D}_{\mathbf{B}}$. En utilisant les propriétés $\mathbf{P}_2$, $\mathbf{P}_3$, $\mathbf{P}_{10}$ - $\mathbf{P}_{14}$,

$\mathbf{P}_{17}$ - $\mathbf{P}_{19}$ dans l'Eq. (4), nous obtenons alors :

$$\begin{aligned} d\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)) &= \sum_{i=1}^{N_m} \left[\text{vec}^T \left(\left(\text{Bdiag}_{(\mathbf{n})} \{ d\mathbf{B}^* \mathbf{X}_i^T \mathbf{B}^T \} \right) \mathbf{B}^* \mathbf{X}_i^* \right. \right. \\ &\quad + \left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B}^* \mathbf{X}_i^T d\mathbf{B}^T \} \right) \mathbf{B}^* \mathbf{X}_i^* \\ &\quad + \text{vec}^T \left(\left(\text{Bdiag}_{(\mathbf{n})} \{ d\mathbf{B}^* \mathbf{X}_i^* \mathbf{B}^T \} \right) \mathbf{B}^* \mathbf{X}_i^T \right. \\ &\quad + \left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B}^* \mathbf{X}_i^* d\mathbf{B}^T \} \right) \mathbf{B}^* \mathbf{X}_i^T \\ &\quad + \text{vec}^T \left(\left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i \mathbf{B}^H \} \right)^T d\mathbf{B}^* \mathbf{X}_i^* \right. \\ &\quad \left. \left. + \left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i^H \mathbf{B}^H \} \right)^T d\mathbf{B}^* \mathbf{X}_i^T \right) \right]. \end{aligned}$$

En utilisant les propriétés $\mathbf{P}_1$, $\mathbf{P}_7$, et $\mathbf{P}_{17}$ nous trouvons que :

$$\begin{aligned} d\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)) &= \sum_{i=1}^{N_m} \left[\left((\mathbf{B}^* \mathbf{X}_i^*)^T \otimes \mathbf{I}_N \right) \text{vec} \left(\text{Bdiag}_{(\mathbf{n})} \{ d\mathbf{B}^* \mathbf{X}_i^T \mathbf{B}^T \} \right) \right. \\ &\quad + \left((\mathbf{B}^* \mathbf{X}_i^*)^T \otimes \mathbf{I}_N \right) \text{vec} \left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B}^* \mathbf{X}_i^T d\mathbf{B}^T \} \right) \\ &\quad + \left((\mathbf{B}^* \mathbf{X}_i^T)^T \otimes \mathbf{I}_N \right) \text{vec} \left(\text{Bdiag}_{(\mathbf{n})} \{ d\mathbf{B}^* \mathbf{X}_i^* \mathbf{B}^T \} \right) \\ &\quad + \left((\mathbf{B}^* \mathbf{X}_i^T)^T \otimes \mathbf{I}_N \right) \text{vec} \left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B}^* \mathbf{X}_i^* d\mathbf{B}^T \} \right) \\ &\quad + \left(\mathbf{I}_N \otimes \left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i \mathbf{B}^H \} \right)^T \right) (\mathbf{X}_i^H \otimes \mathbf{I}_N) \text{vec}(d\mathbf{B}^*) \\ &\quad \left. + \left(\mathbf{I}_N \otimes \left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i^H \mathbf{B}^H \} \right)^T \right) (\mathbf{X}_i \otimes \mathbf{I}_N) \text{vec}(d\mathbf{B}^*) \right]^T. \end{aligned}$$

Les propriétés $\mathbf{P}_2$, $\mathbf{P}_3$, $\mathbf{P}_7$, $\mathbf{P}_8$, $\mathbf{P}_9$, and $\mathbf{P}_{18}$ impliquent que :

$$\begin{aligned} d\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)) &= \sum_{i=1}^{N_m} \left[(\mathbf{X}_i^H \mathbf{B}^H \otimes \mathbf{I}_N) \mathbf{T}_{\text{Bdiag}}(\mathbf{B} \mathbf{X}_i \otimes \mathbf{I}_N) \text{vec}(d\mathbf{B}^*) \right. \\ &\quad + (\mathbf{X}_i^H \mathbf{B}^H \otimes \mathbf{I}_N) \mathbf{T}_{\text{Bdiag}}(\mathbf{I}_N \otimes \mathbf{B}^* \mathbf{X}_i^T) \mathbf{K}_{N,M} \text{vec}(d\mathbf{B}) \\ &\quad + (\mathbf{X}_i \mathbf{B}^H \otimes \mathbf{I}_N) \mathbf{T}_{\text{Bdiag}}(\mathbf{B} \mathbf{X}_i^H \otimes \mathbf{I}_N) \text{vec}(d\mathbf{B}^*) \\ &\quad + (\mathbf{X}_i \mathbf{B}^H \otimes \mathbf{I}_N) \mathbf{T}_{\text{Bdiag}}(\mathbf{I}_N \otimes \mathbf{B}^* \mathbf{X}_i^*) \mathbf{K}_{N,M} \text{vec}(d\mathbf{B}) \\ &\quad + \left(\mathbf{I}_N \otimes \left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i \mathbf{B}^H \} \right)^T \right) (\mathbf{X}_i^H \otimes \mathbf{I}_N) \text{vec}(d\mathbf{B}^*) \\ &\quad \left. + \left(\mathbf{I}_N \otimes \left(\text{Bdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{X}_i^H \mathbf{B}^H \} \right)^T \right) (\mathbf{X}_i \otimes \mathbf{I}_N) \text{vec}(d\mathbf{B}^*) \right]^T. \end{aligned}$$

Les propriétés $\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3, \mathbf{P}_{15}, \mathbf{P}_{16}, \mathbf{P}_{17}, \mathbf{P}_{18}, \mathbf{P}_{19}$, et $\mathbf{P}_{20}$ conduisent au résultat suivant :

$$\begin{aligned}
d\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)) &= \sum_{i=1}^{N_m} \text{vec}^T(d\mathbf{B}) \left[(\mathbf{K}_{N,M})^T (\mathbf{I}_N \otimes \mathbf{X}_i \mathbf{B}^H) \mathbf{T}_{\text{Bdiag}}^T (\mathbf{B}^* \mathbf{X}_i^* \otimes \mathbf{I}_N) \right. \\
&\quad + (\mathbf{K}_{N,M})^T (\mathbf{I}_N \otimes \mathbf{X}_i^H \mathbf{B}^H) \mathbf{T}_{\text{Bdiag}}^T (\mathbf{B}^* \mathbf{X}_i^T \otimes \mathbf{I}_N) \Big] \\
&\quad + \text{vec}^T(d\mathbf{B}^*) \left[(\mathbf{X}_i^T \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{\text{Bdiag}}^T (\mathbf{B}^* \mathbf{X}_i^* \otimes \mathbf{I}_N) \right. \\
&\quad + (\mathbf{X}_i^* \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{\text{Bdiag}}^T (\mathbf{B}^* \mathbf{X}_i^T \otimes \mathbf{I}_N) \\
&\quad + \left(\mathbf{X}_i^* \otimes \left(\text{Bdiag}_{(n)}\{\mathbf{B} \mathbf{X}_i \mathbf{B}^H\} \right) \right) \\
&\quad \left. + \left(\mathbf{X}_i^T \otimes \left(\text{Bdiag}_{(n)}\{\mathbf{B} \mathbf{X}_i^H \mathbf{B}^H\} \right) \right) \right]. \tag{8}
\end{aligned}$$

Finalement, ce vecteur de dimension $1 \times NM$ peut être réécrit de la manière suivante :

$$d\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)) = \begin{bmatrix} d\text{vec}^T(\mathbf{B}^*) & d\text{vec}^T(\mathbf{B}) \end{bmatrix} \begin{bmatrix} \mathbf{A}_{00} \\ \mathbf{A}_{10} \end{bmatrix}, \tag{9}$$

ce qui conduit aux expressions de $\mathbf{A}_{00}$ et $\mathbf{A}_{10}$ données par les Eqs. (2.56)-(2.57).

Concernant maintenant la seconde différentielle $d\mathcal{D}_{\mathbf{B}^*}$, nous tirons à nouveau avantage de l'Eq. (6). En effet, en utilisant la propriété $\mathbf{P}_{23}$, nous obtenons finalement que $d\mathcal{D}_{\mathbf{B}^*} = (d\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)))^*$ ce qui mène alors à :

$$d\mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)) = \begin{bmatrix} d\text{vec}^T(\mathbf{B}) & d\text{vec}^T(\mathbf{B}^*) \end{bmatrix} \begin{bmatrix} \mathbf{A}_{00}^* \\ \mathbf{A}_{10}^* \end{bmatrix}. \tag{10}$$

Cependant ce $(1 \times NM)$ vecteur est censé être écrit comme suit :

$$d\mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{ZBDC}}(\mathbf{B}, \mathbf{B}^*)) = \begin{bmatrix} d\text{vec}^T(\mathbf{B}^*) & d\text{vec}^T(\mathbf{B}) \end{bmatrix} \begin{bmatrix} \mathbf{A}_{01} \\ \mathbf{A}_{11} \end{bmatrix}, \tag{11}$$

ce qui implique que : $\mathbf{A}_{11} = \mathbf{A}_{00}^*$ et $\mathbf{A}_{10}^* = \mathbf{A}_{01}$.

En incorporant les Eqs. (9) et (11) au niveau de l'équation (7), les résultats donnés au niveau de l'équation (2.14), peuvent être trouvés, où les matrices $\mathbf{A}_{ij}$ pour $i, j = 1, \dots, 2$ sont données par les équations (2.56)-(2.57). Finalement, l'expression exacte des quatre matrices Hessiennes complexes données par (2.15), (2.16) peut être établie.

3 Calcul du pas optimal

Pour déterminer à chaque itération le pas optimal, nous devons évaluer $\mathcal{C}_{\text{ZBDC}}(\mathbf{B}^{(m+1)}) = \mathcal{C}_{\text{ZBDC}}(\mathbf{B}^{(m)} - \mu \mathbf{D}^{(m)})$. Dans la suite, pour simplifier les différentes expressions, la dépendance vis à vis des itérations m itérations sera omise. En utilisant les propriétés $\mathbf{P}_{21}$ et $\mathbf{P}_{19}$, la fonction

de coût $\mathcal{C}_{\text{ZBDC}}(\mathbf{B} - \mu\mathbf{D})$ peut être exprimée de la manière suivante :

$$\begin{aligned}
\mathcal{C}_{\text{ZBDC}}(\mathbf{B} - \mu\mathbf{D}) &= \sum_{i=1}^{N_m} \left\| \text{Bdiag}_{(\mathbf{n})} \left\{ \mu^2 \mathbf{D}\mathbf{X}_i \mathbf{D}^H - \mu (\mathbf{B}\mathbf{X}_i \mathbf{D}^H + \mathbf{D}\mathbf{X}_i \mathbf{B}^H) + \mathbf{B}\mathbf{X}_i \mathbf{B}^H \right\} \right\|_F^2 \\
&= \sum_{i=1}^{N_m} \left\| \text{Bdiag}_{(\mathbf{n})} \left\{ \mu^2 \mathbf{C}_2 - \mu \mathbf{C}_1 + \mathbf{C}_0 \right\} \right\|_F^2 \\
&= \sum_{i=1}^{N_m} \text{tr} \left\{ (\mu^2 \mathbf{C}_2 - \mu \mathbf{C}_1 + \mathbf{C}_0)^H \text{Bdiag}_{(\mathbf{n})} \left\{ \mu^2 \mathbf{C}_2 - \mu \mathbf{C}_1 + \mathbf{C}_0 \right\} \right\}, \tag{12}
\end{aligned}$$

avec

$$\begin{aligned}
\mathbf{C}_2 &= \mathbf{D}\mathbf{X}_i \mathbf{D}^H, \\
\mathbf{C}_1 &= (\mathbf{B}\mathbf{X}_i \mathbf{D}^H + \mathbf{D}\mathbf{X}_i \mathbf{B}^H), \\
\mathbf{C}_0 &= \mathbf{B}\mathbf{X}_i \mathbf{B}^H.
\end{aligned} \tag{13}$$

En exploitant les propriétés $\mathbf{P}_0$, $\mathbf{P}_4$, $\mathbf{P}_{22}$, $\mathbf{P}_{21}$, et $\mathbf{P}_{19}$, nous trouvons que :

$$\begin{aligned}
\mathcal{C}_{\text{ZBDC}}(\mathbf{B} - \mu\mathbf{D}) &= \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_0^H \text{Bdiag}_{(\mathbf{n})} \left\{ \mathbf{C}_0 \right\} \right\} \\
&\quad - \mu \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_1^H \text{Bdiag}_{(\mathbf{n})} \left\{ \mathbf{C}_0 \right\} + \mathbf{C}_0^H \text{Bdiag}_{(\mathbf{n})} \left\{ \mathbf{C}_1 \right\} \right\} \\
&\quad + \mu^2 \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_2^H \text{Bdiag}_{(\mathbf{n})} \left\{ \mathbf{C}_0 \right\} + \mathbf{C}_1^H \text{Bdiag}_{(\mathbf{n})} \left\{ \mathbf{C}_1 \right\} + \mathbf{C}_0^H \text{Bdiag}_{(\mathbf{n})} \left\{ \mathbf{C}_2 \right\} \right\} \\
&\quad - \mu^3 \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_2^H \text{Bdiag}_{(\mathbf{n})} \left\{ \mathbf{C}_1 \right\} + \mathbf{C}_1^H \text{Bdiag}_{(\mathbf{n})} \left\{ \mathbf{C}_2 \right\} \right\} \\
&\quad + \mu^4 \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_2^H \text{Bdiag}_{(\mathbf{n})} \left\{ \mathbf{C}_2 \right\} \right\} = a_0 + a_1\mu + a_2\mu^2 + a_3\mu^3 + a_4\mu^4. \tag{14}
\end{aligned}$$

Ce qui conduit aux résultats donnés au niveau des Eqs. (2.62)-(2.66).

Annexe 4 : Bref état de l'art sur la détection automatique de points (α, β) et sur les méthodes de SAS en contexte cyclo-stationnaire

1 Bref état de l'art en matière de détection automatique de points (α, β) pour la construction des ensembles de matrices à décomposer conjointement

Nous considérons un vecteur des sources $\mathbf{s}$ aléatoire ou déterministe. En nous appuyant sur les Eqs. (1.43) et (1.4) et en reprenant les mêmes notations qu'au Chapitre 3, la distribution spatiale quadratique des observations $\mathbf{D}_{\mathbf{xx}}(\alpha, \beta; \mathbf{K}_{\alpha, \beta})$ (qui résulte de l'application d'un opérateur quadratique au vecteur des observations) est dorénavant notée $\mathbf{D}_{\mathbf{xx}}(\alpha, \beta)$ afin de simplifier les notations. Elle admet la décomposition suivante [260] :

$$\mathbf{D}_{\mathbf{xx}}(\alpha, \beta) = \mathbf{A}\mathbf{D}_{\mathbf{ss}}(\alpha, \beta)\mathbf{A}^H, \quad (1)$$

où $\mathbf{D}_{\mathbf{ss}}(\alpha, \beta)$, représente la distribution spatiale quadratique des sources. En présence d'un bruit additif $\mathbf{n}(t)$, cette équation devient :

$$\begin{aligned} \mathbf{D}_{\mathbf{xx}}(\alpha, \beta) &= \mathbf{A}\mathbf{D}_{\mathbf{ss}}(\alpha, \beta)\mathbf{A}^H + \mathbf{D}_{\mathbf{nn}}(\alpha, \beta) + \mathbf{A}\mathbf{D}_{\mathbf{sn}}(\alpha, \beta) + \mathbf{D}_{\mathbf{ns}}(\alpha, \beta)\mathbf{A}^H \\ &= \mathbf{A}\mathbf{D}_{\mathbf{ss}}(\alpha, \beta)\mathbf{A}^H + \mathbf{D}_{\mathbf{nn}}(\alpha, \beta), \end{aligned} \quad (2)$$

le bruit étant supposé être de moyenne nul et indépendant des signaux sources. On rappelle également que dans le cas du modèle convolutif donné par l'Eq. (1.9), l'Eq. (1) s'écrit alors [107] :

$$\mathbf{D}_{\mathbf{xx}}(\alpha, \beta) = \mathbf{H}\mathbf{D}_{\mathbf{ss}}(\alpha, \beta)\mathbf{H}^H, \quad (3)$$

où l'ensemble des matrices carrées $\mathbf{D}_{\mathbf{xx}}(\alpha, \beta) \in \mathbb{C}^{M \times M}$ sera noté $\mathcal{E}$ (ce sont les transformées spatiales quadratiques des observations) tandis que les matrices carrées $\mathbf{D}_{\mathbf{ss}}(\alpha_i, \beta_j) \in \mathbb{C}^{N \times N}$ désignent les transformées spatiales quadratiques des sources et $\mathbf{H} \in \mathbb{C}^{M \times N}$ est une matrice rectangulaire de rang plein ($M \geq N$). Nous rappelons qu'il est possible de passer du modèle convolutif au modèle instantané (cf. §1.1.2.3, en considérant $L = 0$, $L' = 1$ et $\mathbf{A} = \mathbf{H}$).

Il a été montré qu'il existe quatre types différents de points (α_i, β_i) dans le plan $\alpha - \beta$:

- Type I : les points du plan $\alpha - \beta$ qui ne correspondent qu'à des auto-termes sur les sources. En ces points, la distribution spatiale quadratique des sources possède une structure particulière (elle est diagonale dans le cas des mélanges instantanés et bloc-diagonale dans le

cas des mélanges convolutifs. On montre même qu'en raison du théorème des interférences elle ne peut posséder qu'une structure encore plus particulière à savoir diagonale avec un seul terme non nul sur la diagonale ou bloc-diagonale avec un seul bloc non nul sur la diagonale), propriété qui sera généralement perdue sous l'effet du mélange linéaire ; toutefois en de tel points (α_i, β_i) après mélange on ne trouve toujours que des auto-termes sur les sources.

- Type II : les points du plan $\alpha - \beta$ qui ne correspondent qu'à des cross-termes sur les sources. Pour ces points, la distribution spatiale quadratique des sources possède une structure particulière (elle est zéro-diagonale dans le cas des mélanges instantanés) propriété qui sera généralement perdue sous l'effet du mélange linéaire ; toutefois en de tels points (α_i, β_i) après mélange on ne trouve toujours que des cross-termes sur les sources.
- Type III : les points du plan $\alpha - \beta$ où il y a à la fois des auto-termes et des cross-termes sur les sources. En ces points la distribution spatiale quadratique des sources ne possède aucune structure particulière pouvant être exploitée en SAS.
- Type IV : les points du plan $\alpha - \beta$ où il n'y a ni cross-terme ni auto-terme sur les sources (tous les autres points (α_i, β_i)).

Seuls les deux premières catégories de points (I et II) présentent un intérêt pour les SAS. Le problème de la détection automatique de "bons points" dans le plan $\alpha - \beta$ consiste donc en la recherche de points (α_i, β_i) pour lesquels les matrices $\mathbf{D}_{\mathbf{ss}}(\alpha_i, \beta_i)$ ou $\mathbf{D}_{\mathbf{ss}}(\alpha_i, \beta_i)$ possèdent une structure bien particulière. Mais la difficulté est liée au fait que pour y parvenir nous ne pouvons nous appuyer que sur la connaissance des matrices $\mathbf{D}_{\mathbf{xx}}(\alpha_i, \beta_i)$ ou $\mathbf{D}_{\mathbf{xx}}(\alpha_i, \beta_i)$. Cette étape est bien évidemment une étape importante et non-triviale, qui impacte les performances des méthodes proposées. Par ailleurs, le principe de la sélection automatique de points peut différer selon le type de modèle du mélange considéré et selon le type de données (blanchies ou non blanchies). Dans le cas des mélanges instantanés, nous trouvons dans la littérature plusieurs travaux consacrés à ce type de problème, permettant soit la construction d'un seul ensemble de matrices (celui devant être diagonalisé conjointement, que sera noté $\mathcal{M}_{\mathbf{DC}}$) soit la construction de deux ensembles de matrices (celui devant être diagonalisé conjointement et celui devant être zéro diagonalisé conjointement noté quant à lui $\mathcal{M}_{\mathbf{ZDC}}$). Dans [92][216][24], on trouve un panorama des différents détecteurs qui ont été proposés. Tous opèrent en considérant en cas particulier de distributions spatiales quadratiques : les distributions (ou les spectres) temps-fréquence spatiales(aux) quadratiques. Le couple α, β est alors un couple temps-fréquence t, ν . Dans les sous-sections suivantes, nous rappelons le principe de ces principaux détecteurs.

1.1 Détecteurs de points (α_i, β_i) particuliers

1.1.1 En contexte de mélanges instantanés blanchis

L'ensemble des travaux qui ont été conduits exploitent des propriétés particulières des transformées spatiales bilinéaires ou quadratiques et d'autres propriétés telles que l'invariance de la norme de Frobenius ($\|\cdot\|_F$) d'une matrice et l'invariance de sa trace sous transformation unitaire. En effet, comme $\text{tr}\{\mathbf{D}_{\mathbf{x_b x_b}}(t, \nu)\} = \text{tr}\{\mathbf{D}_{\mathbf{ss}}(t, \nu)\}$, cet opérateur doit permettre de décider si des auto-termes sur les sources sont présents ou pas (il y en a si la trace est grande alors qu'il n'y en a pas si la trace est faible).

Dans le premier article ayant porté sur ces techniques de séparation [25], il est suggéré de choisir les matrices $\mathbf{D}_{\mathbf{xx}}(t, \nu)$ de plus haute énergie dans le plan temps-fréquence $(t-f)$. Dans [21] ([65][66] dans le cas bruité), les auteurs proposent de construire deux ensembles de matrices correspondant

à la fois à des matrices diagonales et zéro-diagonales en utilisant le détecteur suivant :

$$\frac{\text{tr}\{\mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu)\}}{\|\mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu)\|_F} \begin{cases} > \epsilon_0 \Rightarrow \mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu) \text{ est ajoutée à } \mathcal{M}_{\text{DC}} \\ < \epsilon_0 \Rightarrow \mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu) \text{ est ajoutée à } \mathcal{M}_{\text{ZDC}} \end{cases} \quad (4)$$

où ϵ_0 est un seuil arbitraire positif de faible valeur et $\mathbf{x}_b$ correspond alors au vecteur des observations blanchies. Avec ce détecteur là, il subsiste un problème potentiel en ce qui concerne les matrices sélectionnées pour la diagonalisation conjointe. Si une valeur non négligeable de la trace atteste bien de la présence d'auto-termes, en aucun cas elle ne certifie l'absence de cross-termes sur les sources.

Ensuite, C. Févotte *et al.* proposent dans [132][96][98] un nouveau critère de sélection plus robuste mais pour les matrices de l'ensemble $\mathcal{M}_{\text{DC}}$ uniquement. Il est basé sur le fait qu'une structure de matrice $\mathbf{D}_{\text{ss}}(t, \nu)$ diagonale n'est obtenue qu'en des points où un auto-terme sur une seule source est présent (si des auto-termes sur deux sources étaient présents alors ces deux sources interféreraient pour donner également des cross-termes et la matrice ne serait alors plus diagonale). Une seule structure de matrice $\mathbf{D}_{\text{ss}}(t, \nu)$ diagonale est donc possible : celle avec un seul terme non-nul sur la diagonale. Le rang d'une telle matrice vaut alors un d'où l'utilisation d'une décomposition en valeurs propres (si le rang vaut 1, la plus grande des valeurs propres est quasiment égale à la somme de toutes les valeurs propres).

$$\begin{cases} \text{tr}\{\mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu)\} > \text{moy}\{\text{tr}\{\mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu)\}\} \\ \max_k \{|\delta_k(t, \nu)|\} \\ \frac{k}{\sum_k |\delta_k^m(t, \nu)|} \geq 1 - \epsilon'_0 \Rightarrow \mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu) \text{ est ajoutée à } \mathcal{M}_{\text{DC}}, \end{cases} \quad (5)$$

où $\delta_k(\cdot)$ désigne k -ème valeur propre, $\text{moy}\{\cdot\}$ désigne l'opérateur de moyenne et ϵ'_0 est une constante positive proche de 0.

Dans [116], le choix du noyau permet d'assurer une propriété de symétrie hermitienne quant aux distributions temps-fréquence spatiales quadratiques utilisées (i.e. $\mathbf{D}_{\mathbf{xx}}(t, \nu) = \mathbf{D}_{\mathbf{xx}}(t, \nu)^H$). Dans ce cas-là, les auto-termes sont alors des quantités réelles alors que les cross-termes possèdent une symétrie hermitienne. Cette propriété est satisfaite par de nombreuses distributions parmi lesquels les distribution de Wigner(-Ville) spatiale, de Wigner(-Ville) lissée spatiale, de Choï-Williams spatial, de Pseudo Wigner(-Ville) spatiale et de Pseudo-Wigner(-Ville) lissée spatiale (moyennant l'ajout d'une contrainte de valeur réelle sur l'une des deux fenêtres de lissage)... Un détecteur combinant à la fois la trace et la partie imaginaire des distributions temps-fréquence spatiales quadratiques est alors proposé. Ainsi, pour construire un ensemble de matrices correspondant à des matrices $\mathbf{D}_{\text{ss}}(t, \nu)$ zéro-diagonales, il suffit de choisir les points (t, ν) associés à des matrices $\mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu)$ de trace faible (ce qui permet d'assurer l'absence d'auto-termes sur les sources) et pour lesquels la partie imaginaire est suffisamment grande (ce qui permet d'assurer la présence de cross-termes sur les sources) :

$$\begin{cases} \|\Im\{\mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu)\}\| > \epsilon_1 \\ |\text{tr}\{\mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu)\}| < \epsilon_2, \end{cases} \quad (6)$$

où ϵ_1 et ϵ_2 deux seuils.

Par ailleurs, pour choisir cette fois-ci les points (t, ν) correspondant à des matrices diagonales, les auteurs proposent de détecter les points de trace élevée (ce qui permet d'assurer la présence d'auto-termes sur les sources) et pour lesquels la partie imaginaire est suffisamment faible (ce qui permet d'éviter la présence de cross-termes sur les sources) :

$$\begin{cases} \|\Im\{\mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu)\}\| < \epsilon_3 \\ |\text{tr}\{\mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu)\}| > \epsilon_4, \end{cases} \quad (7)$$

où ϵ_3 et ϵ_4 deux seuils.

Un problème peut quand même subsister dans le cas de points temps-fréquence où auto-termes et cross-termes sur les sources sont simultanément présents, mais pour lesquels la partie imaginaire des cross-termes est nulle. Ces points devraient être écartés. C'est pourquoi les auteurs ont proposé une procédure pour améliorer le détecteur précédent. Deux solutions à base de filtrage ont été suggérées afin de distinguer les zones où les cross-termes sur les sources sont également nulles dans le voisinage du point considéré des zones des endroits où elles ne font que s'annuler. Le détecteur est alors de la forme suivante :

$$\begin{cases} \|\Im\{\mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu)\}\| < \epsilon_3 \\ |\text{tr}\{\mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu)\}| > \epsilon_4 \\ \|\{G * \Im\{\mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu)\}\}\| < \epsilon'_3 \text{ ou } \|\{H * \Im\{\mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t)\}\}\| < \epsilon'_3, \end{cases} \quad (8)$$

où $G(t, \nu)$ (resp. $H(t, \nu)$) représente un filtre 2D passe-bas (resp. une transformée de Hilbert) et $\{\cdot * \cdot\}(t, \nu)$ pour un produit de convolution. Dans [117], un modèle bruité a également été considéré de manière explicite avec des approches de type Bayésienne ou de Neaman-Pearson : les valeurs des seuils impliqués dans les détecteurs ne sont plus supposés être des constantes mais dépendent alors du rapport signal à bruit.

Ultérieurement, Y. Zhang et M. G. Amin ont proposé dans [261] une sélection de points t - f (auto-termes et cross-termes) plus simple et plus efficace que les méthodes développées dans [21][65][66]. Elle nécessite le calcul de la norme des valeurs propres des matrices de $\mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu)$. Les auteurs proposent alors la procédure suivante :

$$\begin{cases} \frac{\text{tr}\{\mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu)\}}{|\text{tr}\{\mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu)\}|} > \epsilon_5 \\ \frac{|\text{tr}\{\mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu)\}|}{\max\{|\text{tr}\{\mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu)\}|\}} > \epsilon'_5 \Rightarrow \text{est ajoutée à } \mathcal{M}_{\text{DC}}, \end{cases} \quad (9)$$

et

$$\begin{cases} \frac{\text{tr}\{\mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu)\}}{|\text{tr}\{\mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu)\}|} < \epsilon_6 \\ \frac{|\text{tr}\{\mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu)\}|}{\max\{|\text{tr}\{\mathbf{D}_{\mathbf{x}_b\mathbf{x}_b}(t, \nu)\}|\}} > \epsilon'_6 \Rightarrow \text{est ajoutée à } \mathcal{M}_{\text{ZDC}}, \end{cases} \quad (10)$$

avec ϵ'_5, ϵ'_6 sont deux seuils positifs et $\epsilon_6 < \epsilon_5$ où ϵ_5 est proche de 1 et $\max\{\cdot\}$ désigne l'opérateur du maximum.

1.2 En contexte instantané non blanchi

L. Giulieri *et al.* ont montré dans [115] que la SAS fondée sur l'utilisation de distributions temps-fréquence spatiales quadratiques pouvait s'opérer sans l'étape préliminaire classique dite de blanchiment des observations. Ils ont aussi montré qu'il était possible de s'affranchir de l'hypothèse classique d'indépendance des sources et d'ouvrir ces méthodes à d'autres classes de signaux : signaux corrélés, signaux déterministes. Pour y parvenir, les auteurs ont été amenés à généraliser leurs critères de sélection automatique proposés dans [116] pour les adapter au cas de données non blanchies (la matrice $\mathbf{A}$ n'est alors plus nécessairement unitaire)⁴⁵. Cette sélection se fait en supposant que :

45. Notons toutefois, que rien n'empêche de sélectionner les "bons" points temps-fréquence à partir des données blanchies, mais de construire des ensembles de matrices qui elles correspondent à des données non blanchies.

- le noyau de la transformation temps-fréquence spatiale bilinéaire permet d'assurer la propriété de symétrie Hermitienne de celle-ci,
- une propriété discriminante en ce qui concerne les sources est nécessaire afin d'en assurer l'identifiabilité c'est à dire la possible séparation : les signatures des sources dans le plan temps-fréquence seront supposées être suffisamment différentes pour permettre d'assurer l'existence de points qui permettront la construction des ensembles à diagonaliser et/ou à zéro-diagonaliser conjointement,
- la matrice de mélange $\mathbf{A}$ et les sources sont réelles.

Ce détecteur consiste alors à construire l'ensemble $\mathcal{M}_{\text{DC}}$ des matrices à diagonaliser simultanément au moyen de la procédure de sélection automatique suivante :

$$\begin{cases} \|\Im\{\mathbf{D}_{\mathbf{xx}}(t, \nu)\}\| < \epsilon_7 \\ \|\Re\{\mathbf{D}_{\mathbf{xx}}(t, \nu)\}\| > \epsilon_8 \\ \|\{G * \Im\{\mathbf{D}_{\mathbf{xx}}(t, \nu)\}\}\| < \epsilon'_7, \end{cases} \quad (11)$$

où ϵ_7 ϵ_8 ϵ'_7 sont trois seuils.

Enfin, les auteurs améliorent ce dernier détecteur afin de traiter le cas de matrices de mélange $\mathbf{A}$ complexes et de signaux complexes. Les auteurs proposent alors une procédure ne s'appliquant toujours qu'à la sélection automatique de matrices à diagonaliser conjointement (construction de l'ensemble $\mathcal{M}_{\text{DC}}$). Ce détecteur noté $\mathbf{C}_{\mathbf{G}_1}$ est celui que nous utiliserons dans notre contexte applicatif de surveillance maritime. Il exploite à nouveau la propriété de rang un des matrices recherchées. Afin de garantir cette propriété, les auteurs proposent dans [115] d'utiliser une décomposition en valeurs singulières sur la matrice $\mathbf{D}_{\mathbf{xx}}(t, \nu)$ alors réécrite $\mathbf{U}(t, \nu)\Delta(t, \nu)\mathbf{V}(t, \nu)^H$ avec $\mathbf{V}(t, \nu)$ et $\mathbf{U}(t, \nu)$ deux matrices unitaires de dimension $m \times m$ et $\Delta(t, \nu) = \text{diag}\{\delta(t, \nu)\}$ est une matrice diagonale dont les éléments diagonaux sont tous positifs. Notons que $\delta(t, \nu) = (\delta_1(t, \nu), \dots, \delta_m(t, \nu))^T$ et que les valeurs singulières sont triées par ordre décroissant c'est à dire que $\delta_1(t, \nu) \geq \delta_2(t, \nu) \geq \dots \geq \delta_m(t, \nu)$. Ensuite, pour vérifier si la matrice considérée est de rang un, il suffit d'appliquer la règle de détection suivante :

$$\begin{cases} \delta_1(t, \nu) > \epsilon_9, \\ \sum_{i=2}^m \delta_i(t, \nu) < \epsilon_{10}, \end{cases} \quad (12)$$

où ϵ_9 (resp. ϵ_{10}) est une constante positive suffisamment grande (resp. petite).

Une deuxième façon peut être utilisée notée $\mathbf{C}_{\mathbf{G}_2}$:

$$\begin{cases} \mathcal{C}(t, \nu) = \frac{\delta_1(t, \nu)}{\sum_{i=1}^m \delta_i(t, \nu)} > 1 - \epsilon_{11}, \\ \sum_{i=1}^m \delta_i(t, \nu) > \epsilon_{12}, \end{cases} \quad (13)$$

où ϵ_{11} et ϵ_{12} sont des constantes positives. Nous rappelons que le détecteur de sélection automatique de points t - f présenté ci-dessus peut être considéré comme une généralisation de celui proposé dans [98] en ce sens qu'il est applicable sans pré-traitement de type blanchiment des observations.

Un autre détecteur ne requérant pas d'étape de blanchiment et exploitant également cette propriété de rang un des matrices recherchées a été proposé par N. Linh-Trung *et al.* dans [157]. Il permet de construire les deux ensembles de matrices : celui à diagonaliser conjointement et/ou celui à zéro diagonaliser conjointement sous l'hypothèse assez forte toutefois que les sources sont

quasi-disjointes dans le plan t - f (leurs signatures possèdent très peu de points en commun). La procédure suivante est donc suggérée :

$$\left| \frac{\delta_1(t, \nu)}{\|\mathbf{D}_{\mathbf{xx}}(t, \nu)\|} - 1 \right| \begin{cases} < \epsilon_{13} \Rightarrow \mathbf{D}_{\mathbf{xx}}(t, \nu) \text{ est ajoutée à } \mathcal{M}_{\text{DC}} \\ > \epsilon_{13} \Rightarrow \mathbf{D}_{\mathbf{xx}}(t, \nu) \text{ est ajoutée à } \mathcal{M}_{\text{ZDC}}, \end{cases} \quad (14)$$

$$\|\mathbf{D}_{\mathbf{xx}}(t, \nu)\| > \epsilon_{14}, \quad \epsilon_{14} \approx 0.05 \times \mathbf{D}_{\mathbf{xx}}(t_h, \nu_h), \quad (15)$$

où (t_h, ν_h) est le couple temps-fréquence de plus haute énergie. L'Eq. (15) permet d'une manière assez simple de débruiter les matrices $\mathbf{D}_{\mathbf{xx}}(t, \nu)$ dans le cas où elles sont perturbées par des bruits Gaussiens, en ne pré-sélectionnant que celles d'un certain niveau d'énergie.

En ce qui concerne la zéro diagonalisation conjointe, E.M. Fadaili *et al.* améliorent la détection ci-dessus en proposant un nouveau détecteur dans [92] que l'on note $\mathbf{C}_{\text{F-ZD}}$ basé à nouveau sur l'hypothèse des sources quasi-disjointes afin de garder des points t - f avec une énergie suffisante en exploitant la propriété de rang un :

$$\begin{cases} \|\mathbf{D}_{\mathbf{xx}}(t, \nu)\| > \epsilon_{15} \\ \left| \frac{\delta_1(t, \nu)}{\|\mathbf{D}_{\mathbf{xx}}(t, \nu)\|} - 1 \right| \begin{cases} > \epsilon_{16} \Rightarrow \mathbf{D}_{\mathbf{xx}}(t, \nu) \text{ est ajoutée à } \mathcal{M}_{\text{ZDC}}, \\ < \epsilon_{16'} \Rightarrow \mathbf{D}_{\mathbf{xx}}(t, \nu) \text{ est ajoutée à } \mathcal{M}_{\text{DC}}. \end{cases} \end{cases} \quad (16)$$

De plus, Y. Zhang et M. G. Amin ont proposé dans [261] une sélection de points t - f pour la diagonalisation et la zéro diagonalisation conjointe dont la procédure de détection est la suivante :

$$\begin{cases} \frac{\text{tr}\{\mathbf{D}_{\mathbf{xx}}(t, \nu)\}}{|\text{tr}\{\mathbf{D}_{\mathbf{xx}}(t, \nu)\}|} > \epsilon_{17} \\ \frac{|\text{tr}\{\mathbf{D}_{\mathbf{xx}}(t, \nu)\}|}{\max\{|\text{tr}\{\mathbf{D}_{\mathbf{xx}}(t, \nu)\}|\}} > \epsilon'_{17} \Rightarrow \text{est ajoutée à } \mathcal{M}_{\text{DC}}, \end{cases} \quad (17)$$

et

$$\begin{cases} \frac{\text{tr}\{\mathbf{D}_{\mathbf{xx}}(t, \nu)\}}{|\text{tr}\{\mathbf{D}_{\mathbf{xx}}(t, \nu)\}|} < \epsilon_{18} \\ \frac{|\text{tr}\{\mathbf{D}_{\mathbf{xx}}(t, \nu)\}|}{\max\{|\text{tr}\{\mathbf{D}_{\mathbf{xx}}(t, \nu)\}|\}} > \epsilon_{18'} \Rightarrow \text{est ajoutée à } \mathcal{M}_{\text{ZDC}}, \end{cases} \quad (18)$$

avec $\epsilon_{17'}$, $\epsilon_{18'}$ sont deux seuils positifs et $\epsilon_{18} < \epsilon_{17}$ où ϵ_{17} est choisi proche de 1.

Un dernier critère de détection a été proposé par H. Ghennoui dans [114] s'appuyant toujours sur les propriétés du rang d'une matrice pour construire l'ensemble $\mathcal{M}_{\text{DC}}$. Grâce à la décomposition de valeurs singulières, les matrices à diagonaliser conjointement sont choisies de telle façon vérifier la règle ci-dessous :

$$\frac{\delta_1^2(t, \nu)}{\|\mathbf{D}_{\mathbf{xx}}(t, \nu)\|_F^2} > 1 - \eta_1. \quad (19)$$

où $\delta_1(t, \nu)$ est la plus grande valeur singulière de $\mathbf{D}_{\mathbf{xx}}(t, \nu)$ et η_1 est une constante de faible valeur.

1.3 En contexte convolutif blanchi

Dans le cas de mélanges convolutifs, assez peu de solutions ont été proposées. Par extension au cas instantané, on montre qu'il est nécessaire en pratique d'identifier des points t - f pour lesquels la matrice $\mathbf{D}_{\text{SS}}(t, \nu)$ est cette fois bloc-diagonale. C. Févotte *et al.* ont essayé dans [97][96] de généraliser la recherche d'auto-termes simples dans le cas instantané à la recherche de matrices

$\mathbf{D}_{\text{SS}}(t, \nu)$ bloc-diagonales avec un seul bloc non-nul. Ce premier détecteur sélectionne cette fois-ci les matrices $\mathbf{D}_{\mathbf{x}_b \mathbf{x}_b}(t, \nu)$ de rang $L + L'$ de la manière suivante :

$$\frac{\sum_{i=1}^{L+L'} \delta_i(t, \nu)}{\sum_{i=1}^M |\delta_i(t, \nu)|} \simeq 1 \Rightarrow \mathbf{D}_{\mathbf{x}_b \mathbf{x}_b}(t, \nu) \text{ est ajoutée à } \mathcal{M}_{\text{BDC}}, \quad (20)$$

où $\delta_1(t, \nu), \dots, \delta_{L+L'}$ désignent les $L + L'$ plus grandes valeurs propres en valeur absolue de $\mathbf{D}_{\mathbf{x}_b \mathbf{x}_b}(t, \nu)$.

1.4 En contexte convolutif non blanchi

Afin de construire l'ensemble $\mathcal{M}_{\text{BDC}}$ de matrice devant être bloc-diagonalisé conjointement, H. Ghennioui *et al.* proposent dans [109][114] une alternative aux méthodes présentées dans [157][97] lesquelles opèrent sous transformation unitaire. En effet, les auteurs proposent un détecteur généralisant le détecteur introduit en contexte instantané dans [115] exploitant une propriété sur le rang de la matrice. Il s'agit de sélectionner les matrices $\mathbf{D}_{\mathbf{X}\mathbf{X}}(t, \nu)$ dont le rang est égal à $Q = L + L'$. Pour garantir cette propriété, les auteurs proposent une décomposition en valeurs singulières conduisant à $\mathbf{D}_{\mathbf{X}\mathbf{X}}(t, \nu_j) = \mathbf{U}(t, \nu) \mathbf{\Delta}(t, \nu) \mathbf{V}(t, \nu)^H$ avec $\mathbf{V}(t, \nu)$ et $\mathbf{U}(t, \nu)$ matrices unitaires de taille $(M \times M)$ et $\delta(t, \nu) = \text{Diag}\{\mathbf{\Delta}(t, \nu)\}$ un vecteur composé des termes diagonaux de la matrice diagonale $\mathbf{\Delta}(t, \nu)$. Notons que $\delta(t, \nu) = [\delta_1(t, \nu), \dots, \delta_M(t, \nu)]^T$ et en supposant que les valeurs singulières sont rangées dans l'ordre décroissant telle que $\delta_1(t, \nu) \geq \delta_2(t, \nu) \geq \dots \geq \delta_M(t, \nu) \geq 0$. Afin de vérifier si une matrice est de rang Q , les auteurs proposent alors de calculer le rapport suivant :

$$\frac{\sum_{k=1}^{L+L'} \delta_k^2(t, \nu)}{\|\mathbf{D}_{\mathbf{X}}(t, \nu)\|_F^2} \simeq 1 \Rightarrow \mathbf{D}_{\mathbf{X}\mathbf{X}}(t, \nu) \text{ est ajoutée à } \mathcal{M}_{\text{BDC}}. \quad (21)$$

2 Résumé de l'état de l'art sur les méthodes de SAS en contexte cyclo-stationnaire

La théorie de la cyclo-stationnarité voit le jour dans les années 70 à travers les travaux de thèse de Hurd [134]. Ses travaux sont repris ensuite dans les années 80 par Gardner [100][101]. En contexte cyclo-stationnaire, le problème de la séparation aveugle de sources est devenu un axe important de recherche au cours de ces dernières années. Nous rappelons ci-dessous les principaux travaux ayant porté sur ce sujet :

- En 1997, Y.-C. Liang *et al.* sont les premiers dans [155][156] à chercher à exploiter l'information fournie par les fréquences cycliques afin de séparer les sources. Dans [155], les auteurs n'imposent aucune restriction sur les fréquences cycliques.
- K. Abed-Meraïm *et al.* introduisent dans [6] deux critères n'exploitant que les SOD pour séparer les sources avec ou sans fréquences cycliques distinctes.
- En contexte convolutif, Abed-Meraïm *et al.* dans [4] réussissent à séparer les signaux en sélectionnant de manière successive leurs fréquences cycliques et, grâce à la diversité spatiale, ils parviennent à ré-estimer le mélange en s'appuyant sur les statistiques cycliques du second ordre (i.e. la corrélation cyclique que nous définirons plus loin).
- Les auteurs de [77] proposent un algorithme adaptatif utilisant simultanément les SOS et des moments cycliques.
- A. Ferréol et P. Chevalier dans [93][94] proposent une alternative à base de SOD et de cumulants d'ordre quatre.

- A. Ypma *et al.* dans [254] développe une méthode baptisée CycloSOBI qui exploite à la fois la cohérence temporelle et la cyclo-stationnarité des signaux sources.
- Y. Xiang *et al.* dans [247][248] utilisent conjointement la redondance de la phase et la fréquence des signaux sources.
- Les auteurs de [239][140] montrent que l'exploitation de la cyclo-stationnarité au niveau d'un algorithme du gradient naturel améliore les performances de la séparation en terme de vitesse de convergence.
- P. Jallon *et al.* [141] introduisent une fonction de contraste à l'ordre quatre dont l'estimation ne requière pas la connaissance des fréquences cycliques.
- R. Boustany propose dans [35][36] deux méthodes pour la séparation de signaux cyclo-stationnaires en utilisant les statistiques cycliques du second ordre, à savoir CYCLOBLISS (CYCLOstationary BLInd Source Separation) et CYMOD (CYclic Multiple Output Deconvolution).
- Les auteurs dans [201] suggèrent une autre méthode pour identifier la matrice de mélange dans le cas de mélanges sous-déterminés ($N > M$) de signaux cyclostationnaires possédant des fréquences cycliques différentes.
- H. Ghennioui *et al.* généralisent dans [111] quelques idées développées dans [201] pour fournir une solution dans le cas de mélanges convolutifs sur-déterminés de sources ($M \geq N$). Les fréquences cycliques ne sont pas nécessairement connues. Ils exploitent la structure algébrique particulière des matrices de corrélation des sorties après l'application d'une transformation linéaire particulière. Ils utilisent un algorithme de type gradient pour résoudre l'étape de BDC_{NU} .
- Pham propose dans [191] une méthode basée sur la bloc diagonalisation conjointe approchée de la densité inter-spectrale cyclique.

Mis à part le domaine des télécommunications numériques, le modèle des signaux cyclo-stationnaires a également été largement exploité dans d'autres contextes applicatifs et plus particulièrement en :

- biomédical (pour la modélisation des signaux électro-myographiques) [202].
- mécanique : signaux de machines tournantes, diagnostic d'engrenage et roulements à bille [200][106][31][16][17].
- radioastronomie [240].

Documents annexes

Nouvel algorithme de zéro-bloc diagonalisation conjointe par approche de Levenberg-Marquardt

Omar CHERRAK^{1,2}, Hicham GHENNIoui¹, Nadège THIRION-MOREAU², El Hossain ABARKAN¹

¹Université Sidi Mohamed Ben Abdellah

Faculté des Sciences et Techniques de Fès, B.P. 2202, Route d'Immouzzar, Fès, Maroc

²Aix-Marseille Université, CNRS, ENSAM, LSIS, UMR 7296, F-13397 Marseille, France

Université de Toulon, CNRS, LSIS, UMR 7296, F-83957 La Garde, France, Tel: +33 4 94 14 24 56/26 71

{omar.cherrak,hicham.ghennioui}@usmba.ac.ma, thirion@univ-tln.fr, habarkan@yahoo.fr

Résumé – Dans cette communication, nous introduisons un nouveau problème de décomposition matricielle conjointe d'un ensemble de matrices complexes donné appelé Zéro-Bloc Diagonalisation Conjointe Non-Unitaire. Ce problème peut se rencontrer dans différents domaines d'applications dont la séparation aveugle de mélanges convolutifs de sources (ou déconvolution multi-dimensionnelle) notamment dans les méthodes du domaine temps-fréquence spatial quadratique. Nous présentons une nouvelle approche basée sur un schéma d'optimisation numérique itératif de type Levenberg-Marquardt, lequel nécessite les matrices Hessiennes complexes ou des approximations de celles-ci. Nous proposons donc, ici, le calcul exact des matrices Hessiennes puis nous en dérivons l'algorithme suggéré. Les principaux avantages de cette méthode découlent des propriétés de l'algorithme Levenberg-Marquardt lequel est stable, robuste et rapide. Des simulations numériques illustrent le comportement et l'efficacité de l'approche proposée dans différents contextes.

Abstract – In this communication, we address a new problem of joint decomposition of a given set of complex matrices, namely the Non-Unitary Joint Zero-Block Diagonalization. This problem might be encountered in various fields of applications, among which is the blind separation of convolutive mixtures of sources (also known as multidimensional deconvolution) especially when spatial quadratic time-frequency domain methods are considered. We present a new approach based on the Levenberg-Marquardt algorithm which is an iterative numerical optimization scheme that requires the calculation of the complex Hessian matrices or approximations of these latter. We propose, here, the exact calculation of complex Hessian matrices, then the suggested algorithm can be derived. The main advantages of the proposed method stem from the properties of Levenberg-Marquardt algorithm which is stable, more robust and exhibits fast convergence. Computer simulations are provided in order to illustrate the behavior and efficiency of the proposed method in different contexts.

1 Introduction

Durant les dernières années, le problème de la décomposition conjointe d'ensembles de matrices (voire de tenseurs) a suscité de nombreux travaux en traitement du signal, essentiellement dans le cadre d'applications en séparation aveugle de sources (SAS) et en traitement d'antenne (voir [6] pour un panorama assez exhaustif). L'un des tout premiers problèmes à avoir été considéré est celui de la diagonalisation conjointe (DC) d'un ensemble de matrices sous la contrainte d'unitarité du diagonaliseur conjoint recherché, conduisant à de nombreux algorithmes en SAS pour des mélanges instantanés, les plus connus étant JADE [3] et SOBI [2]. Les travaux suivants ont traité du problème de la DC de matrices mais en levant la contrainte d'unitarité sur la matrice recherchée [4][19][20] ou de la DC de tenseurs [9][16]. Un second type de décomposition matricielle, dénommé Bloc-Diagonalisation Conjointe (BDC), a été appliqué dans des problèmes de localisation de sources large bande en présence de bruit corrélé et de SAS pour des mélanges convolutifs (ou déconvolution multidimensionnelle). Plusieurs algorithmes ont été suggérés sous différentes

hypothèses qu'elles concernent l'ensemble des matrices à bloc-diagonaliser (définies positives ou hermitiennes) ou le bloc-diagonaliseur (supposé unitaire [10] ou non [7][13][17]). Enfin, un troisième type de décomposition matricielle, dénommé Zéro Diagonalisation Conjointe (ZDC), s'est avéré utile en SAS, télécommunications numériques et cryptographie. Si le premier algorithme suggéré opérait sous contrainte d'unitarité [1], parce qu'il était appliqué après une étape classique de blanchiment spatial, le fait que cette étape borne les performances a très rapidement conduit au développement d'autres solutions sans contrainte [5][11][12][21].

Nous considérons, ici, le problème de la zéro bloc-diagonalisation conjointe non unitaire (ZBDC_{NU}) d'un ensemble de matrices complexes. Il généralise celui de la ZDC non unitaire. Nous proposons un nouvel algorithme basé sur un schéma d'optimisation itératif de Levenberg-Marquardt lequel nécessite les matrices Hessiennes complexes ou des approximations de celles-ci. L'algorithme que nous dérivons ici est fondé sur un calcul exact des matrices Hessiennes. Des simulations numériques illustrent le comportement de l'approche proposée dans diffé-

rentes situations. Les matrices sont d'abord considérées comme exactement zéro bloc-diagonales, puis elles sont ensuite progressivement perturbées par un bruit additif Gaussien. Enfin, ce nouvel algorithme de ZBDC_{NU} est comparé à l'algorithme proposé dans [8] reposant sur le gradient conjugué non linéaire lequel ne nécessite pas le calcul des matrices Hessiennes. Cette comparaison illustre l'intérêt de la méthode suggérée.

2 Énoncé du problème & hypothèses

Le problème de la zéro-bloc diagonalisation conjointe non unitaire (ZBDC_{NU}) s'énonce comme suit. Considérons un ensemble $\mathcal{N}$ de N_m ($N_m \in \mathbb{N}^*$) matrices carrées $\mathbf{X}_i \in \mathbb{C}^{M \times M}$, $i \in \{1, \dots, N_m\}$ admettant toutes une factorisation de la forme $\mathbf{X}_i = \mathbf{A} \mathbf{Z}_i \mathbf{A}^H$ où $(\cdot)^H$ désigne l'opérateur de transposition conjuguaison et où les matrices

$$\mathbf{Z}_i = \begin{pmatrix} \mathbf{0}_{11} & \mathbf{Z}_{i,12} & \dots & \mathbf{Z}_{i,1r} \\ \mathbf{Z}_{i,21} & \mathbf{0}_{22} & \ddots & \mathbf{Z}_{i,2r} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{Z}_{i,r1} & \dots & \vdots & \mathbf{0}_{rr} \end{pmatrix} \forall i \in \{1, \dots, N_m\},$$

sont des matrices zéro-bloc diagonales de taille $N \times N$, où $r \in \mathbb{N}^*$ est le nombre de blocs considérés. Les matrices $\mathbf{Z}_{i,jk}$, $i \in \{1, \dots, N_m\}$, $j, k \in \{1, \dots, r\}$ sont des matrices rectangulaires de taille $n_j \times n_k$ de telle sorte que $n_1 + \dots + n_r = N$ alors que $\mathbf{0}_{jj}$ désigne la matrice nulle de taille $n_j \times n_j$. $\mathbf{A}$ est une matrice de rang plein de taille $M \times N$ ($M \geq N$) tandis que sa pseudo-inverse (ou inverse généralisée de Moore-Penrose) de taille $N \times M$ est notée $\mathbf{B}$. L'ensemble des N_m matrices $\mathbf{Z}_i \in \mathbb{C}^{N \times N}$ zéro-bloc diagonales est noté $\mathcal{Z}$. La taille des blocs n_j pour tout $j = 1, \dots, r$ est supposée connue.

Le problème de la ZBDC_{NU} consiste à estimer soit l'ensemble des matrices zéro-bloc diagonales $\mathcal{Z}$ et la matrice $\mathbf{A}$ soit la matrice $\mathbf{B}$, ce uniquement à partir des matrices de l'ensemble $\mathcal{N}$. Pour résoudre ce problème, nous suggérons, ici, de considérer la fonction de coût suivante :

$$\mathcal{C}_{ZBD}(\mathbf{B}) = \sum_{i=1}^{N_m} \|\text{Bdiag}_{(n)}\{\mathbf{B} \mathbf{X}_i \mathbf{B}^H\}\|_F^2, \quad (1)$$

où $\|\cdot\|_F$ désigne la norme de Frobenius et où l'opérateur matriciel $\text{Bdiag}_{(n)}\{\cdot\}$ est défini de la manière suivante :

$$\text{Bdiag}_{(n)}\{\mathbf{X}\} = \begin{pmatrix} \mathbf{X}_{11} & \mathbf{0}_{12} & \dots & \mathbf{0}_{1r} \\ \mathbf{0}_{21} & \mathbf{X}_{22} & \ddots & \mathbf{0}_{2r} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0}_{r1} & \mathbf{0}_{r2} & \dots & \mathbf{X}_{rr} \end{pmatrix}, \quad (2)$$

avec $\mathbf{X}$ matrice carrée $N \times N$ composée de blocs $\mathbf{X}_{ij}$ de taille $n_i \times n_j$, pour tout i et j allant de 1 à r (avec $n_1 + \dots + n_r = N$). On note enfin $\mathbf{n} = (n_1, n_2, \dots, n_r)$.

3 Nouvel algorithme de ZBDC_{NU}

La fonction de coût donnée au niveau de l'Eq. (1) doit être minimisée afin d'estimer le zéro-bloc diagonaliseur conjoint $\mathbf{B} \in \mathbb{C}^{N \times M}$. Pour ce faire, nous proposons un nouvel algorithme basé sur une approche itérative de type Levenberg-Marquardt (LMA) [15]. Cette approche connue pour sa robustesse

et son efficacité a été largement utilisée dans divers domaines dont les réseaux de neurones, le traitement d'antenne, la séparation aveugle de source, etc...

3.1 Principe & calculs

Partant d'une matrice $\mathbf{B}$ initiale, notée $\mathbf{B}^{(0)}$, à chaque itération m , nous devons ré-estimer cette matrice que nous notons alors $\mathbf{B}^{(m)}$. Cette matrice peut aussi être stockée dans un vecteur $\mathbf{b}^{(m)} = \text{vec}\{\mathbf{B}^{(m)}\}$, $\text{vec}\{\cdot\}$ étant l'opérateur de vectorisation qui permet de ranger les colonnes d'une matrice $N \times M$ dans un vecteur de taille $(N \times M) \times 1$. Dans un schéma de type LMA, ce vecteur est mis à jour selon la règle d'adaptation suivante, pour tout $m = 1, 2, \dots$:

$$\mathbf{b}^{(m)} = \mathbf{b}^{(m-1)} - [\mathbf{H}^{(m-1)} + \lambda \mathbf{I}_{MN}]^{-1} \mathbf{g}^{(m-1)}, \quad (3)$$

où $[\cdot]^{-1}$ désigne l'inverse d'une matrice, λ est un petit coefficient (positif) d'amortissement, $\mathbf{I}_{MN}$ est la matrice identité de taille $MN \times MN$, $\mathbf{H}$ désigne la matrice Hessienne et le vecteur $\mathbf{g}$ résulte de la vectorisation du gradient matriciel complexe noté $\mathbf{G}^{(m-1)} = \nabla_{\mathbf{a}} \mathcal{C}_{ZBD}(\mathbf{B}^{(m-1)})$ i.e. $\mathbf{g}^{(m-1)} = \text{vec}\{\mathbf{G}^{(m-1)}\}$. Le gradient matriciel a été calculé dans [8] où l'on démontre que cette matrice vaut :

$$2 \sum_{i=1}^{N_m} \left[(\text{Bdiag}\{\mathbf{B} \mathbf{X}_i \mathbf{B}^H\}) \mathbf{B} \mathbf{X}_i^H + (\text{Bdiag}\{\mathbf{B} \mathbf{X}_i \mathbf{B}^H\})^H \mathbf{B} \mathbf{X}_i \right], \quad (4)$$

$(\cdot)^T$ étant l'opérateur de transposition et $(\cdot)^*$ celui de conjuguaison. Il nous reste donc à calculer la matrice Hessienne $\mathbf{H} =$

$\begin{pmatrix} \mathbf{H}_{\mathbf{B}, \mathbf{B}^*} & \mathbf{H}_{\mathbf{B}^*, \mathbf{B}^*} \\ \mathbf{H}_{\mathbf{B}, \mathbf{B}} & \mathbf{H}_{\mathbf{B}^*, \mathbf{B}} \end{pmatrix}$ elle-même composée de quatre matrices complexes définies comme dans [14], par :

$$\mathbf{H}_{\mathbf{B}, \mathbf{B}^*} = \frac{\mathbf{A}_{00} + \mathbf{A}_{11}^T}{2} = [\mathbf{H}_{\mathbf{B}^*, \mathbf{B}}]^T, \quad (5)$$

$$\mathbf{H}_{\mathbf{B}^*, \mathbf{B}^*} = \frac{\mathbf{A}_{01} + \mathbf{A}_{01}^T}{2}, \quad \mathbf{H}_{\mathbf{B}, \mathbf{B}} = \frac{\mathbf{A}_{10} + \mathbf{A}_{10}^T}{2}, \quad (6)$$

où les quatre matrices complexes $\mathbf{A}_{00}$, $\mathbf{A}_{01}$, $\mathbf{A}_{10}$ et $\mathbf{A}_{11}$ impliquées sont déterminées au moyen du calcul de la dérivée seconde de la fonction de coût notée $d^2 \mathcal{C}_{ZBD}(\mathbf{B})$ que l'on commence par mettre sous la forme :

$$d^2 \mathcal{C}_{ZBD}(\mathbf{B}) = [(\text{dvec}\{\mathbf{B}^*\})^T, (\text{dvec}\{\mathbf{B}\})^T] \begin{bmatrix} \mathbf{A}_{00} & \mathbf{A}_{01} \\ \mathbf{A}_{10} & \mathbf{A}_{11} \end{bmatrix} \begin{bmatrix} \text{dvec}\{\mathbf{B}\} \\ \text{dvec}\{\mathbf{B}^*\} \end{bmatrix}. \quad (7)$$

En annexe, nous établissons qu'elles valent :

$$\mathbf{A}_{00} = \mathbf{A}_{11}^* = (\mathbf{X}_i^T \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{\text{Bdiag}}^T (\mathbf{B}^* \mathbf{X}_i^* \otimes \mathbf{I}_N) + (\mathbf{X}_i^* \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{\text{Bdiag}}^T (\mathbf{B}^* \mathbf{X}_i^T \otimes \mathbf{I}_N) \quad (8)$$

$$+ \mathbf{X}_i^* \otimes (\text{Bdiag}_{(n)}\{\mathbf{B} \mathbf{X}_i \mathbf{B}^H\}) + \mathbf{X}_i^T \otimes \text{Bdiag}_{(n)}\{\mathbf{B} \mathbf{X}_i^H \mathbf{B}^H\},$$

$$\mathbf{A}_{10} = \mathbf{A}_{01}^* = \mathbf{T}_{\text{vec}}^T (\mathbf{I}_N \otimes \mathbf{X}_i \mathbf{B}^H) \mathbf{T}_{\text{Bdiag}}^T (\mathbf{B}^* \mathbf{X}_i^* \otimes \mathbf{I}_N) + \mathbf{T}_{\text{vec}}^T (\mathbf{I}_N \otimes \mathbf{X}_i^H \mathbf{B}^H) \mathbf{T}_{\text{Bdiag}}^T (\mathbf{B}^* \mathbf{X}_i^T \otimes \mathbf{I}_N), \quad (9)$$

où l'opérateur $\otimes$ désigne le produit de Kronecker, $\mathbf{T}_{\text{vec}}$ est une matrice carrée de permutation de taille $(MN \times MN)$ et $\mathbf{T}_{\text{diag}} = \text{diag}\{\text{vec}\{\text{Bdiag}\{\mathbf{1}_N\}\}\}$, $\mathbf{1}_N$ est la matrice $N \times N$ dont tous les éléments sont égaux à un et $\text{diag}\{\mathbf{a}\}$ est une matrice diagonale carrée qui contient les éléments du vecteur $\mathbf{a}$ sur sa diagonale.

Selon [18], il est suggéré de ne considérer que la diagonale (partie réelle) de la matrice Hessienne (i.e. Eq. (5)) afin d'avoir facilement son inverse et assurer la minimisation de la fonction de coût.

3.2 L'algorithme en résumé

Le principe de l'algorithme proposé, noté JBDC_{LM} , est récapitulé ci-dessous :

Data: N_m matrices carrées $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{N_m}$, critère d'arrêt ϵ (nbre positif petit), nombre max. d'itérations M_{max}

Result: Estimation du zéro bloc-diagonaliseur conjoint $\mathbf{B}$
Initialisations : $\mathbf{B}^{(0)}$; $m = 0$; facteur d'amortissement $\lambda^{(0)}$

repeat

 Calcul du gradient matriciel complexe

$\mathbf{g}^{(m)} = \text{vec}\{\nabla_a \mathcal{C}_{\text{ZBD}}(\mathbf{B}^{(m)})\}$ (Eq. (4))

 Calcul des deux blocs de la diagonale de la Hessienne (Eq. (5))

 Calcul de $\mathbf{b}^{(m+1)} = \text{vec}\{\mathbf{B}^{(m+1)}\}$ (Eq. (3))

 Calcul de l'erreur $e^{(m)} = \frac{1}{N_m} \mathcal{C}_{\text{ZBD}}(\mathbf{B}^{(m+1)})$

$m = m + 1$;

if $e^{(m)} \geq e^{(m-1)}$ **then**

$\lambda^{(m)} = \frac{\lambda^{(m-1)}}{10}$, $e^{(m)} = e^{(m-1)}$

else

$\lambda^{(m)} = 10\lambda^{(m-1)}$

end

until ($\|\mathbf{B}^{(m+1)} - \mathbf{B}^{(m)}\|_F^2 \leq \epsilon$) ou ($m \geq M_{\text{max}}$);

4 Simulations numériques

Le comportement de l'algorithme est maintenant évalué grâce à des simulations numériques. Nous commençons par considérer des ensembles de N_m matrices exactement zéro-bloc diagonales générées aléatoirement (*i.e.* les éléments des matrices que nous construisons suivent tous une loi Gaussienne de moyenne 0 et de variance 1). Ensuite, nous les perturbons afin d'évaluer la robustesse de l'algorithme dans le cas d'erreurs vis à vis de l'hypothèse de zéro bloc-diagonalité. Dans ce second cas, une matrice dont les termes aléatoires sont tirés selon une loi Gaussienne de moyenne 0 et de variance σ_b^2 est ajoutée à la précédente. Le rapport signal à bruit (RSB) est alors défini en dB comme $\text{RSB} = 10 \log(\frac{1}{\sigma_b^2})$. Pour juger de la qualité de l'estimation, nous utilisons l'indice d'erreur suivant $I_{\text{conv}}(\mathbf{G})$ - généralisant celui introduit dans [16] :

$$\frac{1}{r(r-1)} \left[\sum_{i=1}^r \left(\sum_{j=1}^r \frac{\|(\mathbf{G})_{i,j}\|_F^2}{\max_{\ell} \|(\mathbf{G})_{i,\ell}\|_F^2} - 1 \right) + \sum_{j=1}^r \left(\sum_{i=1}^r \frac{\|(\mathbf{G})_{i,j}\|_F^2}{\max_{\ell} \|(\mathbf{G})_{\ell,j}\|_F^2} - 1 \right) \right] \quad (10)$$

où $(\mathbf{G})_{i,j} \forall i, j \in \{1, \dots, r\}$ est la (i, j) -ème matrice zéro-bloc de $\mathbf{G} = \hat{\mathbf{B}}\mathbf{A}$. Quand l'estimation est parfaite, l'indice d'erreur doit tendre vers $-\infty$ en échelle logarithmique (0 en échelle linéaire). Les résultats affichés résultent du moyennage par méthode de Monte-Carlo sur 100 tirages, en initialisant toujours à partir de la décomposition en valeurs propres généralisées [17]. Au niveau de la Fig. 1, partie gauche (resp. partie droite), nous comparons les performances de l'algorithme proposé (JBDC_{LM}) à celles de l'algorithme JBDC_{GC} [8] (basé sur la méthode du Gradient Conjugué qui ne nécessite pas le calcul des matrices Hessiennes) en fonction du nombre d'itérations pour une taille de bloc $r = 4$ en considérant différents nombres de matrices ($N_m = 5, 20$ et 100) dans le cas non bruité c'est-à-dire pour un RSB de 100 dB (resp. dans le cas bruité avec un RSB de 25 dB).

Nous montrons ensuite au niveau de la courbe de gauche de la Fig. 2 l'influence de la taille N_m de l'ensemble à zéro-bloc dia-

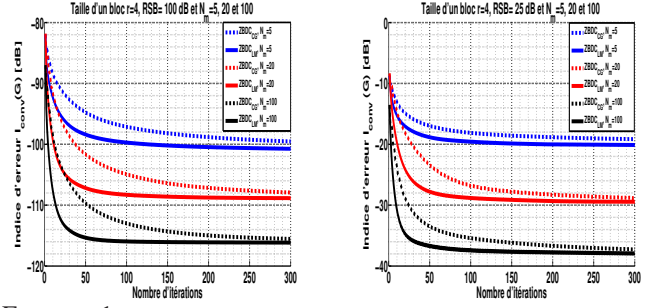


FIGURE 1 – $I_{\text{conv}}(\mathbf{G})$ en dB en fonction du nombre d'itérations pour $N_m = 5, 20$ et 100 matrices, en considérant la taille d'un bloc $r = 4$. Gauche : RSB = 100 dB. Droite : RSB = 25 dB.

gonaliser conjointement sur les résultats. Nous traçons l'évolution des performances en fonction de N_m pour un nombre de blocs $r = 4$ et pour différentes valeurs de RSB = 20, 40 et 100 dB. Sur la courbe de droite de la Fig. 2, nous illustrons l'évolution des performances par rapport à différentes valeurs de RSB pour un nombre de blocs $r = 4$ et pour différents nombres de matrices à zéro bloc-diagonaliser conjointement : $N_m = 5, 20$ et 100 .

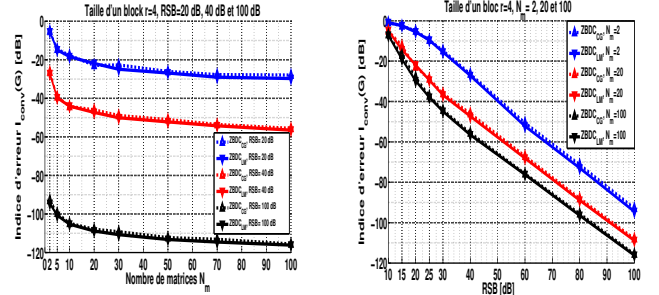


FIGURE 2 – Comportement de l'algorithme JBDC_{LM} . Gauche : $I_{\text{conv}}(\mathbf{G})$ en dB en fonction du nombre N_m de matrices considérées pour différentes valeurs de RSB = 20 dB, 40 dB et 100 dB. Droite : $I_{\text{conv}}(\mathbf{G})$ en fonction du RSB pour des ensembles de taille variable ($N_m = 2, 20$ et 100 matrices).

Ces simulations illustrent le bon comportement de l'algorithme proposé. Nous constatons que plus le RSB est favorable et/ou plus le nombre de matrices N_m à zéro bloc-diagonaliser est grand, meilleurs sont les résultats.

5 Conclusion

Dans cette communication, nous avons proposé un nouvel algorithme pour la zéro bloc diagonalisation conjointe non unitaire, noté JBDC_{LM} , car il repose sur un schéma d'optimisation de type Levenberg-Marquardt. Cet algorithme repose sur un calcul exact des matrices Hessiennes complexes. Le principal avantage de cet algorithme est qu'il généralise le problème de la zéro-diagonalisation conjointe tout en témoignant de bonnes performances dans des situations complexes (cas bruité et/ou faible nombre de matrices dans l'ensemble à zéro bloc-diagonaliser conjointement). Les extensions futures de ces travaux consisteront à illustrer l'intérêt de cet algorithme dans un véritable contexte de séparation aveugle de mélanges convo-

lutifs de sources (plus particulièrement pour les méthodes du domaine temps-fréquence spatial quadratique).

ANNEXE

Considérons quatre matrices carrées $\mathbf{D}_1, \mathbf{D}_2, \mathbf{D}_3$ de taille $M \times M$ et $\mathbf{D}_4$ de taille $M \times N$. $\text{tr}\{\cdot\}$ désigne l'opérateur de trace. Nos développements sont basés sur les propriétés suivantes [14] :

- P₁. $\|\text{Bdiag}\{\mathbf{D}_1\}\|_F^2 = \text{tr}\{\mathbf{D}_1^H \text{Bdiag}\{\mathbf{D}_1\}\}$.
- P₂. $\text{tr}\{\mathbf{D}_1\} = \text{tr}\{\mathbf{D}_1^T\}$.
- P₃. $\text{tr}\{\mathbf{D}_1 + \mathbf{D}_2\} = \text{tr}\{\mathbf{D}_1\} + \text{tr}\{\mathbf{D}_2\}$.
- P₄. $\text{tr}\{\mathbf{D}_1 \mathbf{D}_2 \mathbf{D}_3\} = \text{tr}\{\mathbf{D}_3 \mathbf{D}_1 \mathbf{D}_2\} = \text{tr}\{\mathbf{D}_2 \mathbf{D}_3 \mathbf{D}_1\}$.
- P₅. $\text{tr}\{\mathbf{D}_1^H \mathbf{D}_2\} = (\text{vec}\{\mathbf{D}_1\})^H \text{vec}\{\mathbf{D}_2\}$.
- P₆. $\text{d}(\text{tr}\{\mathbf{D}_1\}) = \text{tr}\{\text{d}(\mathbf{D}_1)\}$.
- P₇. $\text{vec}\{\mathbf{D}_1 + \mathbf{D}_2\} = \text{vec}\{\mathbf{D}_1\} + \text{vec}\{\mathbf{D}_2\}$.
- P₈. $\text{vec}\{\mathbf{D}_1 \mathbf{D}_2 \mathbf{D}_3\} = (\mathbf{D}_3^T \otimes \mathbf{I}_1) \text{vec}\{\mathbf{D}_2\}$.
- P₉. $\text{vec}\{\mathbf{D}_1^T\} = \mathbf{T}_{vec} \text{vec}\{\mathbf{D}_1\}$.
- P₁₀. $\text{vec}\{\text{Bdiag}\{\mathbf{D}_1\}\} = \mathbf{T}_{Bdiag} \text{vec}\{\mathbf{D}_1\}$.
- P₁₁. $\text{d}(\text{vec}\{\mathbf{D}_1\}) = \text{vec}\{\text{d}(\mathbf{D}_1)\}$.
- P₁₂. $(\mathbf{D}_1 \otimes \mathbf{D}_2)^T = \mathbf{D}_1^T \otimes \mathbf{D}_2^T$.
- P₁₃. $(\mathbf{D}_1 \otimes \mathbf{D}_2)^H = \mathbf{D}_1^H \otimes \mathbf{D}_2^H$.
- P₁₄. $(\mathbf{D}_1 \otimes \mathbf{D}_2)(\mathbf{D}_3 \otimes \mathbf{D}_4) = (\mathbf{D}_1 \mathbf{D}_3 \otimes \mathbf{D}_2 \mathbf{D}_4)$.

La dérivée seconde de la fonction de coût (Eq. (1)) est définie comme dans [14] par :

$$d^2 C_{ZBD}(\mathbf{B}) = d\mathcal{D}_B(C_{ZBD}(\mathbf{B})) \text{dvec}\{\mathbf{B}\} + d\mathcal{D}_B^*(C_{ZBD}(\mathbf{B})) \text{dvec}\{\mathbf{B}^*\}, \quad (11)$$

où $\mathcal{D}_B(C_{ZBD}(\mathbf{B}))$ et $\mathcal{D}_B^*(C_{ZBD}(\mathbf{B}^*))$ sont les matrices jacobienues de $C_{ZBD}(\mathbf{B})$, définies comme dans [14],

$$\begin{aligned} \mathcal{D}_B(C_{ZBD}(\mathbf{B})) &= \left(\text{vec} \left\{ \frac{\partial C_{ZBD}(\mathbf{B})}{\partial \mathbf{B}} \right\} \right)^T, \\ \mathcal{D}_B^*(C_{ZBD}(\mathbf{B})) &= \left(\text{vec} \left\{ \frac{\partial C_{ZBD}(\mathbf{B})}{\partial \mathbf{B}^*} \right\} \right)^T. \end{aligned} \quad (12)$$

Les dérivées partielles $\frac{\partial C_{ZBD}(\mathbf{B})}{\partial \mathbf{B}}$ and $\frac{\partial C_{ZBD}(\mathbf{B})}{\partial \mathbf{B}^*}$ de la fonction de coût (Eq. (1)) ont été calculées dans [8] :

$$\frac{\partial C_{ZBD}(\mathbf{B})}{\partial \mathbf{B}} = \sum_{i=1}^{N_m} (\text{Bdiag}_{(n)}\{\mathbf{B} \mathbf{X}_i \mathbf{B}^H\})^T \mathbf{B}^* \mathbf{X}_i^* + \sum_{i=1}^{N_m} (\text{Bdiag}_{(n)}\{\mathbf{B} \mathbf{X}_i \mathbf{B}^H\})^* \mathbf{B}^* \mathbf{X}_i^T, \quad (13)$$

$$\frac{\partial C_{ZBD}(\mathbf{B})}{\partial \mathbf{B}^*} = \sum_{i=1}^{N_m} (\text{Bdiag}_{(n)}\{\mathbf{B} \mathbf{X}_i \mathbf{B}^H\}) \mathbf{B} \mathbf{X}_i^H + \sum_{i=1}^{N_m} (\text{Bdiag}_{(n)}\{\mathbf{B} \mathbf{X}_i \mathbf{B}^H\})^H \mathbf{B} \mathbf{X}_i. \quad (14)$$

En utilisant les propriétés ci-dessus ($\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_{14}$) et l'expression des dérivées partielles $\frac{\partial C_{ZBD}(\mathbf{B})}{\partial \mathbf{B}}$ et $\frac{\partial C_{ZBD}(\mathbf{B})}{\partial \mathbf{B}^*}$, la dérivée des matrices jacobienues vaut alors :

$$\begin{aligned} d\mathcal{D}_B(C_{ZBD}(\mathbf{B})) &= \sum_{i=1}^{N_m} (\text{vec}\{\text{d}\mathbf{B}\})^T \left[(\mathbf{T}_{vec})^T (\mathbf{I}_N \otimes \mathbf{X}_i \mathbf{B}^H) \mathbf{T}_{Bdiag}^T (\mathbf{B}^* \mathbf{X}_i^* \otimes \mathbf{I}_N) \right. \\ &\quad + (\mathbf{T}_{vec})^T (\mathbf{I}_N \otimes \mathbf{X}_i^H \mathbf{B}^H) \mathbf{T}_{Bdiag}^T (\mathbf{B}^* \mathbf{X}_i^T \otimes \mathbf{I}_N) \\ &\quad + (\text{vec}\{\text{d}\mathbf{B}^*\})^T \left[(\mathbf{X}_i^T \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{Bdiag}^T (\mathbf{B}^* \mathbf{X}_i^* \otimes \mathbf{I}_N) \right. \\ &\quad + (\mathbf{X}_i^* \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{Bdiag}^T (\mathbf{B}^* \mathbf{X}_i^T \otimes \mathbf{I}_N) \\ &\quad + (\mathbf{X}_i^* \otimes (\text{Bdiag}_{(n)}\{\mathbf{B} \mathbf{X}_i \mathbf{B}^H\})) \\ &\quad \left. + (\mathbf{X}_i^T \otimes (\text{Bdiag}_{(n)}\{\mathbf{B} \mathbf{X}_i^H \mathbf{B}^H\})) \right]. \end{aligned} \quad (15)$$

$$\begin{aligned} d\mathcal{D}_B^*(C_{ZBD}(\mathbf{B})) &= \sum_{i=1}^{N_m} \left[(\mathbf{X}_i^* \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{Bdiag} (\mathbf{B}^* \mathbf{X}_i^T \otimes \mathbf{I}_N) \text{vec}\{\text{d}\mathbf{B}\} \right. \\ &\quad + (\mathbf{X}_i^* \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{Bdiag} (\mathbf{I}_N \otimes \mathbf{B} \mathbf{X}_i) (\mathbf{T}_{vec}) \text{vec}\{\text{d}\mathbf{B}^*\} \\ &\quad + (\mathbf{X}_i^T \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{Bdiag} (\mathbf{B}^* \mathbf{X}_i^* \otimes \mathbf{I}_N) \text{vec}\{\text{d}\mathbf{B}\} \\ &\quad + (\mathbf{X}_i^T \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{Bdiag} (\mathbf{I}_N \otimes \mathbf{B} \mathbf{X}_i^H) (\mathbf{T}_{vec}) \text{vec}\{\text{d}\mathbf{B}^*\} \\ &\quad + (\mathbf{I}_N \otimes (\text{Bdiag}_{(n)}\{\mathbf{B} \mathbf{X}_i \mathbf{B}^H\})) (\mathbf{X}_i^* \otimes \mathbf{I}_N) \text{vec}\{\text{d}\mathbf{B}\} \\ &\quad \left. + (\mathbf{I}_N \otimes (\text{Bdiag}_{(n)}\{\mathbf{B} \mathbf{X}_i \mathbf{B}^H\})^H) (\mathbf{X}_i^T \otimes \mathbf{I}_N) \text{vec}\{\text{d}\mathbf{B}\} \right]^T. \end{aligned} \quad (16)$$

Les Eqs. (7) et (11) conduisent aux résultats présentés au niveau des Eqs. (8) et (9). Enfin, nous pouvons en déduire les quatre matrices Hessiennes complexes sous la forme des Eqs. (5) et (6).

Références

- [1] A. Belouchrani, K. Abed-Meraïm, M. G. Amin, and A. Zoubir. Joint anti-diagonalization for blind source separation. In *Proc. IEEE ICASSP*, pages 2196–2199, Salt Lake City, UT, May 2001.
- [2] A. Belouchrani, K. Abed-Meraïm, J.-F. Cardoso, and E. Moulines. A blind source separation technique using second order statistics. *IEEE Trans. Signal Proc.*, 45 :434–444, Feb. 1997.
- [3] J.-F. Cardoso and A. Souloumiac. Blind beamforming for non-gaussian signals. *IEEE Proc.-F*, 40(6) :362–370, 1993.
- [4] G. Chabriel and J. Barrère. A direct algorithm for nonorthogonal approximate joint diagonalization. *IEEE Trans. Signal Proc.*, 60(1) :39–47, 2012.
- [5] G. Chabriel, J. Barrère, N. Thirion-Moreau, and E. Moreau. Algebraic joint zero-diagonalization and blind source separation. *IEEE Trans. Signal Proc.*, 56(3) :980–989, March 2008.
- [6] G. Chabriel, M. Kleinstueber, E. Moreau, H. Shen, P. Tichavsky, and A. Yeredor. Joint matrices decompositions and blind source separation : A survey of methods, identification, and applications. *IEEE Signal Proc. Mag.*, 31(3) :34–43, 2014.
- [7] O. Cherrak, H. Ghennioui, E.-H. Abarkan, and N. Thirion-Moreau. Non-unitary joint block diagonalization of matrices using a levenberg-marquardt algorithm. In *Proc. EUSIPCO*, pages 1–5, Sept. 2013.
- [8] O. Cherrak, H. Ghennioui, E.-H. Abarkan, N. Thirion-Moreau, and E. Moreau. Non-unitary joint zero-block diagonalization of matrices using a conjugate gradient algorithm. In *Proc. EUSIPCO*, Sept. 2015.
- [9] P. Comon. Independent component analysis, a new concept ? *Signal Proc.*, 36 :287–314, 1994.
- [10] L. De Lathauwer, C. Févotte, B. De Moor, and J. Vandewalle. Jacobi algorithm for block diagonalization in blind identification. In *Proc. Symp. Inform. Theory in the Benelux*, pages 155–162, Louvain-la-Neuve, Belgium, May 2002.
- [11] E. M. Fadaïli, N. Thirion-Moreau, and E. Moreau. Non orthogonal zero-diagonalization for sources separation based on time-frequency representation. In *Proc. IEEE ICASSP*, volume V, pages 297–300, Philadelphia, USA, March 2005.
- [12] E. M. Fadaïli, N. Thirion-Moreau, and E. Moreau. Non-orthogonal joint diagonalization/zero-diagonalization for source separation based on time-frequency distributions. *IEEE Trans. Signal Proc.*, 55(5), May 2007.
- [13] H. Ghennioui, N. Thirion-Moreau, E. Moreau, and D. Aboutajdine. Gradient-based joint block diagonalization algorithms : Application to blind separation of fir convolutive mixtures. *Signal Proc.*, 90(6) :1836–1849, June 2010.
- [14] A. Hjørungnes. *Complex-Valued Matrix Derivatives With Applications in Signal Processing and Communications*. Cambridge, University Press, Feb. 2011.
- [15] D. W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal on Applied Math.*, 11 :431–441, 1963.
- [16] E. Moreau. A generalization of joint-diagonalization criteria for source separation. *IEEE Trans. Signal Proc.*, 49(3) :530–541, March 2001.
- [17] D. Nion. A tensor framework for non-unitary joint block diagonalization. *IEEE Trans. Signal Proc.*, 59(10) :4585–4594, Oct. 2011.
- [18] J. R. Shewchuk. Introduction to the conjugate gradient method without the agonizing pain. Technical report, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA, 1994.
- [19] A. Souloumiac. Nonorthogonal joint diagonalization by combining givens and hyperbolic rotations. *IEEE Trans. Signal Proc.*, 57(6) :2222–2231, June 2009.
- [20] A. Yeredor, A. Ziehe, and K.-R. Müller. Approximate joint diagonalization using natural gradient approach. In *Lecture Notes in Computer Science, in Proceedings ICA*, pages 89–96, Granada, Spain, May 2004.
- [21] W. T. Zhang and S. T. Lou. A low complexity iterative algorithm for joint zero-diagonalization. *IEEE Signal Proc. Lett.*, 19(2) :115–118, Jan. 2012.

NON-UNITARY JOINT ZERO-BLOCK DIAGONALIZATION OF MATRICES USING A CONJUGATE GRADIENT ALGORITHM

O. Cherrak, H. Ghennioui, E.-H. Abarkan

Université Sidi Mohamed Ben Abdellah
FSTF, LSSC
Faculté des Sciences et Techniques de Fès,
Route Immouzzar, B.P. 2202, Fès-Maroc
{cherrakomar, ghennioui}@gmail.com

N. Thirion-Moreau, E. Moreau

Université Aix-Marseille, CNRS,
ENSAM, LSIS, UMR 7296,
13397 Marseille-France
Université de Toulon, CNRS, LSIS,
UMR 7296, 83957 La Garde-France
{thirion, moreau}@univ-tln.fr

ABSTRACT

This communication addresses a new problem which is the Non-Unitary Joint Zero-Block Diagonalization of a given set of complex matrices. This problem can occur in fields of applications such as blind separation of convolutive mixtures of sources and generalizes the non unitary Joint Zero-Diagonalization problem.

We present a new method based on the Conjugate Gradient algorithm. Our algorithm uses a numerical diagram of optimization which requires the calculation of the complex gradient matrix. The main advantages of the proposed method stem from the conjugate gradient properties: it is fast, stable and robust. Computer simulations are provided in order to illustrate the good behavior of the proposed method in different contexts. Two cases are studied: in the first scenario, a set of exactly zero-block-diagonal matrices are considered, then these matrices are progressively perturbed by an additive gaussian noise.

Index Terms— Joint zero-block diagonalization, matrix decompositions, conjugate gradient algorithm, linear convolutive mixtures.

1. INTRODUCTION

In the recent years, the problem of the joint decomposition of matrices (or tensors) sets has often arisen in the signal processing field, especially in blind source separation and array processing applications. One of the first considered problems was the Joint Diagonalization (JD) of a given matrix set under the unitary constraint, leading to the nowadays well-known JADE (Joint Approximate Diagonalization of Eigenmatrices) [1] and SOBI (Second Order Blind Identification) [2] algorithms. The following works have addressed either the problem of the JD of tensors [3][4] or the problem of JD of matrices but discarding the unitary constraint [5][6][7]. This first particular type of matrices decompositions is useful both

in sources localization and direction finding problems and in blind sources separation of instantaneous mixtures.

A second type of matrices decompositions, namely the joint block-diagonalization (JBD), is encountered both in the wide-band sources localization in the presence of a correlated noise and in the blind separation of convolutive mixtures (or multi-dimensional deconvolution) problems. Several algorithms have been developed, under different assumptions about the considered matrix set (the matrices can be either positive definite or Hermitian) and about the block-diagonalizer (it is assumed unitary [8] or not [9][10][11][12]). A third type of matrices decompositions, namely the joint zero-diagonalization (JZD), has proven to be useful in blind source separation, telecommunication and cryptography. The first suggested algorithms operated under the unitary constraint [13], since they were applied after a classical pre-whitening stage. But such a preliminary pre-whitening step establishes a bound with regard to the best reachable performances in the context of BSS that is the reason why the unitary constraint was soon discarded, leading to several other solutions [14][15][16]. In this communication, our purpose is to generalize the non unitary joint zero diagonalization approach suggested in [17][18] to the non-unitary joint zero-block diagonalization (JZBD) of several complex (not necessarily Hermitian) matrices (the zero-block-diagonalizer is not assumed unitary). It involves the choice of a well-chosen cost function and the calculation of quantities such as the complex gradient matrix. The main advantage of this approach remains its rather generic aspect since it encompasses the aforementioned JZD problem.

2. PROBLEM STATEMENT

2.1. Non-unitary joint zero-block diagonalization

The problem of the non-unitary joint zero-block diagonalization is stated in the following way. We consider a set $\mathcal{N}$ of N_m ($N_m \in \mathbb{N}^*$) square matrices $\mathbf{X}_i \in \mathbb{C}^{M \times M}$. For $i \in \{1, \dots, N_m\}$, these matrices all admit the following

decomposition form: $\mathbf{X}_i = \mathbf{A}\mathbf{Z}_i\mathbf{A}^H$, where $(\cdot)^H$ stands for the transpose conjugate operator and the matrices $\mathbf{Z}_i =$

$$\begin{pmatrix} \mathbf{0}_{11} & \mathbf{Z}_{i,12} & \cdots & \mathbf{Z}_{i,1r} \\ \mathbf{Z}_{i,21} & \mathbf{0}_{22} & \ddots & \mathbf{Z}_{i,2r} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{Z}_{i,r1} & \cdots & \vdots & \mathbf{0}_{rr} \end{pmatrix}, \text{ for all } i \in \{1, \dots, N_m\}$$

are $(N \times N)$ zero-block diagonal matrices with r the number of considered blocks ($r \in \mathbb{N}^*$), $\mathbf{Z}_{i,kl}$, for all $k, l \in \{1, \dots, r\}$ are $n_k \times n_l$ matrices so that $n_1 + \dots + n_r = N$ where $\mathbf{0}_{kk}$ denotes the $n_k \times n_k$ square null matrix. $\mathbf{A}$ is a $M \times N$ ($M \geq N$) full rank matrix and the $N \times M$ matrix $\mathbf{B}$ is its pseudo-inverse (or generalized Moore-Penrose inverse). The set of the N_m square matrices $\mathbf{Z}_i \in \mathbb{C}^{N \times N}$ is denoted by $\mathcal{Z}$. The block sizes n_j for all $j = 1, \dots, r$ are assumed known. The NU – JZBD problem consists of estimating $\mathbf{A}$ and the zero-block-diagonal matrices belonging to $\mathcal{Z}$ from only the matrix set $\mathcal{N}$. To tackle that problem, we propose, here, to consider the following cost function:

$$\mathcal{C}_{ZBD}(\mathbf{B}) = \sum_{i=1}^{N_m} \|\text{Bdiag}_{(\mathbf{n})}\{\mathbf{B}\mathbf{X}_i\mathbf{B}^H\}\|_F^2, \quad (1)$$

where the matrix operator $\text{Bdiag}_{(\mathbf{n})}\{\cdot\}$ is defined as follows:

$$\text{Bdiag}_{(\mathbf{n})}\{\mathbf{X}\} = \begin{pmatrix} \mathbf{X}_{11} & \mathbf{0}_{12} & \cdots & \mathbf{0}_{1r} \\ \mathbf{0}_{21} & \mathbf{X}_{22} & \ddots & \mathbf{0}_{2r} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0}_{r1} & \mathbf{0}_{r2} & \cdots & \mathbf{X}_{rr} \end{pmatrix}, \quad (2)$$

where $\mathbf{X}$ is a $N \times N$ square matrix whose block components $\mathbf{X}_{ij}$ for all $i, j = 1, \dots, r$ are $n_i \times n_j$ matrices (and $n_1 + \dots + n_r = N$). Finally, we denote by $\mathbf{n} = (n_1, n_2, \dots, n_r)$. Using the cost function given Eq. (1) is rather advantageous since one single matrix is finally involved (direct estimation of the joint zero-block-diagonalizer matrix).

3. AN ALGORITHM TO SOLVE THE NU-JZBD

The cost function given in Eq. (1) has to be minimized to estimate the zero-block-diagonalizer matrix $\mathbf{B}$. To that aim, we suggest a new algorithm based on the conjugate gradient approach [19]. This iterative optimization method is well-known for its robustness and effectiveness. It has been widely used in various fields (*e.g.* neural networks [20], array processing [21] or blind sources separation [22]). But first, the complex gradient matrix $\mathbf{G} = \nabla_a \mathcal{C}_{ZBD}(\mathbf{B}) = 2 \frac{\partial \mathcal{C}_{ZBD}(\mathbf{B})}{\partial \mathbf{B}^*}$ or its vectorization denoted by $\mathbf{g} = \text{vec}(\nabla_a \mathcal{C}_{ZBD}(\mathbf{B}))$ has to be calculated. The operator $\text{vec}(\cdot)$ stacks the columns of a matrix into a vector. This calculation is performed in Appendix A. It is established that $\nabla_a \mathcal{C}_{ZBD}(\mathbf{B})$ is equal to:

$$2 \sum_{i=1}^{N_m} \left[\text{Bdiag}_{(\mathbf{n})}\{\mathbf{B}\mathbf{X}_i\mathbf{B}^H\}\mathbf{B}\mathbf{X}_i^H + \left(\text{Bdiag}_{(\mathbf{n})}\{\mathbf{B}\mathbf{X}_i\mathbf{B}^H\} \right)^H \mathbf{B}\mathbf{X}_i \right].$$

3.1. Updating rule of the conjugate gradient algorithm

The matrix $\mathbf{B}$ is re-evaluated at each iteration m . Thus, it is denoted by $\mathbf{B}^{(m)}$ or by $\mathbf{b}^{(m)}$ when the vector $\mathbf{b}^{(m)} =$

$\text{vec}(\mathbf{B}^{(m)})$ is considered instead of the matrix. The CG algorithm (with restarts) used to estimate $\mathbf{b}$ thus reads:

Step 1. Given $\mathbf{b}^{(0)}$, compute $\mathbf{g}^{(0)}$ and set $\mathbf{d}^{(0)} = -\mathbf{g}^{(0)}$

Step 2. For $m = 0, 1, \dots, n-1$

$$\begin{cases} \mathbf{b}^{(m+1)} &= \mathbf{b}^{(m)} - \mu^{(m)} \mathbf{d}_B^{(m)}, \\ \mathbf{d}_B^{(m+1)} &= -\mathbf{g}^{(m+1)} + \beta^{(m)} \mathbf{d}_B^{(m)}. \end{cases} \quad (3)$$

Step 3. If $m = n-1$ replace $\mathbf{b}^{(0)}$ by $\mathbf{b}^{(n)}$ and go back to Step 1 (*i.e.* restart).

μ is a positive small factor called the step-size and $\mathbf{d}_B$ is the direction of search. We show in Section 3.2 how the optimal step-size μ_{opt} can be calculated at each iteration. In exact line search methods, several expressions of β have been suggested among which are the Fletcher-Reeves (β_{FR}), the Polak-Ribière (β_{PR}) [23][24] and the Dai-Yuan (β_{DY}) formula [25]:

$$\beta_{\text{FR}}^{(m+1)} = \frac{(\mathbf{g}^{(m+1)})^H \mathbf{g}^{(m+1)}}{(\mathbf{g}^{(m)})^H \mathbf{g}^{(m)}}, \quad (4)$$

$$\beta_{\text{PR}}^{(m+1)} = \frac{(\mathbf{g}^{(m+1)} - \mathbf{g}^{(m)})^H \mathbf{g}^{(m+1)}}{(\mathbf{g}^{(m)})^H \mathbf{g}^{(m)}}, \quad (5)$$

$$\beta_{\text{DY}}^{(m+1)} = \frac{(\mathbf{g}^{(m+1)})^H \mathbf{g}^{(m+1)}}{(\mathbf{d}_B^{(m)})^H (\mathbf{g}^{(m+1)} - \mathbf{g}^{(m)})}. \quad (6)$$

The restarting aspect is important for the convergence of the algorithm since in general one cannot guarantee that the directions $\mathbf{d}^{(k)}$ are effectively descent directions. But since a pure steepest descent step is taken at each restart by setting $\mathbf{d}_B^{(m)} = -\mathbf{g}^{(m)}$ the local convergence is assured.

3.2. Exact Line Search: seek of the optimal step-size

The optimal step size, say μ_{opt} , is used to decrease the total number of iterations needed to reach convergence. It minimizes the polynomial $\mathcal{C}_{ZBD}(\mathbf{P})$ where $\mathbf{P} = \mathbf{B}^{(m)} - \mu \mathbf{D}_B^{(m)}$ and $\mathbf{d}_B^{(m)} = \text{vec}(\mathbf{D}_B^{(m)})$. In our case, this quantity is a 4th order polynomial, and its derivative a 3rd order polynomial. They are respectively given by:

$$\mathcal{C}_{ZBD}(\mathbf{B} - \mu \mathbf{D}_B) = a_0 + a_1\mu + a_2\mu^2 + a_3\mu^3 + a_4\mu^4, \quad (7)$$

$$\frac{\partial \mathcal{C}_{ZBD}(\mathbf{B} - \mu \mathbf{D}_B)}{\partial \mu} = 4a_4\mu^3 + 3a_3\mu^2 + 2a_2\mu + a_1, \quad (8)$$

where the five coefficients a_i for $i = 0, \dots, 4$ are equal to (see Appendix B):

$$a_0 = \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{K}_3^H \text{Bdiag}_{(\mathbf{n})}\{\mathbf{K}_3\} \right\},$$

$$a_1 = - \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{K}_1 \text{Bdiag}_{(\mathbf{n})}\{\mathbf{K}_3\} + \mathbf{K}_3^H \mathbf{K}_2 \right\},$$

$$a_2 = \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{K}_0^H \text{Bdiag}_{(\mathbf{n})}\{\mathbf{K}_3\} + \mathbf{K}_3^H \text{Bdiag}_{(\mathbf{n})}\{\mathbf{K}_0\} + \mathbf{K}_1 \mathbf{K}_2 \right\},$$

$$a_3 = - \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{K}_0^H \mathbf{K}_2 + \mathbf{K}_1 \text{Bdiag}_{(\mathbf{n})}\{\mathbf{K}_0\} \right\},$$

$$a_4 = \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{K}_0^H \text{Bdiag}_{(\mathbf{n})}\{\mathbf{K}_0\} \right\},$$

where:

$$\begin{aligned} \mathbf{K}_0 &= \mathbf{D}_B^{(m)} \mathbf{X}_i (\mathbf{D}_B^{(m)})^H, \\ \mathbf{K}_1 &= \mathbf{D}_B^{(m)} \mathbf{X}_i^H \mathbf{B}^H + \mathbf{B} \mathbf{X}_i^H (\mathbf{D}_B^{(m)})^H, \\ \mathbf{K}_2 &= \text{Bdiag}_{(n)} \{ \mathbf{B} \mathbf{X}_i (\mathbf{D}_B^{(m)})^H + \mathbf{D}_B^{(m)} \mathbf{X}_i \mathbf{B}^H \}, \\ \mathbf{K}_3 &= \mathbf{B} \mathbf{X}_i \mathbf{B}^H. \end{aligned}$$

The optimal step-size μ_{opt} corresponds to the root of (8) which minimizes (7).

3.3. Algorithm

The non-unitary JZBD algorithm (denoted by JZBD_{CG}) is summarized below:

Data: N_m square matrices $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{N_m}$, stopping criterion ϵ , step-size μ , max. number of iterations M_{max}

Result: Estimation of joint zero block diagonalizer $\mathbf{B}$

initialize: $\mathbf{B}^{(0)}; \mathbf{D}^{(0)}; m = 0;$

repeat

if $m \bmod M_0 = 0$ **then**

restart

else

Calculate optimal step-size $\mu_{\text{opt}}^{(m)}$

Compute gradient matrix $\nabla_a C_{ZBD}(\mathbf{B}^{(m+1)})$

Compute matrix $\mathbf{B}^{(m+1)}$

Compute $\beta_{\text{PR}}^{(m)}$

Compute the search direction $\mathbf{D}_B^{(m+1)}$

$m = m + 1;$

end

until $(\|\mathbf{B}^{(m+1)} - \mathbf{B}^{(m)}\| \leq \epsilon) \text{ or } (m \geq M_{\text{max}});$

4. COMPUTER SIMULATIONS

We present simulations to illustrate the effectiveness of the suggested algorithm. We consider a set $\mathcal{Z}$ of $N_m = 5$ (resp. 20 and 100) matrices, randomly chosen according to a Gaussian law of mean 0 and variance 1. First, these matrices are considered as exactly zero block-diagonal, then a random matrix of noise is added (drawn from a gaussian law of mean 0 and variance σ_b^2). The signal to noise ratio is defined as $\text{SNR} = 10 \log(\frac{1}{\sigma_b^2})$. To measure the quality of the estimation, the ensuing error index (which is an extension of the one introduced in [4]) is used:

$$\begin{aligned} I_{\text{conv}}(\mathbf{G}) &= \frac{1}{r(r-1)} \left[\sum_{i=1}^r \left(\sum_{j=1}^r \frac{\|(\mathbf{G})_{i,j}\|_F^2}{\max_{\ell} \|(\mathbf{G})_{i,\ell}\|_F^2} - 1 \right) \right. \\ &\quad \left. + \sum_{j=1}^r \left(\sum_{i=1}^r \frac{\|(\mathbf{G})_{i,j}\|_F^2}{\max_{\ell} \|(\mathbf{G})_{\ell,j}\|_F^2} - 1 \right) \right], \end{aligned}$$

where $(\mathbf{G})_{i,j}$ for all $i, j \in \{1, \dots, r\}$ is the (i, j) -th block matrix of $\mathbf{G} = \hat{\mathbf{B}}\mathbf{A}$. The best results are obtained when the error index $I_{\text{conv}}(\cdot)$ is close to 0 in a linear scale ($-\infty$ in a logarithmic scale). All the displayed results have been averaged over 100 Monte-Carlo trials. For all simulations we have considered a size r of blocks equal to 4.

First, we plot on the Fig.1 the evolution of the performance of the suggested algorithm versus the iterations for different sizes of the matrix sets ($N_m = 5, 20$ and 100). In this example, we consider the noiseless case where $\text{SNR} = 100$ dB. Whereas, on the Fig. 2, the same study is performed in a noisy context ($\text{SNR} = 25$ dB). In both cases, the algorithm is initialized thanks to the generalized eigenvalue decomposition [10].

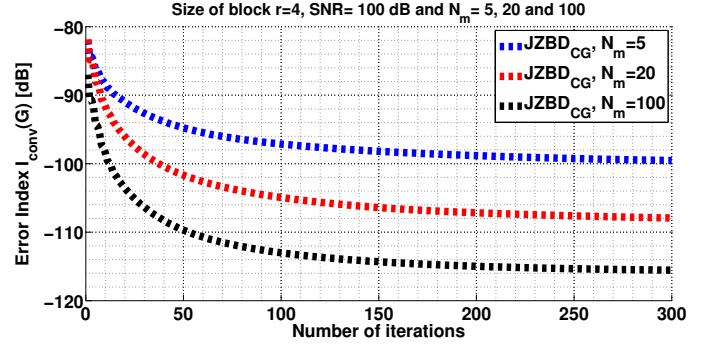


Fig. 1. Evolution of the error index $I_{\text{conv}}(\mathbf{G})$ versus the number of iterations in the noiseless case ($\text{SNR} = 100$ dB) and for $N_m = 5, 20$ and 100 matrices. We consider the case of $r = 4$ blocks.

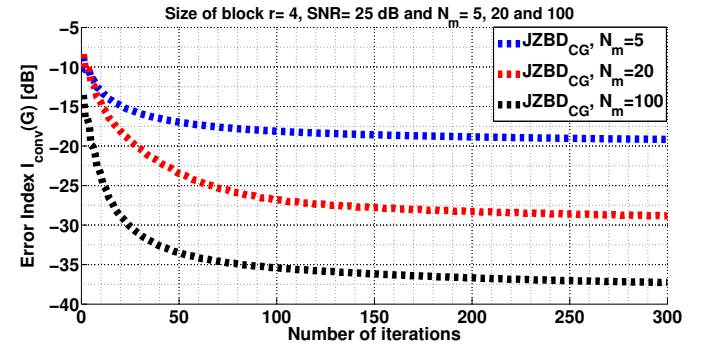


Fig. 2. Evolution of the error index $I_{\text{conv}}(\mathbf{G})$ versus the number of iterations in the noisy case ($\text{SNR} = 25$ dB) and for $N_m = 5, 20$ and 100 matrices. We consider the case of $r = 4$ blocks.

In the noiseless case, we can observe the good stability and convergence of the JZBD_{CG} algorithm for $N_m = 5$ (resp. 20 and 100), since it reaches -100 dB (resp. -108 dB and -115 dB). In a noisy context, we observe the same kind of behavior, except that the performance are deteriorated and bounded (-37 dB instead of -115 dB).

Finally, on Fig.3, we emphasize the influence of the SNR. We display the evolution of the error index versus the number of matrices N_m . Different values of the SNR are considered (25 dB, 40 dB and 100 dB). After that, we show the influence of the size N_m of the matrix set to be joint zero block-diagonalized. We display the evolution of the error index versus the SNR for different sizes of the matrix sets ($N_m = 5, 20$ and 100). These charts illustrate the behavior of the proposed algorithm *i.e.* an increase of the performance when bigger subsets of matrices are considered or when the SNR is

improved.

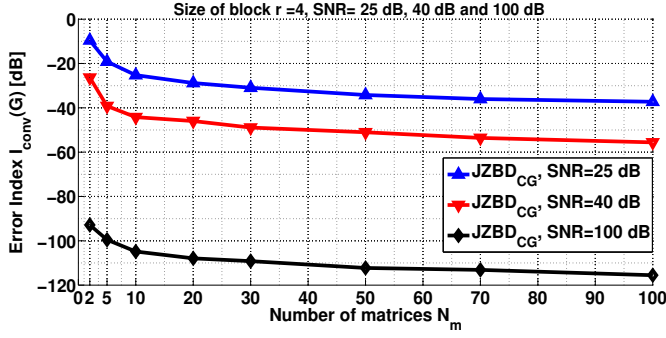


Fig. 3. Evolution of the error index $I_{conv}(\mathbf{G})$ versus the number, N_m , of matrices that are used for SNR = 25 dB, 40 dB and 100 dB. We consider the case of $r = 4$ blocks.

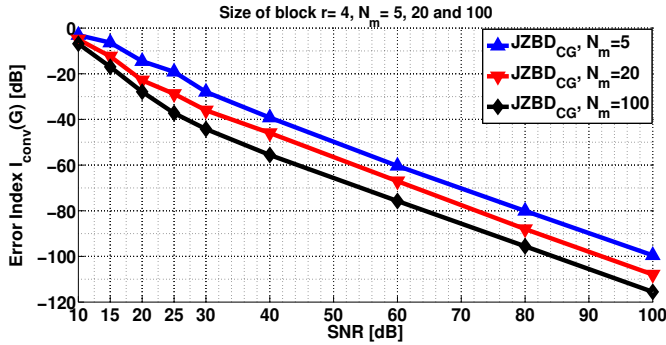


Fig. 4. Evolution of the error index $I_{conv}(\mathbf{G})$ versus the SNR for different sizes of the matrix sets $N_m = 5, 20$ and 100 . We consider the case of $r = 4$ blocks.

5. CONCLUSION

In this communication, we have suggested a new NU – JZBD algorithm called JZBD_{CG}. It relies on a conjugate gradient optimization scheme which requires the calculation of the complex gradient matrix. The main advantage of the JZBD_{CG} algorithm is its more general aspect as well as the fact that it leads to rather good performances even in difficult situations (noise and/or few matrices to be joint zero block-diagonalized). Further works will consist of showing the interest of this algorithm in a true blind multi-dimensional deconvolution context (blind separation of convolutive mixtures of sources).

APPENDIX

Considering three $(M \times M)$ square matrices $\mathbf{D}_1$, $\mathbf{D}_2$ and $\mathbf{D}_3$ and two rectangular matrices $\mathbf{D}_4$ ($M \times N$) and $\mathbf{D}_5$ ($N \times M$) and a square $N \times N$ matrix $\mathbf{D}_6$, let $\text{tr}\{\cdot\}$, $\text{d}\{\cdot\}$, $\text{vec}\{\cdot\}$, $\text{Bdiag}_{(n)}\{\cdot\}$ respectively denote the trace operator, the differential operator, the vec-operator, the block-diagonal operator defined in Eq. (2) and $\mathbf{T}_{\text{Bdiag}} = \text{diag}\{\text{vec}(\text{Bdiag}\{\mathbf{1}_N\})\}$,

$\mathbf{1}_N$ is the $N \times N$ matrix whose components are all ones. Our developments are based on the ensuing properties:

- P₁. $\|\text{Bdiag}_{(n)}\{\mathbf{D}_1\}\|_F^2 = \text{tr}\{(\text{Bdiag}_{(n)}\{\mathbf{D}_1\})^H \text{Bdiag}_{(n)}\{\mathbf{D}_1\}\} = \text{tr}\{\mathbf{D}_1^H \text{Bdiag}_{(n)}\{\mathbf{D}_1\}\}.$
- P₂. $\text{tr}\{\mathbf{D}_1\} = \text{tr}\{\mathbf{D}_1^T\}.$
- P₃. $\text{tr}\{\mathbf{D}_1 + \mathbf{D}_2\} = \text{tr}\{\mathbf{D}_1\} + \text{tr}\{\mathbf{D}_2\}.$
- P₄. $\text{tr}\{\mathbf{D}_1 \mathbf{D}_2 \mathbf{D}_3\} = \text{tr}\{\mathbf{D}_3 \mathbf{D}_1 \mathbf{D}_2\} = \text{tr}\{\mathbf{D}_2 \mathbf{D}_3 \mathbf{D}_1\} \Rightarrow \text{tr}\{\mathbf{D}_1 \mathbf{D}_2\} = \text{tr}\{\mathbf{D}_2 \mathbf{D}_1\}.$
- P_{4'}. $\text{tr}\{\mathbf{D}_4 \mathbf{D}_5\} = \text{tr}\{\mathbf{D}_5 \mathbf{D}_4\}.$
- P₅. $\text{tr}\{\mathbf{D}_1^H \mathbf{D}_2\} = (\text{vec}\{\mathbf{D}_1\})^H \text{vec}\{\mathbf{D}_2\}.$
- P₆. $\text{vec}\{\text{Bdiag}_{(n)}\{\mathbf{D}_6\}\} = \mathbf{T}_{\text{Bdiag}} \text{vec}\{\mathbf{D}_6\}.$
- P₇. $\text{d}\{\mathbf{D}_1^H\} = (\text{d}\{\mathbf{D}_1\})^H.$
- P₈. $\text{d}\{\mathbf{D}_1^*\} = (\text{d}\{\mathbf{D}_1\})^*.$
- P₉. $\text{d}\{\mathbf{D}_1 \mathbf{D}_2\} = \text{d}\{\mathbf{D}_1\} \mathbf{D}_2 + \mathbf{D}_1 \text{d}\{\mathbf{D}_2\}.$
- P₁₀. $\text{d}\{\mathbf{D}_1 + \mathbf{D}_2\} = \text{d}\{\mathbf{D}_1\} + \text{d}\{\mathbf{D}_2\}.$
- P₁₁. $\text{d}\{\text{tr}\{\mathbf{D}_1\}\} = \text{tr}\{\text{d}\{\mathbf{D}_1\}\}.$
- P₁₂. $\text{d}\{\text{vec}\{\mathbf{D}_1\}\} = \text{vec}\{\text{d}\{\mathbf{D}_1\}\}.$
- P₁₃. if $f(\mathbf{Z}, \mathbf{Z}^*) = \text{tr}\{\mathbf{D}_1^T \mathbf{Z} + \mathbf{Z}^H \mathbf{D}_2\}$ then $\text{d}\{f(\mathbf{Z}, \mathbf{Z}^*)\} = \text{tr}\{\mathbf{D}_1^T d\mathbf{Z} + \mathbf{D}_2^T d\mathbf{Z}^*\} \Rightarrow \frac{\partial f}{\partial \mathbf{Z}} = \mathbf{D}_1$ and $\frac{\partial f}{\partial \mathbf{Z}^*} = \mathbf{D}_2.$
- P₁₄. $(\mathbf{D}_1 \mathbf{D}_2)^H = \mathbf{D}_2^H \mathbf{D}_1^H.$
- P₁₅. $(\mathbf{D}_1 + \mathbf{D}_2)^H = (\mathbf{D}_1^H + \mathbf{D}_2^H).$
- P₁₆. $\text{Bdiag}_{(n)}\{\mathbf{D}_1 + \mathbf{D}_2\} = \text{Bdiag}_{(n)}\{\mathbf{D}_1\} + \text{Bdiag}_{(n)}\{\mathbf{D}_2\}.$

A. Calculation of the complex gradient matrix of the cost function $\mathcal{C}_{ZBD}(\mathbf{B})$

Using the property P₁, P₃, P₉, P₁₀ and P₁₁, the differential of cost function is rewritten as:

$$\begin{aligned} d\{\mathcal{C}_{ZBD}(\mathbf{B})\} &= \sum_{i=1}^{N_m} \text{tr}\left\{d\{(\mathbf{B}\mathbf{X}_i \mathbf{B}^H)^H \text{Bdiag}_{(n)}\{\mathbf{B}\mathbf{X}_i \mathbf{B}^H\}\}\right\}, \\ &= \mathcal{F}(\mathbf{B}) + \mathcal{G}(\mathbf{B}), \end{aligned} \quad (9)$$

where, $\mathcal{F}(\mathbf{B}) = \sum_{i=1}^{N_m} \text{tr}\left\{d\{(\mathbf{B}\mathbf{X}_i \mathbf{B}^H)^H\} \text{Bdiag}_{(n)}\{\mathbf{B}\mathbf{X}_i \mathbf{B}^H\}\right\},$ and $\mathcal{G}(\mathbf{B}) = \sum_{i=1}^{N_m} \text{tr}\left\{(\mathbf{B}\mathbf{X}_i \mathbf{B}^H)^H d\{\text{Bdiag}_{(n)}\{\mathbf{B}\mathbf{X}_i \mathbf{B}^H\}\}\right\}.$

After using the properties P₂, P₃, P₄, P_{4'}, P₇, P₉, P₈, P₁₂ $\mathcal{F}(\mathbf{B})$ and $\mathcal{G}(\mathbf{B})$ in Eq. (9) and using the property P₁₃ we obtain $\frac{\partial \mathcal{C}_{ZBD}(\mathbf{B})}{\partial \mathbf{B}}$ and $\frac{\partial \mathcal{C}_{ZBD}(\mathbf{B})}{\partial \mathbf{B}^*}$ respectively equal as follows:

$$\begin{aligned} \sum_{i=1}^{N_m} (\text{Bdiag}_{(n)}\{\mathbf{B}\mathbf{X}_i \mathbf{B}^H\})^T \mathbf{B}^* \mathbf{X}_i^* + \sum_{i=1}^{N_m} (\text{Bdiag}_{(n)}\{\mathbf{B}\mathbf{X}_i \mathbf{B}^H\})^* \mathbf{B}^* \mathbf{X}_i^T, \\ \sum_{i=1}^{N_m} (\text{Bdiag}_{(n)}\{\mathbf{B}\mathbf{X}_i \mathbf{B}^H\}) \mathbf{B}\mathbf{X}_i^H + \sum_{i=1}^{N_m} (\text{Bdiag}_{(n)}\{\mathbf{B}\mathbf{X}_i \mathbf{B}^H\})^H \mathbf{B}\mathbf{X}_i. \end{aligned}$$

It finally leads to the result stated in Section 3.

B. Coefficients of the 4th-degree polynomial

Using the properties P₁₄ and P₁₅, the cost function can be expressed as:

$$\begin{aligned} \mathcal{C}_{ZBD}(\mathbf{B} - \mu \mathbf{D}_B^{(m)}) &= \sum_{i=1}^{N_m} \left\| \text{Bdiag}_{(n)}\left\{ \mu^2 \left(\mathbf{D}_B^{(m)} \mathbf{X}_i \left(\mathbf{D}_B^{(m)} \right)^H \right) \right. \right. \\ &\quad \left. \left. - \mu \left(\mathbf{B}\mathbf{X}_i \left(\mathbf{D}_B^{(m)} \right)^H + \mathbf{D}_B^{(m)} \mathbf{X}_i \mathbf{B}^H \right) + \mathbf{B}\mathbf{X}_i \mathbf{B}^H \right\} \right\|_F^2. \end{aligned} \quad (10)$$

From the properties P₁, P₁₄, P₁₅ and P₁₆ when introducing the four following matrices $\mathbf{K}_0$, $\mathbf{K}_1$, $\mathbf{K}_2$ and $\mathbf{K}_3$:

$$\mathbf{K}_0 = \mathbf{D}_B^{(m)} \mathbf{X}_i \left(\mathbf{D}_B^{(m)} \right)^H, \quad (11)$$

$$\mathbf{K}_1 = \mathbf{D}_B^{(m)} \mathbf{X}_i^H \mathbf{B}^H + \mathbf{B} \mathbf{X}_i^H \left(\mathbf{D}_B^{(m)} \right)^H, \quad (12)$$

$$\mathbf{K}_2 = \text{Bdiag}_{(n)} \{ \mathbf{B} \mathbf{X}_i \left(\mathbf{D}_B^{(m)} \right)^H + \mathbf{D}_B^{(m)} \mathbf{X}_i \mathbf{B}^H \}, \quad (13)$$

$$\mathbf{K}_3 = \mathbf{B} \mathbf{X}_i \mathbf{B}^H. \quad (14)$$

We finally find that:

$$\begin{aligned} C_{BD} \left(\mathbf{B} - \mu \mathbf{D}_B^{(m)} \right) &= \sum_{i=1}^{N_m} \text{tr} \left\{ (\mathbf{K}_3)^H \text{Bdiag}_{(n)} \{ \mathbf{K}_3 \} \right\} \\ &- \mu \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{K}_1 \text{Bdiag}_{(n)} \{ \mathbf{K}_3 \} + (\mathbf{K}_3)^H \mathbf{K}_2 \right\} \\ &+ \mu^2 \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{K}_0^H \text{Bdiag}_{(n)} \{ \mathbf{K}_3 \} + (\mathbf{K}_3)^H \text{Bdiag}_{(n)} \{ \mathbf{K}_0 \} + \mathbf{K}_1 \mathbf{K}_2 \right\} \\ &- \mu^3 \sum_{i=1}^{N_m} \text{tr} \left\{ (\mathbf{K}_0)^H \mathbf{K}_2 + \mathbf{K}_1 \text{Bdiag}_{(n)} \{ \mathbf{K}_0 \} \right\} \\ &+ \mu^4 \sum_{i=1}^{N_m} \text{tr} \left\{ (\mathbf{K}_0)^H \text{Bdiag}_{(n)} \{ \mathbf{K}_0 \} \right\} \end{aligned} \quad (15)$$

It finally leads to the results stated in Section 3.2.

6. REFERENCES

- [1] J.-F. Cardoso and A. Souloumiac, "Blind beamforming for non-gaussian signals," *IEEE Proc.-F*, vol. 40, no. 6, pp. 362–370, 1993.
- [2] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique using second order statistics," *IEEE Trans. Signal Proc.*, vol. 45, pp. 434–444, Feb. 1997.
- [3] P. Comon, "Independent component analysis, a new concept?," *Signal Proc.*, vol. 36, pp. 287–314, 1994.
- [4] E. Moreau, "A generalization of joint-diagonalization criteria for source separation," *IEEE Trans. Signal Proc.*, vol. 49, no. 3, pp. 530–541, March 2001.
- [5] A. Yeredor, A. Ziehe, and K.-R. Müller, "Approximate joint diagonalization using natural gradient approach," in *Lecture Notes in Computer Science, in Proceedings ICA*, Granada, Espagne, May 2004, pp. 89–96.
- [6] A. Souloumiac, "Nonorthogonal joint diagonalization by combining givens and hyperbolic rotations," *IEEE Trans. Signal Proc.*, vol. 57, no. 6, pp. 2222–2231, June 2009.
- [7] G. Chabriel and J. Barrère, "A direct algorithm for nonorthogonal approximate joint diagonalization," *IEEE Trans. Signal Proc.*, vol. 60, no. 1, pp. 39–47, 2012.
- [8] L. De Lathauwer, C. Févotte, B. De Moor, and J. Vandewalle, "Jacobi algorithm for block diagonalization in blind identification," in *Proc. Symp. Inform. Theory in the Benelux*, Louvain-la-Neuve, Belgium, May 2002, pp. 155–162.
- [9] H. Ghennioui, N. Thirion-Moreau, E. Moreau, and D. Aboutajdine, "Gradient-based joint block diagonalization algorithms: Application to blind separation of fir convolutive mixtures," *Signal Proc.*, vol. 90, no. 6, pp. 1836–1849, June 2010.
- [10] D. Nion, "A tensor framework for non-unitary joint block diagonalization," *IEEE Trans. Signal Proc.*, vol. 59, no. 10, pp. 4585–4594, Oct. 2011.
- [11] P. Tichavsky and Z. Koldovsky, "Algorithms for nonorthogonal approximate joint block-diagonalization," in *Proceedings EUSIPCO*, Bucharest, Romania, August 2012, pp. 2094–2098.
- [12] O. Cherrak, H. Ghennioui, E.-H. Abarkan, and N. Thirion-Moreau, "Non-unitary joint block diagonalization of matrices using a levenberg-marquardt algorithm," in *Signal Processing Conference (EUSIPCO), 2013 Proceedings of the 21st European*, Sept 2013, pp. 1–5.
- [13] A. Belouchrani, K. Abed-Meraim, M. G. Amin, and A. Zoubir, "Joint anti-diagonalization for blind source separation," in *Proc. IEEE ICASSP*, Salt Lake City, UT, May 2001, pp. 2196–2199.
- [14] E. M. Fadaili, N. Thirion-Moreau, and E. Moreau, "Non-orthogonal joint diagonalization/zero-diagonalization for source separation based on time-frequency distributions," *IEEE Trans. Signal Proc.*, vol. 55, no. 5, May 2007.
- [15] G. Chabriel, J. Barrère, N. Thirion-Moreau, and E. Moreau, "Algebraic joint zero-diagonalization and blind source separation," *IEEE Trans. Signal Proc.*, vol. 56, no. 3, pp. 980–989, March 2008.
- [16] W. T. Zhang and S. T. Lou, "A low complexity iterative algorithm for joint zero-diagonalization," *IEEE Signal Proc. Lett.*, vol. 19, no. 2, pp. 115–118, January 2012.
- [17] E. M. Fadaili, N. Thirion-Moreau, and E. Moreau, "Algorithme de zéro-diagonalisation conjointe pour la sparation de sources dterministes," in *20me Colloque GRETSI*, Louvain-La-Neuve, Belgique, Septembre 2005.
- [18] E. M. Fadaili, N. Thirion-Moreau, and E. Moreau, "Combined non-orthogonal joint zero-diagonalization and joint diagonalization for source separation," in *IEEE Workshop on Statistical Signal Processing (SSP'2005)*, Bordeaux, France, Juillet 2005.
- [19] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *Journal of research of the National Bureau of Standards*, vol. 49, pp. 409–436, 1952.
- [20] N. Gong, W. Shao, and H. Xu, "The conjugate gradient method with neural network control," in *Intelligent Systems and Knowledge Engineering (ISKE), 2010 International Conference on*, Nov 2010, pp. 82–84.
- [21] A. Boonpoonga, S. Buritramart, P. Sirisuk, M. Krairiksh, and T.K. Sarkar, "Adaptive array processing unit based on direct data domain least square approach using conjugate gradient method," in *Microwave Conference, 2007. APMC 2007. Asia-Pacific*, Dec 2007, pp. 1–4.
- [22] L. Zhang, "Conjugate gradient approach to blind separation of temporally correlated signals," in *Communications, Circuits and Systems, 2004. ICCAS 2004. 2004 International Conference on*, June 2004, vol. 2, pp. 1008–1012.
- [23] E. Polak, *Optimization algorithms and consistent approximations*, Springer, 1997.
- [24] W. W. Hager and H. Zhang, "A survey of nonlinear conjugate gradient methods," *Pacific J. Optimiz.*, vol. 2, pp. 35–58, 2006.
- [25] Y. H. Dai and Y. Yuan, "A nonlinear conjugate gradient method with a strong global convergence property," *SIAM J. Optim.*, vol. 10, pp. 177–182, 1999.

NON-UNITARY JOINT BLOCK DIAGONALIZATION OF MATRICES USING A LEVENBERG-MARQUARDT ALGORITHM

O. Cherrak, H. Ghennioui, E.-H. Abarkan

Université Sidi Mohamed Ben Abdellah
FSTF, LSSC
Faculté des Sciences et Techniques de Fès,
Route Immouzzar, B.P. 2202, Fès-Maroc
{cherrakomar, ghennioui}@gmail.com

N. Thirion-Moreau

Aix-Marseille Université, CNRS,
ENSAM, LSIS, UMR 7296,
13397 Marseille-France
Université de Toulon, CNRS, LSIS,
UMR 7296, 83957 La Garde-France
thirion@univ-tln.fr

ABSTRACT

This communication addresses the problem of the Non-Unitary Joint Block Diagonalization (NU – JBD) of a given set of complex matrices. This problem occurs in various fields of applications, among which is the blind separation of convolutive mixtures of sources. We present a new method for the NU – JBD based on the Levenberg-Marquardt algorithm (LMA). Our algorithm uses a numerical diagram of optimization which requires the calculation of the complex Hessian matrices. The main advantages of the proposed method stem from the LMA properties: it is powerful, stable and more robust. Computer simulations are provided in order to illustrate the good behavior of the proposed method in different contexts. Two cases are studied: in the first scenario, a set of exactly block-diagonal matrices are considered, then these matrices are progressively perturbed by an additive gaussian noise. Finally, this new NU – JBD algorithm is compared to others put forward in the literature: one based on an optimal step-size relative gradient-descent algorithm [1] and one based on a nonlinear conjugate gradient algorithm [2]. This comparison emphasizes the good behavior of the proposed method.

Index Terms— Joint block-diagonalization algorithms, complex Hessian Matrices, Levenberg-Marquardt algorithm.

1. INTRODUCTION

In the recent years, the problem of the joint decomposition of matrix (or tensor) sets has often arisen in the signal processing field, especially in blind source separation and array processing applications. One of the first considered problem was the Joint Diagonalization (JD) of a given matrix set under the unitary constraint, leading to the nowadays well-known JADE (Joint Approximate Diagonalization of Eigenmatrices) [3] and SOBI (Second Order Blind Identification) [4] algorithms. The following works have addressed either the problem of the JD of tensors [5][6] or the problem of JD of ma-

trices but discarding the unitary constraint [7][8][9][10]. This first particular type of matrix decomposition is useful both in sources localization and direction finding problems and in blind sources separation of instantaneous mixtures.

A second type of matrix decompositions, namely joint zero-diagonalization, has proven to be useful in blind source separation, telecommunication [11] and cryptography [12]. The first suggested algorithms operated under the unitary constraint [13], since they were applied after a classical pre-whitening stage. But such a preliminary pre-whitening step establishes a bound with regard to the best reachable performances in the context of BSS that is the reason why the unitary constraint was soon discarded, leading to several solutions.

A third type of matrix decompositions, namely the joint block-diagonalization, is encountered both in the wide-band sources localization in the presence of a correlated noise and in the blind separation of convolutive mixtures (or multi-dimensional deconvolution) problems. Several algorithms have been developed, under different assumptions about the considered matrix set (the matrices can be either positive definite or hermitian) and about the block-diagonalizer (it is assumed unitary [14][15] or not [1][2][16] [17]).

In this communication, we present a new approach for the NU – JBD based on a Levenberg-Marquardt optimization scheme. It requires that quantities such as the Hessian matrices are calculated. The most important advantage of this approach is that it is general (the considered matrices are complex and not necessarily hermitian and the block-diagonalizer is not necessarily an unitary matrix) and robust.

2. PROBLEM STATEMENT

2.1. Non-unitary joint block-diagonalization problem

The problem of the non-unitary joint block diagonalization is stated in the following way [1][16]. We consider a set $\mathcal{M}$ of

N_m ($N_m \in \mathbb{N}^*$) square matrices $\mathbf{M}_i \in \mathbb{C}^{M \times M}$, for all $i \in \{1, \dots, N_m\}$ which all admit the following decomposition,

$$\mathbf{M}_i = \mathbf{A} \mathbf{D}_i \mathbf{A}^H, \quad (1)$$

where $(\cdot)^H$ stands for the transpose conjugate operator and

$$\text{the matrices } \mathbf{D}_i = \begin{pmatrix} \mathbf{D}_{i,11} & \mathbf{0}_{12} & \dots & \mathbf{0}_{1r} \\ \mathbf{0}_{21} & \mathbf{D}_{i,22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}_{r-1r} \\ \mathbf{0}_{r1} & \dots & \mathbf{0}_{rr-1} & \mathbf{D}_{i,rr} \end{pmatrix}, \text{ for}$$

all $i \in \{1, \dots, N_m\}$ are $(N \times N)$ block diagonal matrices with r the number of considered blocks ($r \in \mathbb{N}^*$), $\mathbf{D}_{i,jj}$, $i \in \{1, \dots, N_m\}$, $j \in \{1, \dots, r\}$ are $(n_j \times n_j)$ square matrices so that $n_1 + \dots + n_r = N$ where $\mathbf{0}_{ij}$ denotes the $(n_i \times n_j)$ null matrix. $\mathbf{A}$ is a $M \times N$ ($M \geq N$) full rank matrix and the $N \times M$ matrix $\mathbf{B}$ is its pseudo-inverse (or generalized Moore-Penrose inverse). The set of the N_m square matrices $\mathbf{D}_i \in \mathbb{C}^{N \times N}$ is denoted by $\mathcal{D}$. The block sizes n_j for all $j = 1, \dots, r$ are assumed known.

The NU - JBD problem consists of estimating the matrix $\mathbf{A}$ and the block-diagonal matrix set $\mathcal{D}$ from only the matrix set $\mathcal{M}$. Several cases were explored for matrix $\mathbf{A}$ and $\mathbf{M}_i$: the case of a unitary matrix $\mathbf{A}$ has been considered in [15] where a first solution is developed. More recently, other solutions have been suggested in [1][2][16][18], for a non-unitary matrix $\mathbf{A}$ and under different assumptions about the matrix set $\mathcal{M}$. To tackle that problem, we propose, here, to consider the following cost function like in [1][16] (in [2] another cost function is used). The reason is that it enables to simplify the calculations:

$$\mathcal{C}_{BD}(\mathbf{B}) = \sum_{i=1}^{N_m} \|\text{OffBdiag}_{(n)}\{\mathbf{B} \mathbf{M}_i \mathbf{B}^H\}\|_F^2, \quad (2)$$

where the matrix operator $\text{OffBdiag}_{(n)}\{\cdot\}$ is defined in the following way:

$$\text{OffBdiag}_{(n)}\{\mathbf{M}\} = \begin{pmatrix} \mathbf{0}_{11} & \mathbf{M}_{12} & \dots & \mathbf{M}_{1r} \\ \mathbf{M}_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{M}_{r1} & \mathbf{M}_{r2} & \dots & \mathbf{0}_{rr} \end{pmatrix}. \quad (3)$$

where $\mathbf{M}$ is a $N \times N$ square matrix whose components $\mathbf{M}_{ij}$ for all $i, j = 1, \dots, r$ are $n_j \times n_j$ matrices (and $n_1 + \dots + n_r = N$). We denote by $\mathbf{n} = (n_1, n_2, \dots, n_r)$. From now on, the notation $\text{OffBdiag}_{(n)}\{\cdot\}$ is simplified into $\text{OffBdiag}\{\cdot\}$.

3. A NEW JOINT BLOCK-DIAGONALIZATION ALGORITHM

The cost function given in Eq. (2) has to be minimized in order to estimate the matrix $\mathbf{B} \in \mathbb{C}^{N \times M}$. To do this, we suggest a new algorithm based on a Levenberg-Marquardt approach [19]. In the literature, the Levenberg-Marquardt optimization method is well-known for its robustness and its effectiveness. It has been widely used in various fields (e.g. neural networks [20], array processing, blind sources separation and so on).

3.1. Principle

We propose, here, to use a Levenberg-Marquardt optimization scheme [19] to minimize the cost function given by Eq. (2) in order to estimate the block-diagonalizer matrix $\mathbf{B}$. It means

that $\mathbf{B}$ is re-estimated at each iteration m . From now on, it is denoted by $\mathbf{B}^{(m)}$. This matrix can be stored in a vector $\mathbf{b}^{(m)}$ which is obtained by $\mathbf{b}^{(m)} = \text{vec}(\mathbf{B}^{(m)})$, where $\text{vec}(\cdot)$ is the vec-operator (applied on a matrix $\mathbf{B}^{(m)} \in \mathbb{C}^{M \times N}$ it stacks its columns into a column vector belonging to $\mathbb{C}^{(M \times N) \times 1}$). This vector is, thus, updated according to the following adaptation rule for all $m = 1, 2, \dots$,

$$\mathbf{b}^{(m)} = \mathbf{b}^{(m-1)} - \mu [\mathbf{H}^{(m-1)} + \lambda \mathbf{I}_{M^2}]^{-1} \mathbf{g}^{(m-1)}, \quad (4)$$

where $[\cdot]^{-1}$ denotes the inverse of a matrix, λ is positive a small damping factor. μ is small positive factor called the step-size, in our case $\mu = 1$, $\mathbf{I}_{M^2}$ is the $M^2 \times M^2$ identity matrix, $\mathbf{H}$ is the Hessian matrix of $\mathcal{C}_{BD}(\mathbf{B})$. The vector $\mathbf{g}$ is defined as $\mathbf{g} = \text{vec}(\nabla_a \mathcal{C}_{BD}(\mathbf{B}^{(m-1)}))$ where $\nabla_a \mathcal{C}_{BD}(\mathbf{B})$ stands for the complex (absolute) gradient matrix. It was previously calculated in [1] and found to be equal to:

$$\begin{aligned} \nabla_a \mathcal{C}_{BD}(\mathbf{B}) &= 2 \sum_{i=1}^{N_m} \left(\text{OffBdiag}\{\mathbf{B} \mathbf{M}_i \mathbf{B}^H\} \right) \mathbf{B} \mathbf{M}_i^H \\ &+ 2 \sum_{i=1}^{N_m} \left(\text{OffBdiag}\{\mathbf{B} \mathbf{M}_i \mathbf{B}^H\} \right)^H \mathbf{B} \mathbf{M}_i. \end{aligned} \quad (5)$$

If $(\cdot)^T$ stands for the transpose operator and $(\cdot)^*$ for the conjugate operator, we now have to calculate the four complex Hessian matrices that are defined in [21] as:

$$\mathbf{H}_{\mathbf{B}, \mathbf{B}^*}(\mathcal{C}_{BD}(\mathbf{B})) = \frac{\mathbf{A}_{0,0} + \mathbf{A}_{1,1}^T}{2} = [\mathbf{H}_{\mathbf{B}^*, \mathbf{B}}(\mathcal{C}_{BD}(\mathbf{B}))]^T,$$

$$\mathbf{H}_{\mathbf{B}^*, \mathbf{B}^*}(\mathcal{C}_{BD}(\mathbf{B})) = \frac{\mathbf{A}_{0,1} + \mathbf{A}_{0,1}^T}{2}, \mathbf{H}_{\mathbf{B}, \mathbf{B}}(\mathcal{C}_{BD}(\mathbf{B})) = \frac{\mathbf{A}_{1,0} + \mathbf{A}_{1,0}^T}{2}, \quad (6)$$

To that aim, we have to derive the expression of $d^2 \mathcal{C}_{BD}(\mathbf{B})$ which is written as:

$$d^2 \mathcal{C}_{BD}(\mathbf{B}) = [\text{dvec}^T(\mathbf{B}^*), \text{dvec}^T(\mathbf{B})] \begin{bmatrix} \mathbf{A}_{0,0} & \mathbf{A}_{0,1} \\ \mathbf{A}_{1,0} & \mathbf{A}_{1,1} \end{bmatrix} \begin{bmatrix} \text{dvec}(\mathbf{B}) \\ \text{dvec}(\mathbf{B}^*) \end{bmatrix}. \quad (7)$$

The four complex matrices $\mathbf{A}_{0,0}$, $\mathbf{A}_{0,1}$, $\mathbf{A}_{1,0}$ and $\mathbf{A}_{1,1}$ involved in this expression are shown to be equal to (see the Appendix for the detailed calculations),

$$\begin{aligned} \mathbf{A}_{0,0} &= (\mathbf{M}_i^H \mathbf{B}^H \otimes \mathbf{I}_N) \mathbf{T}_{\text{Boff}}(\mathbf{B} \mathbf{M}_i \otimes \mathbf{I}_N) \\ &+ (\mathbf{M}_i \mathbf{B}^H \otimes \mathbf{I}_N) \mathbf{T}_{\text{Boff}}(\mathbf{B} \mathbf{M}_i^H \otimes \mathbf{I}_N) \\ &+ (\mathbf{M}_i \otimes (\text{OffBdiag}\{\mathbf{B} \mathbf{M}_i \mathbf{B}^H\})) \\ &+ (\mathbf{M}_i^H \otimes (\text{OffBdiag}\{\mathbf{B} \mathbf{M}_i^H \mathbf{B}^H\})). \end{aligned} \quad (8)$$

$$\begin{aligned} \mathbf{A}_{0,1} &= (\mathbf{T}_{\text{vec}})^T \left((\mathbf{I}_N \otimes \mathbf{M}_i^T \mathbf{B}^T) \mathbf{T}_{\text{Boff}}(\mathbf{B}^* \mathbf{M}_i^T \otimes \mathbf{I}_N) \right. \\ &\left. + (\mathbf{I}_N \otimes \mathbf{M}_i^* \mathbf{B}^T) \mathbf{T}_{\text{Boff}}(\mathbf{B}^* \mathbf{M}_i^* \otimes \mathbf{I}_N) \right). \end{aligned} \quad (9)$$

$$\begin{aligned} \mathbf{A}_{1,0} &= (\mathbf{T}_{\text{vec}})^T \left((\mathbf{I}_N \otimes \mathbf{M}_i \mathbf{B}^H) \mathbf{T}_{\text{Boff}}(\mathbf{B} \mathbf{M}_i \otimes \mathbf{I}_N) \right. \\ &\left. + (\mathbf{I}_N \otimes \mathbf{M}_i^H \mathbf{B}^H) \mathbf{T}_{\text{Boff}}(\mathbf{B} \mathbf{M}_i^H \otimes \mathbf{I}_N) \right). \end{aligned} \quad (10)$$

$$\begin{aligned} \mathbf{A}_{1,1} &= (\mathbf{M}_i^* \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{\text{Boff}}(\mathbf{B}^* \mathbf{M}_i^T \otimes \mathbf{I}_N) \\ &+ (\mathbf{M}_i^T \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{\text{Boff}}(\mathbf{B}^* \mathbf{M}_i^* \otimes \mathbf{I}_N) \\ &+ (\mathbf{M}_i^T \otimes (\text{OffBdiag}\{\mathbf{B} \mathbf{M}_i \mathbf{B}^H\})) \\ &+ (\mathbf{M}_i^* \otimes (\text{OffBdiag}\{\mathbf{B} \mathbf{M}_i^H \mathbf{B}^H\})). \end{aligned} \quad (11)$$

where the operator $\otimes$ denotes the Kronecker product, $\mathbf{T}_{\text{vec}}$ is a permutation square matrix of size $(MN \times MN)$ and

$\mathbf{T}_{\text{Boff}} = \mathbf{I}_{N^2} - \mathbf{T}_{\text{Diag}}$, is the $N^2 \times N^2$ “transformation” matrix, with $\mathbf{T}_{\text{Diag}} = \text{diag}\{\text{vec}(\text{BDiag}\{\mathbf{1}_N\})\}$, $\mathbf{1}_N$ is the $N \times N$ matrix whose components are all ones, $\mathbf{I}_{N^2} = \text{Diag}\{\mathbf{1}_{N^2}\}$ is the $N^2 \times N^2$ identity matrix, and $\text{Diag}\{\mathbf{a}\}$ is a square diagonal matrix which contains the elements of the vector $\mathbf{a}$.

3.2. Summary

The principle of the proposed non-unitary joint block diagonalization algorithm (denoted JBD_{LM}) is summarized in the following table,

Denote the N_m square matrices as $\mathbf{M}_1, \dots, \mathbf{M}_{N_m}$. Given initial estimate $\mathbf{B}^{(0)}$. Given initial estimate for damping value $\lambda^{(0)}$. For $m = 1, 2, \dots$ - Compute $\nabla_a \mathcal{C}_{BD}(\mathbf{B})$ whose expression is given by Eq. (5). - Compute the four complex Hessian matrices whose expressions are given by Eq. (6). - Compute the error $\mathbf{e}^{(m)}$. - Compute the updated solution whose expression is given by Eq. (4). - Evaluate the error at the updated parameter $\mathbf{e}_{lm}$. If $\mathbf{e}^{(m)} \geq \mathbf{e}_{lm}$, $\lambda^{(m)} = \frac{\lambda^{(m-1)}}{10}$, $\mathbf{e}^{(m)} = \mathbf{e}_{lm}$. Else, $\lambda^{(m)} = 10\lambda^{(m-1)}$. EndIf - Stop after a fixed number of iterations or when $\mathbf{B}^{(m)} - \mathbf{B}^{(m-1)} \leq \varepsilon$ where ε is a small positive threshold. EndFor

4. COMPUTER SIMULATIONS

We present simulations to illustrate the effectiveness of the suggested algorithm. We consider a set $\mathcal{D}$ of $N_m = 2, 10$ and 100 matrices, randomly chosen (according to a Gaussian law) of mean 0 and variance 1. Initially these matrices are exactly block-diagonal, then a random noise matrix of mean 0 and variance σ_b^2 is added. The signal to noise ratio is defined as $\text{SNR} = 10 \log(\frac{1}{\sigma_b^2})$. To measure the quality of the estimation, the ensuing performance index (which is an extension of the one introduced in [6]) is used,

$$I_{\text{conv}}(\mathbf{G}) = \frac{1}{r(r-1)} \left[\sum_{i=1}^r \left(\sum_{j=1}^r \frac{\|(\mathbf{G})_{i,j}\|_F^2}{\max_{\ell} \|(\mathbf{G})_{i,\ell}\|_F^2} - 1 \right) + \sum_{j=1}^r \left(\sum_{i=1}^r \frac{\|(\mathbf{G})_{i,j}\|_F^2}{\max_{\ell} \|(\mathbf{G})_{\ell,j}\|_F^2} - 1 \right) \right],$$

where $(\mathbf{G})_{i,j}$ for all $i, j \in \{1, \dots, r\}$ is the (i, j) -th block matrix of $\mathbf{G} = \hat{\mathbf{B}}\mathbf{A}$. The better results are obtained when the index performance $I_{\text{conv}}(\cdot)$ is found to be close to 0 in linear scale ($-\infty$ in logarithmic scale). All the displayed results have been averaged over 100 Monte-Carlo trials.

First, we compare the results (in terms of performance index) obtained thanks to the three algorithms JBD_{ORG} (Optimal Relative Descent-Gradient) [1], JBD_{CG} (Non-linear Conjugate Gradient) [2] and the proposed algorithm JBD_{LM} versus the number of iterations for number of blocks $r = 2$ for different numbers of matrices ($N_m = 2, N_m = 10$ and $N_m = 100$). The obtained results are displayed in the Fig.1 in a noiseless context (SNR = 100 dB) whereas, in the Fig.2 the same study is performed in a noisy context (SNR = 40 dB). All these methods were initialized using the generalized eigenvalue decomposition [2].

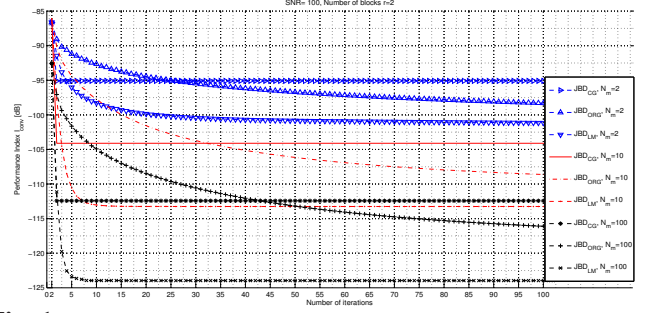


Fig. 1. Performance index versus number of iterations for number of matrices : $N_m = 2, N_m = 10$ and $N_m = 100$, number of blocks $r = 2$ and SNR = 100 dB.

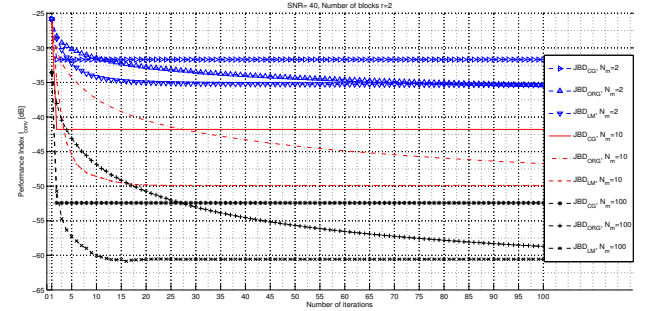


Fig. 2. Performance index versus number of iterations for number of matrices : $N_m = 2, N_m = 10$ and $N_m = 100$, number of blocks $r = 2$ and SNR = 40 dB.

In the noiseless case, the JBD_{LM} algorithm outperforms the two other algorithms (JBD_{CG}, JBD_{ORG}) for $N_m = 2$ (resp. 10 and 100), -101 dB instead of -96 dB and -98 dB (resp. -113 dB instead of -104 dB and -108 dB and -124 dB instead of -112 dB and -113 dB). In the noisy context (SNR = 40 dB) for different number of $N_m = 2$ (resp. 10 and 100), JBD_{LM} outperforms JBD_{CG}, JBD_{ORG}. JBD_{LM} and JBD_{ORG} reach nearly the same performances which is not surprising since these two algorithms minimize the same cost function. -35 dB is reached instead of -31 dB for JBD_{CG} (resp. -50 dB (JBD_{LM}) instead of -42 (JBD_{CG}) dB and -47 dB (JBD_{ORG}) and -61 dB instead of -53 dB and -58 dB).

Then, we show the influence of the size N_m of the matrix set to be block-diagonalized. We display the evolution of the performance index versus the number of matrices N_m for the number of blocks $r = 2$ and 4 in a noiseless context in the Fig.3 and in the noisy context in the Fig.4.

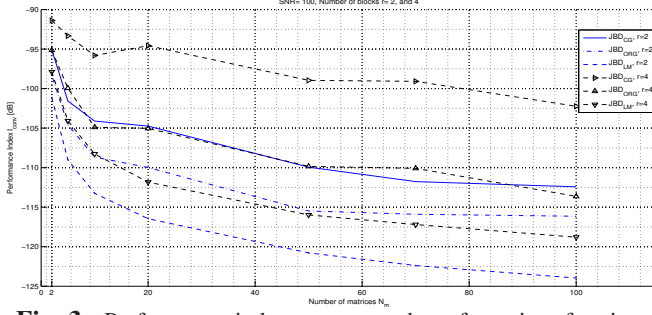


Fig. 3. Performance index versus number of matrices for size of blocks: $r = 2$, $r = 4$ and SNR = 100 dB.

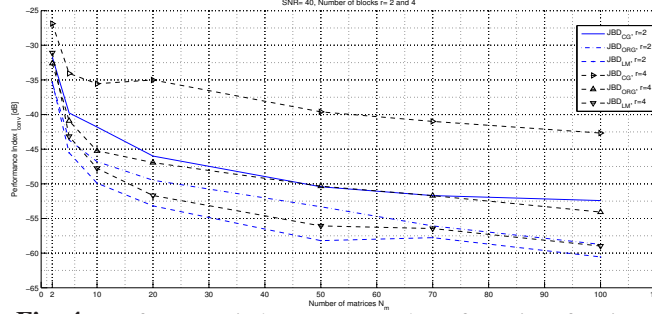


Fig. 4. Performance index versus number of matrices for size of blocks: $r = 2$, $r = 4$ and SNR = 40 dB.

These charts illustrate the good behavior of the proposed algorithm. We observe a decrease in the performances when smaller subsets of matrices are used, even though the results remain relatively good.

5. CONCLUSION

In this communication, we have suggested a new NU – JBD algorithm called JBD_{LM}. It relies on a Levenberg-Marquardt optimization scheme which requires the calculation of the complex Hessian matrices. The main advantage of the JBD_{LM} is that it offers the best performances in complicated situations (noisy cases, and/or few matrices to be joint block-diagonalized). Extensions for future research would be to demonstrate the interest of this algorithm in a true blind multi-dimensional deconvolution context (blind separation of convolutive mixtures of sources).

APPENDIX

Considering four square matrices $\mathbf{D}_1$, $\mathbf{D}_2$, $\mathbf{D}_3$ of size $M \times M$ and $\mathbf{D}_4$ of size $N \times N$. $\text{tr}\{\cdot\}$ denotes the trace operator. Our developments are based on the following properties [21][22],

- P₁. $\|\text{OffBdiag}\{\mathbf{D}_1\}\|_F^2 = \text{tr}\{\mathbf{D}_1^H \text{OffBdiag}\{\mathbf{D}_1\}\}$.
- P₂. $\text{tr}\{\mathbf{D}_1\} = \text{tr}\{\mathbf{D}_1^T\}$.
- P₃. $\text{tr}\{\mathbf{D}_1 + \mathbf{D}_2\} = \text{tr}\{\mathbf{D}_1\} + \text{tr}\{\mathbf{D}_2\}$.
- P₄. $\text{tr}\{\mathbf{D}_1 \mathbf{D}_2 \mathbf{D}_3\} = \text{tr}\{\mathbf{D}_3 \mathbf{D}_1 \mathbf{D}_2\} = \text{tr}\{\mathbf{D}_2 \mathbf{D}_3 \mathbf{D}_1\}$.
- P₅. $\text{tr}\{\mathbf{D}_1^H \mathbf{D}_2\} = (\text{vec}(\mathbf{D}_1))^H \text{vec}(\mathbf{D}_2)$.
- P₆. $d(\text{tr}\{\mathbf{D}_1\}) = \text{tr}\{d(\mathbf{D}_1)\}$.
- P₇. $\text{vec}(\mathbf{D}_1 + \mathbf{D}_2) = \text{vec}(\mathbf{D}_1) + \text{vec}(\mathbf{D}_2)$.

- P₈. $\text{vec}(\mathbf{D}_1 \mathbf{D}_2 \mathbf{D}_3) = (\mathbf{D}_1^H \otimes \mathbf{D}_1) \text{vec}(\mathbf{D}_2)$.
- P₉. $\text{vec}(\mathbf{D}_1^T) = \mathbf{T}_{\text{vec}} \text{vec}(\mathbf{D}_1)$.
- P₁₀. $\text{vec}(\text{OffBdiag}\{\mathbf{D}_1\}) = \mathbf{T}_{\text{Boff}} \text{vec}(\mathbf{D}_1)$.
- P₁₁. $d(\text{vec}(\mathbf{D}_1)) = \text{vec}(d(\mathbf{D}_1))$.
- P₁₂. $(\mathbf{D}_1 \otimes \mathbf{D}_2)^T = \mathbf{D}_1^T \otimes \mathbf{D}_2^T$.
- P₁₃. $(\mathbf{D}_1 \otimes \mathbf{D}_2)^H = \mathbf{D}_1^H \otimes \mathbf{D}_2^H$.
- P₁₄. $(\mathbf{D}_1 \otimes \mathbf{D}_2)(\mathbf{D}_3 \otimes \mathbf{D}_4) = (\mathbf{D}_1 \mathbf{D}_3 \otimes \mathbf{D}_2 \mathbf{D}_4)$.

The second-order derivative of the cost function (Eq. (2)) is defined like in [21],

$$d^2 \mathcal{C}_{BD}(\mathbf{B}) = d\mathcal{D}_B(\mathcal{C}_{BD}(\mathbf{B})) d\text{vec}(\mathbf{B}) + d\mathcal{D}_B^*(\mathcal{C}_{BD}(\mathbf{B})) d\text{vec}(\mathbf{B}^*), \quad (12)$$

where $\mathcal{D}_B(\mathcal{C}_{BD}(\mathbf{B}))$ and $\mathcal{D}_B^*(\mathcal{C}_{BD}(\mathbf{B}^*))$ are the Jacobian matrices of the $\mathcal{C}_{BD}(\mathbf{B})$, defined like in [21],

$$\begin{aligned} \mathcal{D}_B(\mathcal{C}_{BD}(\mathbf{B})) &= \text{vec}^T \left(\frac{\partial \mathcal{C}_{BD}(\mathbf{B})}{\partial \mathbf{B}} \right), \\ \mathcal{D}_B^*(\mathcal{C}_{BD}(\mathbf{B})) &= \text{vec}^T \left(\frac{\partial \mathcal{C}_{BD}(\mathbf{B})}{\partial \mathbf{B}^*} \right). \end{aligned} \quad (13)$$

The partial derivatives $\frac{\partial \mathcal{C}_{BD}(\mathbf{B})}{\partial \mathbf{B}}$ and $\frac{\partial \mathcal{C}_{BD}(\mathbf{B})}{\partial \mathbf{B}^*}$ of the cost function $\mathcal{C}_{BD}(\mathbf{B})$ (2) are calculated in [1],

$$\begin{aligned} \frac{\partial \mathcal{C}_{BD}(\mathbf{B})}{\partial \mathbf{B}} &= \sum_{i=1}^{N_m} \left(\text{OffBdiag}_{(n)}\{\mathbf{B}\mathbf{M}_i \mathbf{B}^H\} \right)^T \mathbf{B}^* \mathbf{M}_i^* \\ &+ \sum_{i=1}^{N_m} \left(\text{OffBdiag}\{\mathbf{B}\mathbf{M}_i \mathbf{B}^H\} \right)^* \mathbf{B}^* \mathbf{M}_i^T, \end{aligned} \quad (14)$$

$$\begin{aligned} \frac{\partial \mathcal{C}_{BD}(\mathbf{B})}{\partial \mathbf{B}^*} &= \sum_{i=1}^{N_m} \left(\text{OffBdiag}\{\mathbf{B}\mathbf{M}_i \mathbf{B}^H\} \right) \mathbf{B}\mathbf{M}_i^H \\ &+ \sum_{i=1}^{N_m} \left(\text{OffBdiag}\{\mathbf{B}\mathbf{M}_i \mathbf{B}^H\} \right)^H \mathbf{B}\mathbf{M}_i. \end{aligned} \quad (15)$$

Using the previous properties (P₁, P₂, ..., P₁₄) and the expression of the partial derivatives $\frac{\partial \mathcal{C}_{BD}(\mathbf{B})}{\partial \mathbf{B}}$ (14) and $\frac{\partial \mathcal{C}_{BD}(\mathbf{B})}{\partial \mathbf{B}^*}$ (15), the derivate of the Jacobian matrices can be expressed as,

$$\begin{aligned} d\mathcal{D}_B(\mathcal{C}_{BD}(\mathbf{B})) &= \sum_{i=1}^{N_m} \text{vec}^T(d\mathbf{B}) \left[(\mathbf{T}_{\text{vec}})^T (\mathbf{I}_N \otimes \mathbf{M}_i \mathbf{B}^H) \mathbf{T}_{\text{Boff}} (\mathbf{B}\mathbf{M}_i \otimes \mathbf{I}_N) \right. \\ &+ (\mathbf{T}_{\text{vec}})^T (\mathbf{I}_N \otimes \mathbf{M}_i^H \mathbf{B}^H) \mathbf{T}_{\text{Boff}} (\mathbf{B}\mathbf{M}_i^H \otimes \mathbf{I}_N) \\ &+ \text{vec}^T(d\mathbf{B}^*) \left[(\mathbf{M}_i^H \mathbf{B}^H \otimes \mathbf{I}_N) \mathbf{T}_{\text{Boff}} (\mathbf{B}\mathbf{M}_i \otimes \mathbf{I}_N) \right. \\ &+ (\mathbf{M}_i \mathbf{B}^H \otimes \mathbf{I}_N) \mathbf{T}_{\text{Boff}} (\mathbf{B}\mathbf{M}_i^H \otimes \mathbf{I}_N) \\ &+ (\mathbf{M}_i \otimes (\text{OffBdiag}\{\mathbf{B}\mathbf{M}_i \mathbf{B}^H\})) \\ &\left. + (\mathbf{M}_i^H \otimes (\text{OffBdiag}\{\mathbf{B}\mathbf{M}_i^H \mathbf{B}^H\})) \right], \end{aligned} \quad (16)$$

$$\begin{aligned} d\mathcal{D}_B^*(\mathcal{C}_{BD}(\mathbf{B})) &= \sum_{i=1}^{N_m} \text{vec}^T(d\mathbf{B}^*) \left[(\mathbf{T}_{\text{vec}})^T (\mathbf{I}_N \otimes \mathbf{M}_i^T \mathbf{B}^T) \mathbf{T}_{\text{Boff}} (\mathbf{B}^* \mathbf{M}_i^* \otimes \mathbf{I}_N) \right. \\ &+ (\mathbf{T}_{\text{vec}})^T (\mathbf{I}_N \otimes \mathbf{M}_i^* \mathbf{B}^T) \mathbf{T}_{\text{Boff}} (\mathbf{B}^* \mathbf{M}_i^* \otimes \mathbf{I}_N) \\ &+ \text{vec}^T(d\mathbf{B}) \left[(\mathbf{M}_i^* \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{\text{Boff}} (\mathbf{B}^* \mathbf{M}_i^* \otimes \mathbf{I}_N) \right. \\ &+ (\mathbf{M}_i^T \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{\text{Boff}} (\mathbf{B}^* \mathbf{M}_i^* \otimes \mathbf{I}_N) \\ &+ (\mathbf{M}_i^T \otimes (\text{OffBdiag}\{\mathbf{B}\mathbf{M}_i \mathbf{B}^H\})) \\ &\left. + (\mathbf{M}_i^* \otimes (\text{OffBdiag}\{\mathbf{B}\mathbf{M}_i^H \mathbf{B}^H\}))^* \right]. \end{aligned} \quad (17)$$

So from Eq. (7) and Eq. (12), it leads to the results stated by Eqs. (8, ..., 11). Finally, we can express the four complex Hessian matrices in the form of Eq. (6).

6. REFERENCES

- [1] H. Ghennioui, N. Thirion-Moreau, E. Moreau, and D. Aboutajdine, "Gradient-based joint block diagonalization algorithms: Application to blind separation of fir convolutive mixtures," *Signal Proc.*, vol. 90, no. 6, pp. 1836–1849, June 2010.
- [2] D. Nion, "A tensor framework for non-unitary joint block diagonalization," *IEEE Trans. Signal Proc.*, vol. 59, no. 10, pp. 4585–4594, Oct. 2011.
- [3] J.-F. Cardoso and A. Souloumiac, "Blind beamforming for non-gaussian signals," *IEEE Proc.-F*, vol. 40, no. 6, pp. 362–370, 1993.
- [4] A. Belouchrani, K. Abed-Meraïm, J.-F. Cardoso, and E. Moulines, "A blind source separation technique using second order statistics," *IEEE Trans. Signal Proc.*, vol. 45, pp. 434–444, Feb. 1997.
- [5] P. Comon, "Independant component analysis, a new concept ?," *Signal Proc.*, vol. 36, pp. 287–314, 1994.
- [6] E. Moreau, "A generalization of joint-diagonalization criteria for source separation," *IEEE Trans. Signal Proc.*, vol. 49, no. 3, pp. 530–541, March 2001.
- [7] A. Yeredor, A. Ziehe, and K.-R. Müller, "Approximate joint diagonalization using natural gradient approach," in *Lecture Notes in Computer Science, in Proceedings ICA*, Granada, Espagne, May 2004, pp. 89–96.
- [8] A. Souloumiac, "Nonorthogonal joint diagonalization by combining givens and hyperbolic rotations," June 2009, vol. 57, pp. 2222–2231.
- [9] X.-L. Li and X. D. Zhang, "Nonorthogonal joint diagonalization free of degenerate solution," *IEEE Trans. Signal Proc.*, vol. 55, no. 5, pp. 1803–1814, May 2007.
- [10] G. Chabriel and J. Barrère, "A direct algorithm for nonorthogonal approximate joint diagonalization," *IEEE Trans. Signal Proc.*, vol. 60, no. 1, pp. 39–47, 2012.
- [11] G. Chabriel and J. Barrère, "Non-symmetrical joint zero-diagonalization and mimo zero-division multiple access," *IEEE Trans. Signal Proc.*, vol. 59, no. 5, pp. 2296–2307, 2011.
- [12] J. Walgate, A. J. Short, L. Hardy, and V. Vedral, "Local distinguishability of multipartite orthogonal quantum states," *Physical Review Letter*, vol. 85, pp. 4972–4975, 2000.
- [13] A. Belouchrani, K. Abed-Meraïm, M. G. Amin, and A. Zoubir, "Joint anti-diagonalization for blind source separation," in *Proc. IEEE ICASSP*, Salt Lake City, UT, May 2001, pp. 2196–2199.
- [14] A. Belouchrani, M. G. Amin, and K. Abed-Meraïm, "Direction finding in correlated noise fields based on joint block-diagonalization of spatio-temporal correlation matrices," *IEEE Signal Lett.*, vol. 4, no. 9, pp. 266–268, Sept. 1997.
- [15] L. De Lathauwer, C. Févotte, B. De Moor, and J. Vandewalle, "Jacobi algorithm for block diagonalization in blind identification," in *Proc. Symp. Inform. Theory in the Benelux*, Louvain-la-Neuve, Belgium, May 2002, pp. 155–162.
- [16] H. Ghennioui, E.-M. Fadaili, N. Thirion-Moreau, A. Adib, and E. Moreau, "A nonunitary joint block diagonalization algorithm for blind separation of convolutive mixtures of sources," *IEEE Signal Lett.*, vol. 14, no. 11, pp. 860–863, November 2007.
- [17] P. Tichavsky and Z. Koldovsky, "Algorithms for nonorthogonal approximate joint block-diagonalization," in *Proceedings EUSIPCO*, Bucharest, Romania, August 2012, pp. 2094–2098.
- [18] D.-T. Pham, "Blind separation of cyclostationary sources using joint block approximate diagonalization," in *ICA, Springer-Verlag Berlin Heidelberg LNCS 4666, Mike E. Davies et al. (Eds)*, Sept. 2007, pp. 244–251.
- [19] D. G. Luenberger, *Optimization by Vector Space Methods*, John Wiley, New York, 1969.
- [20] B. M. Wilamowski and H. Yu, "Improved computation for levenberg-marquardt training," *IEEE trans. on Neural Networks*, vol. 21, no. 6, pp. 25–29, June 2010.
- [21] A. Hjørungnes, *Complex-Valued Matrix Derivatives With Applications in Signal Processing and Communications*, Cambridge, University Press, February 2011.
- [22] M. Joho and K. Rahbar, "Joint diagonalization of correlation matrices by using newton methods with application to blind signal separation," in *SAM Workshop. IEEE*, 2002, pp. 403–407.

SEPARATION AVEUGLE DE SOURCES BASEE SUR LA BLOC DIAGONALISATION CONJOINTE NON-UNITAIRE DE MATRICES ISSUES DE TRANSFORMEES SPATIALES QUADRATIQUES : APPLICATION A LA SURVEILLANCE MARITIME

O. Cherrak ^(1, 2), H.Ghennioui ⁽²⁾, N.Thirion-Moreau ⁽¹⁾, E. Abarkan ⁽²⁾

(1) Laboratoire Sciences de l'Information et des Systèmes, LSEET, UMR CNRS 6017, Université du Sud Toulon Var, Avenue G. Pompidou, BP. 56, 83162 La Valette du Var Cédex, France, Email: omar.cherrak@usmba.ac.ma, thirion@univ-tln.fr.

(2) Laboratoire Signaux, Systèmes et Composants, Université Sidi Mohamed Ben Abdellah, B.P. 2202, Faculté des Sciences et Techniques de Fès, Route d'Immuizzer, Fès, Maroc, Email: hicham.ghennioui@usmba.ac.ma, habarkan@yahoo.fr.

Résumé:

Nous considérons le problème de la séparation aveugle de mélanges de sources par bloc-diagonalisation conjointe non unitaire d'un ensemble de matrices issues de transformées temps-fréquence spatiales quadratiques. Nous utilisons un critère de sélection automatique de points temps-fréquence permettant la construction de l'ensemble des matrices à bloc-diagonaliser conjointement. Les principaux avantages de cette méthode sont d'être plus directe en ce sens qu'elle ne requière plus de blanchiment préalable des observations et de pouvoir être appliquée à des signaux corrélés. Puisqu'il est indiqué que l'algorithme de bloc diagonalisation peut aussi résoudre le problème de séparation de sources AIS (Automatic Identification System) des mélanges instantanés dédiée à la surveillance maritime puisque que la diagonalisation conjointe n'est rien d'autre qu'un cas particulier de bloc diagonalisation conjointe. Donc, nous proposons une nouvelle méthode consistant à diagonaliser conjointement l'ensemble des distributions quadratique temps-fréquence spatiales ou des matrices de spectres. A l'aide de simulations informatiques, nous illustrons le comportement de la méthode proposée ainsi que son efficacité et la comparons à une autre méthode basée quant à elle sur un algorithme de bloc-diagonalisation conjointe non-unitaire.

Mots clés: Séparation aveugle de sources, mélanges instantanés de sources AIS, (bloc) diagonalisation conjointe, distributions temps-fréquence, critère de sélection automatique.

1. INTRODUCTION

L'objet de la séparation aveugle de sources (SAS) est de parvenir à ré-estimer signaux source à partir de la seule connaissance de m signaux d'observation correspondant à des mélanges de ces sources. Elle possède plusieurs degrés de difficulté selon les caractéristiques des sources considérées (signaux déterministes ou aléatoires, statistiquement mutuellement indépendants ou corrélés, i.i.d. ou non, stationnaires, cyclo-stationnaires ou non stationnaires...) et également selon le type de système mélangeant envisagé (système linéaire ou non linéaire, mélange instantané ou convolutif, modèle de mélange sur-déterminé ou sous-déterminé). Elle possède de nombreuses applications parmi lesquelles la téléphonie mobile, les télécommunications numériques, le traitement

de la parole, le génie biomédical, la géophysique interne ou externe, le radar, le sonar, etc...

Dans ce papier, nous nous intéressons au problème de la SAS basée sur l'utilisation de :

- représentations temps-fréquence spatiales quadratiques (RTFSQ) dans le contexte des mélanges linéaires, convolutifs à réponse impulsionnelle finie (RIF) de sources déterministes,
- ou de spectres temps-fréquence spatiales quadratiques (STFSQ) dans le cas des mélanges linéaires convolutifs RIF de sources aléatoires non-stationnaires, éventuellement corrélées.

Le problème de la séparation de mélanges linéaires instantanés de sources au moyen de RTFSQ a déjà fait l'objet de nombreux travaux parmi lesquels [1][3][4]. Les méthodes de SAS basées sur l'utilisation de RTFSQ reposent toutes sur une étape préalable de sélection (automatique) de points temps-fréquence $(t-f)$ particuliers permettant la

construction d'ensembles de matrices devant ensuite être bloc diagonalisées conjointement et/ou zéro-bloc diagonalisées conjointement. Pour que la séparation soit possible, il faut que les RTFSQ des sources diffèrent en un nombre suffisant de points t - f , en d'autres termes que les signatures des sources soient suffisamment différentes dans le plan t - f pour permettre la construction des ensembles de matrices auxquels nous nous intéressons. En contexte instantané, de nombreux travaux sont dédiés au problème de la sélection automatique des points t - f permettant la construction soit d'un seul ensemble [4] (celui des matrices qui devront être diagonalisées conjointement) soit de deux ensembles (ceux des matrices qui devront être diagonalisées conjointement et zéro-diagonalisées conjointement) [1][3]. Certains de ces détecteurs opèrent sur les données blanchies, d'autres directement sur les données non blanchies. En contexte convolutif, un premier détecteur a été proposé dans [2]. En ce qui concerne la mise au point d'algorithmes de bloc-diagonalisation conjointe, cette problématique a suscité un certain nombre de travaux parmi lesquels [5][6]. Ces algorithmes ont ensuite été appliqués au problème de la SAS basée sur l'utilisation de RTFSQ en mélange convolutif dans [1][2].

Dans ce papier, nous utilisons un détecteur automatique de points t - f pour la construction de l'ensemble des matrices devant être diagonaliser conjointement [4]. Nous proposons également une nouvelle méthode pour la séparation aveugle de mélanges instantanés de sources AIS de surveillance maritime basée sur l'utilisation de spectres t - f spatiaux quadratiques : elle sera baptisée DC_{LM-TF} . Son intérêt principal est de ne plus requérir de blanchiment préalable des observations puisqu'elle combine ce détecteur automatique de points t - f à l'algorithme de bloc diagonalisation conjointe non unitaire à base de l'algorithme de Levenberg-Marquardt proposé dans [8]. Enfin, au moyen des simulations informatiques, nous montrons l'intérêt de cette méthode en la comparant à une autre approche existante dans la littérature [9].

2. SÉPARATION AVEUGLE DE SOURCES

2.1 Formulation du problème

On considère le système multi-entrées multi-sorties RIF à temps discret décrit par l'équation suivante,

$$x_i(t) = \sum_{j=1}^n \sum_{l=0}^L h_{ij}(l) s_j(t-l) + n_j(t), \quad \forall i = 1, \dots, m, \quad (1)$$

où $s_j(t)$ représente la j -ème n sources, $x_i(t)$ est la i -ème des $m \geq n$ observations, $h_{ij}(t)$ est la réponse impulsionnelle réelle entre la j -ème source et le i -ème capteur. L'ordre du filtre à RIF $\mathbf{H}(t)$ considéré est supposé être égale à L et les $n_i(t)$ pour tout $i = 1, \dots, m$ sont des bruits blancs Gaussiens stationnaires, centrés, mutuellement indépendants et indépendants des sources. Il est

classique de réécrire ce modèle convolutif RIF sous la forme d'un modèle instantané de type $\mathbf{X}(t) = \mathbf{A}\mathbf{S}(t) + \mathbf{N}(t)$. Le vecteur $\mathbf{S}(t)$ de taille $n(L+L') \times 1$ et les vecteurs $\mathbf{X}(t)$ et $\mathbf{N}(t)$ tous deux de taille $mL' \times 1$ sont alors définis de la manière suivante :

$$\begin{aligned} \mathbf{S}(t) &= [s_1(t), \dots, s_1(t-(L+L')+1), \dots, s_n(t-(L+L'+1))]^T, \\ \mathbf{X}(t) &= [x_1(t), \dots, x_1(t-L'+1), \dots, x_m(t-L'+1)]^T, \\ \mathbf{N}(t) &= [n_1(t), \dots, n_1(t-L'+1), \dots, n_m(t-L'+1)]^T, \end{aligned} \quad (2)$$

avec $(.)^T$ opérateur de transposition et L' le nombre de retards considérés (suffisamment grand, i.e., $mL' \geq nL'$ pour que le modèle considéré soit sur-déterminé). La matrice $\mathbf{A}$ de dimension $mL' \times n(L+L')$ s'écrit,

$$\begin{pmatrix} \mathbf{A}_{11} & \dots & \mathbf{A}_{1n} \\ \vdots & \ddots & \vdots \\ \mathbf{A}_{m1} & \dots & \mathbf{A}_{mn} \end{pmatrix}. \quad (3)$$

Elle est constituée de blocs $\mathbf{A}_{ij}$ de dimension $L \times (L+L')$, possédant alors la structure particulière suivante,

$$\mathbf{A}_{ij} = \begin{pmatrix} h_{ij}(0) & \dots & h_{ij}(L) & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \dots & h_{ij}(0) & 0 & \dots & h_{ij}(L) \end{pmatrix}. \quad (4)$$

Notez que lorsque $L = 0, L' = 1$, et $\mathbf{A} = \mathbf{H}$, nous retrouvons le modèle instantané qui peut être réécrit sous la forme matricielle $\mathbf{X}(t) = \mathbf{A}\mathbf{S}(t) + \mathbf{N}(t)$, où la $(M \times N)$ matrice de mélange $\mathbf{A}$ est une matrice définie comme $\mathbf{A} = (\mathbf{A}_{i,j})$ pour tout $i = 1, \dots, m$ et $j = 1, \dots, n$.

2.2 Principe des méthodes basées sur les STFSQ ou les RTFSQ

Le spectre temps-fréquence spatial quadratique d'un vecteur aléatoire complexe $\mathbf{X}(t)$ est définie par,

$$\mathbf{D}_{\mathbf{X}}(t, \nu) = \int_{\mathbb{R}^2} \mathbf{R}_{\mathbf{X}}(\theta, \theta') K(\theta, \theta'; t, \nu) d\theta d\theta', \quad (5)$$

où la fonction $K(\theta, \theta'; t, \nu)$ désigne le noyau de la transformation et $\mathbf{R}_{\mathbf{X}}(\theta, \theta')$ représente la matrice de corrélation de $\mathbf{X}(t)$. Elle est définie par,

$$\mathbf{R}_{\mathbf{X}}(\theta, \theta') = \mathbf{E} \left\{ \mathbf{X} \left(\theta + \frac{\theta'}{2} \right) \mathbf{X}^H \left(\theta - \frac{\theta'}{2} \right) \right\}, \quad (6)$$

où $\mathbf{E}\{\cdot\}$ désigne l'espérance mathématique et $(.)^H$ l'opérateur de transposition conjuguaison. Dans le cas de signaux déterministes, cela devient,

$$\mathbf{R}_{\mathbf{X}}(\theta, \theta') = \mathbf{X} \left(\theta + \frac{\theta'}{2} \right) \mathbf{X}^H \left(\theta - \frac{\theta'}{2} \right). \quad (7)$$

Notons que le Spectre de Pseudo Wigner Spatial (SPWS) que nous utiliserons pour les simulations informatiques (Section 4) car il permet la réduction des interférences entre sources. Il est défini par,

$$\mathbf{D}_{\text{SPWS}, \mathbf{X}}(t, \nu) = \int_{\mathbb{R}} \mathbf{R}_{\mathbf{X}}(t, \tau) h(\tau) e^{-2i\pi\nu\tau} d\tau, \quad (8)$$

$i^2 = -1$ et $h(\tau)$ est une fenêtre de lissage. Notons qu'en considérant $h(\tau) = 1$, quel que soit τ , on retrouve le Spectre de Wigner Spatial (SWS) tandis qu'en enlevant l'espérance mathématique on retrouve la distribution de Pseudo-Wigner Spatiale (PWS) ou celle de Wigner Spatiale (WS) si en plus $h(\tau) = 1$ pour tout τ . Le STFSQ des observations noté $\mathbf{D}_{\mathbf{XX}}(t, \nu)$ de dimension $mL' \times mL'$ vaut alors:

$$\mathbf{D}_{\mathbf{XX}}(t, \nu) = \mathbf{A}\mathbf{D}_{\mathbf{SS}}(t, \nu)\mathbf{A}^H + \mathbf{D}_{\mathbf{NN}}(t, \nu) + \mathbf{A}\mathbf{D}_{\mathbf{SN}}(t, \nu) + \mathbf{D}_{\mathbf{NS}}(t, \nu)\mathbf{A}^H, \quad (9)$$

où $\mathbf{D}_{\mathbf{SS}}(t, \nu)$ représente STFSQ des sources, $\mathbf{D}_{\mathbf{NN}}(t, \nu)$ celui des bruits tandis que $\mathbf{D}_{\mathbf{SN}}(t, \nu)$ et $\mathbf{D}_{\mathbf{NS}}(t, \nu)$ sont les STFSQ croisés entre sources et bruits. $\mathbf{D}_{\mathbf{SS}}(t, \nu)$ est une matrice carrée de taille $n(L + L') \times n(L + L')$. Les bruits étant indépendants des sources alors le STFSQ des observations devient,

$$\mathbf{D}_{\mathbf{XX}}(t, \nu) = \mathbf{A}\mathbf{D}_{\mathbf{SS}}(t, \nu)\mathbf{A}^H + \mathbf{D}_{\mathbf{NN}}(t, \nu). \quad (10)$$

Remarquons que si l'on considère comme STFSQ celui de Wigner alors, $\mathbf{D}_{\mathbf{NN}}(t, \nu) = \int_{\mathbb{R}} \mathbf{R}_{\mathbf{N}}(\tau) h(\tau) e^{-2i\Pi\nu\tau} d\tau = \mathbf{S}_{\mathbf{NN}}(\nu)$, où $\mathbf{S}_{\mathbf{NN}}(\nu)$ est la matrice inter-spectrale du vecteur des bruits. Pour obtenir une estimation des matrices t - f correspondant à des observations non-bruitées, nous éliminons les points t - f de faible énergie en ne gardant que les points tels que,

$$\min \left(|D_{x_i x_j}(t, \nu)|^2 \right) > \varepsilon, \quad i, j \in \{1, \dots, mL'\} \quad (11)$$

où ε est un réel de faible valeur ($\varepsilon > 0.01$ par exemple) et $D_{x_i x_j}(t, \nu) = (\mathbf{D}_{\mathbf{XX}}(t, \nu))_{i,j}$. On peut alors considérer que,

$$\mathbf{D}_{\mathbf{XX}}(t, \nu) \approx \mathbf{A}\mathbf{D}_{\mathbf{SS}}(t, \nu)\mathbf{A}^H \quad (12)$$

La matrice $\mathbf{D}_{\mathbf{SS}}(t, \nu)$ est composée de blocs carrés de dimensions $(L + L') \times (L + L')$. En des points t - f ne correspondant qu'à des auto-termes sur les sources - et en raison du principe d'interférences, il ne peut y avoir qu'un seul auto-terme, c'est la raison pour laquelle nous pourrions traiter le cas de sources corrélées -, cette matrice est (bloc)-diagonale avec un seul (bloc ou terme) non nul sur la diagonale. C'est de cette structure particulière dont nous allons tirer avantage au niveau du détecteur que nous proposons dans la section suivante.

3. UNE NOUVELLE MÉTHODE

3.1 Détecteur de "bons" points t - f

Nous utilisons le détecteur proposé dans [4] noté C_{Ins} pour le contexte instantané considéré. L'idée est de trouver des points t - f "utiles". Pour ce faire, une décomposition en valeurs singulières peut être utilisée (DVS) conduisant à $(\mathbf{D}_{\mathbf{X}}(t, \nu) = (\mathbf{U}(t, \nu)\mathbf{\Delta}(t, \nu)\mathbf{V}(t, \nu)^H)$ avec $\mathbf{V}(t, \nu)$ et $\mathbf{U}(t, \nu)$ sont $(M \times M)$ matrices unitaires et $\mathbf{\Delta}(t, \nu) = \text{diag}\{\delta(t, \nu)\}$ est une matrice diagonale dont les éléments diagonaux sont positifs. Notons que $\delta(t, \nu) = (\delta_1(t, \nu), \dots, \delta_M(t, \nu))^T$ et en supposant que les valeurs singulières sont triées dans un ordre décroissant $\delta_1(t, \nu) \geq \delta_2(t, \nu) \geq \dots \geq \delta_M(t, \nu) \geq 0$. Ensuite, un moyen de vérifier si une matrice de rang un est donnée par,

$$\begin{cases} \mathcal{R}(t, \nu) = \delta_1(t, \nu) > \varepsilon_1, \\ \sum_{i=2}^M \delta_i(t, \nu) < \varepsilon_2, \end{cases} \quad (13)$$

où ε_1 (resp. ε_2) est une constante positive (suffisamment) grande (resp. petite). La

généralisation de ce détecteur [4] en contexte convolutif pour des signaux non-blanchis peut être vue dans [7] dont le rang est égale à $L + L'$.

Les matrices de l'ensemble $\mathcal{M}$ que l'on construit (et que l'on supposera de dimension N_m avec $N_m \in \mathbb{N}^*$) ont alors une structure particulière puisqu'elles se factorisent toutes en $\mathbf{D}_{\mathbf{XX}}(t, \nu) = \mathbf{A}\mathbf{D}_{\mathbf{SS}}(t, \nu)\mathbf{A}^H$ avec $\mathbf{D}_{\mathbf{SS}}(t, \nu)$ matrice diagonale possédant un seul terme non nul sur la diagonale.

3.2 Bloc diagonalization conjointe non-unitaire

Considérons l'ensemble $\mathcal{M}$ construit au moyen du détecteur qui vient d'être présenté,

$$\mathcal{M} = \{\mathbf{M}_i = \mathbf{D}_{\mathbf{XX}}(t_i, \nu_i) \in \mathbb{C}^{mL' \times mL'}, \forall i, j = 1, \dots, N_m\}.$$

Les matrices de l'ensemble $\mathcal{M}$ admettent toutes la factorisation suivante,

$$\mathbf{M}_i = \mathbf{A}\mathbf{D}_i\mathbf{A}^H \text{ ou } \mathbf{D}_i = \mathbf{B}\mathbf{M}_i\mathbf{B}^H, \forall i, j \in \{1, \dots, N_m\}, \quad (14)$$

$$\mathbf{D}_i = \begin{pmatrix} \mathbf{D}_{i,11} & \mathbf{0}_{L+L'} & \dots & \mathbf{0}_{L+L'} \\ \mathbf{0}_{L+L'} & \mathbf{D}_{i,22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}_{L+L'} \\ \mathbf{0}_{L+L'} & \dots & \mathbf{0}_{L+L'} & \mathbf{D}_{i,nn} \end{pmatrix} \forall i \in \{1, \dots, N_m\} \text{ sont}$$

des matrices bloc-diagonales de dimension $n(L + L') \times n(L + L')$ et les matrices $\mathbf{D}_{i,jj}$, $i \in \{1, \dots, N_m\}, j \in \{1, \dots, n\}$ sont des matrices carrées de dimension $(L + L') \times (L + L')$ et $\mathbf{0}_{L+L'}$ désigne la matrice carrée, nulle, de dimension $(L + L') \times (L + L')$. La matrice $\mathbf{A}$ est une matrice de rang plein de dimension $mL' \times n(L + L')$ et $\mathbf{B}$ est sa pseudo-inverse (ou inverse généralisée de Moore-Penrose). Le problème de la bloc diagonalisation conjointe non-unitaire est alors d'estimer la matrice de séparation $\mathbf{B}$ uniquement à partir des matrices de l'ensemble $\mathcal{M}$. Nous proposons donc de considérer le critère suivant,

$$C_{BD}(\mathbf{B}) = \sum_{i=1}^{N_m} \|\text{OffBdiag}_{L+L'}\{\mathbf{B}\mathbf{M}_i\mathbf{B}^H\}\|_F^2, \quad (15)$$

où $\|\cdot\|_F$ désigne la norme de Frobenius et l'opérateur $\text{OffBdiag}_{L+L'}\{\cdot\}$ la matrice zéro bloc-diagonale construite à partir des éléments hors bloc-diagonal de son argument (ses éléments bloc-diagonaux sont donc des matrices $\mathbf{0}_{L+L'}$). Pour trouver l'estimée $\hat{\mathbf{B}}$, il faut donc minimiser la somme des carrés des normes des blocs hors diagonale de $\mathbf{B}\mathbf{M}_i\mathbf{B}^H$. Ceci est réalisé comme dans [8] à savoir au moyen d'une solution algébrique (notée NU - JBD_{LM}) basée sur le calcul des matrices hessiennes complexes exactes. La combinaison de cet algorithme à la sélection automatique de points t - f précédemment décrite conduit à un algorithme baptisé DC_{LM-TF} . Notons que lorsque $\mathbf{D}_{i,jj}$, $i \in \{1, \dots, N_m\}, j \in \{1, \dots, n\}$ ne sont plus de matrices mais des scalaires ($n_j = 1, \forall j \in \{1, \dots, r\} \Rightarrow r = N$), nous retrouvons le problème de diagonalisation conjointe non unitaire [7].

4. SIMULATIONS & CONCLUSION

Nous présentons des simulations informatiques afin d'illustrer le comportement de la méthode proposée DC_{LM-TF} ainsi que son efficacité et la

comparons avec le même genre d'approches existantes, celui dénommé par $DC_{CGNION-TF}$ proposé dans [9] avec les points t - f du détecteur C_{Ins} . Nous considérons $m = 3$ mélanges de $n = 2$ trames de 256 bits correspondent à deux navires avec des caractéristiques différentes. Les trames sont générées conformément à la recommandation AIS comme mentionné dans [10]. Par conséquent, les différentes trames sont composées d'une séquence d'apprentissage (TS) composé de 0 et 1 et qui prend 24 bits. Le fanion de début/Fin qui prend 8 bits. Une séquence de contrôle de redondance cyclique (CRC) d'une longueur de 16 bits est ajoutée aux différents types d'information de 168 bits propres aux navires qui appartiennent à la zone de navigation. Ces trames sont codées en NRZI et modulées en GMSK avec le produit temps bande spécifié dans le système AIS est $BT_s = 0.4$ avec $T_s = \frac{1}{9600}$ s et l'ordre du filtre gaussien est $OF = 21$. Nous choisissons pour la première source (resp. la deuxième source) en tant que paramètres d'entrée, le retard causé par la distance entre le navire et le satellite est de 5 ms et l'effet Doppler provoqué par la vitesse du satellite est de 10 Hz (resp. le retard est 0 ms et le décalage Doppler est de 0 Hz), $L = 0$, $L' = 1$ et ces sources correspondent à 1208 échantillon de temps.

Les sources sont mélangées selon une matrice de mélange $\mathbf{A}(t)$ dont les composantes représentent,

$$\mathbf{A} = \begin{pmatrix} 0.2082 & 0.1450 \\ -0.8666 & -0.1220 \\ -0.7612 & 0.8398 \end{pmatrix}. \quad (16)$$

Nous utilisons le spectre de Pseudo Wigner-Ville avec une fenêtre de Hamming de taille 31 points en temps et nous considérons 121 fréquences. Pour le détecteur C_{Ins} nous avons choisi $\varepsilon_1 = 6$ et $\varepsilon_2 = 0.3$. Afin de pouvoir juger de la qualité de la séparation, nous utilisons l'indice de performance suivant:

$$I(\mathbf{G}) = \frac{1}{r(r-1)} \left[\sum_{i=1}^r \left(\sum_{j=1}^r \frac{\|(\mathbf{G})_{i,j}\|_F^2}{\max_l \|(\mathbf{G})_{i,l}\|_F^2} - 1 \right) + \sum_{j=1}^r \left(\sum_{i=1}^r \frac{\|(\mathbf{G})_{i,j}\|_F^2}{\max_l \|(\mathbf{G})_{l,j}\|_F^2} - 1 \right) \right], \quad (17)$$

où $(\mathbf{G})_{i,j}$, $\forall i, j \in \{1, \dots, r\}$ est la (i, j) -ème matrice bloc (carrée) de $\mathbf{G} = \mathbf{B}\mathbf{A}$. Quand la séparation est parfaite, l'indice de performance doit tendre vers $-\infty$ en échelle logarithmique (0 en échelle linéaire). Le rapport signal à bruit (RSB) est alors défini en dB comme $RSB = 10 \log \left(\frac{1}{\sigma_b^2} \right)$. I en dB est représenté au niveau de la Fig. (1) en fonction du RSB (≈ 150 matrices temps-fréquence ont été sélectionnées pour $RSB \geq 20$ dB. Pour $RSB < 20$ dB, plus le RSB diminue, plus ce nombre de matrices diminue aussi). Ces résultats ont été obtenus par la méthode de Monte-Carlo sur 30 réalisations. La Fig. (1) montre que le DC_{LM-TF} montre les meilleures performances même si l'environnement bruité est considéré. Cette méthode donne les meilleurs résultats en la comparant à $DC_{CGNION-TF}$.

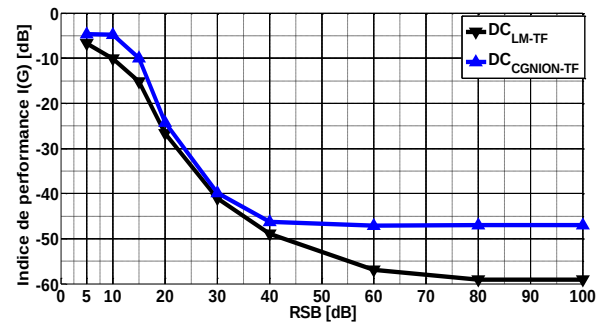


Figure 1: Comportement de l'algorithme DC_{LM-TF} : $I(\mathbf{G})$ en dB en fonction du RSB.

RÉFÉRENCES

- [1] K. Abed-Meraïm, A. Belouchrani et R. Leyman, "Time-frequency signal analysis and processing : a comprehensive reference", *chapter in "Blind source separation using time-frequency distributions"*, Prentice-Hall, B. Boashash Editor, Oxford, UK, January 2003.
- [2] C. Févotte et C. Doncarli, "A unified presentation of blind source separation methods for convolutive mixtures using block-diagonalization", in *Proc. Independent Component Analysis (ICA)*, pp. 349-354, Nara, Japan, April 2003.
- [3] E.-M. Fadaili, N. Thirion-Moreau et E. Moreau, "Non orthogonal joint diagonalization/ zero-diagonalization for source separation based on time-frequency distributions", *IEEE Trans. On Signal Processing*, Vol. 55, No. 5, pp. 1673-1687, May 2007.
- [4] L. Giulieri, H. Ghennioui, N. Thirion-Moreau et E. Moreau, "Non-orthogonal joint diagonalization of spatial quadratic time-frequency matrices for source separation", *IEEE Signal Processing Letter*, Vol. 12, No. 5, pp. 415-418, May 2005.
- [5] H. Bousbiah-Salah, A. Belouchrani et K. Abed Meraïm, "Blind separation of non stationary sources using joint block diagonalization", in *Proc. IEEE Workshop on Statistical Signal Proc.*, pp. 448-451, 2001.
- [6] L. De Lathauwer, C. Févotte, B. De Moor et J. Vandewalle, "Jacobi algorithm for joint block diagonalization in blind identification", in *Proc. 23rd Symposium on Information Theory in the Benelux*, Louvain-la-Neuve, Belgium, May 2002.
- [7] H. Ghennioui, E.M. Fadaili, N. Thirion-Moreau, A. Adib et E. Moreau, "Une nouvelle méthode sans blanchiment pour la séparation aveugle de mélanges convolutifs de sources basée sur l'utilisation de distributions temps-fréquence spatiales quadratiques", pp. 465-468, Colloque GRETSI, septembre 2007, Troyes.
- [8] O. Cherrak, H. Ghennioui, E.H. Abarkan, N. Thirion-Moreau, "Non-Unitary Joint Block Diagonalization of matrices using a Levenberg-Marquardt algorithm", *European Signal Processing Conference EUSIPCO'2013*, pp. 1-5, September 2013, Marrakech.
- [9] D. Nion, "A Tensor Framework for Nonunitary Joint Block Diagonalization", *IEEE Trans. Signal Proc.*, vol. 59, no. 10, pp. 4585-4594, Oct. 2011.
- [10] R. Prévost, M. Coulon, D. Bonacci, J. LeMaitre, J.-P. Millerioux and J.-Y. Tournet, "Interference mitigation and error correction method for AIS signals received by satellite", in *Proc. European Signal Processing Conference (EUSIPCO'12)*, Bucharest, Romania, pp. 46-50, August 2012.

Joint Maximum Likelihood Sequence Estimation algorithm for decollision of AIS signals in maritime surveillance context

CHERRAK Omar

Laboratory of Signals, Systems and Components (LSSC)/ Laboratory of Information Sciences and Systems (LISS)
Faculty of Sciences and Technologies of Fez, Sidi Mohamed Ben Abdellah University, Fez, Morocco/Sud-Toulon-Var, University, France
E-mail: cherrakomar@gmail.com

GHENNIQUI Hicham

Laboratory of Signals, Systems and Components (LSSC)
Faculty of Sciences and Technologies of Fez, Sidi Mohamed Ben Abdellah University, Fez, Morocco
E-mail: ghennioui@gmail.com

THIRION-MOREAU Nadège

Laboratory of Information Sciences and Systems (LISS)
Sud-Toulon-Var University, France
E-mail : thirion@univ-tln.fr

ABARKAN El Hossain

Laboratory of Signals, Systems and Components (LSSC)
Faculty of Sciences and Technologies of Fez, Sidi Mohamed Ben Abdellah University, Fez, Morocco
E-mail : habarkan@yahoo.fr

Abstract—This paper addresses the Joint Maximum-Likelihood Sequence Estimation (JMLSE) algorithm of AIS (Automatic Identification System) signals decollision in maritime surveillance context. The AIS is a Self-Organized Time Division Multiple Access (SO-TDMA) system, in which vessels periodically transmit information. The main idea of this paper is the decollision of signals received by a Low Earth Orbit (LEO) satellite lead to a global surveillance of maritime traffic using JMLSE algorithm. Simulations are provided in order to evaluate the behavior of the JMLSE algorithm for separation and estimation of sources.

IndexTerms— AIS signal, Maritime surveillance, GMSK modulation, Joint Maximum Likelihood Sequence Estimation, (De)Collision, Blind Source Separation.

I. INTRODUCTION

The Moroccan activities of economic order insured by sea have experienced a highly significant increase, proportionately far greater than that of the world economic development: Maritime transport, Marina, etc. This change imposes an enhanced presence of the state to ensure and enhance the safety and the security at sea. It is necessary to adopt a strategy, which emphasizes the monitoring in depth and reactivity. Whenever possible, an action at the nearest source of the threats or dangers is sought to detect the precursors. Maritime protection consists to develop and make available monitoring devices and intervention, in order to guard against certain dangers, to maintain the optimum performance for the control of ships and to prevent or treat a wide range of existing threats at the sea or from the sea which are susceptible undermine our national territory, our interests and the national or international public order. It includes, of course, the maritime defense of the territory. This paper is a part of this problematic.

The major objective of this paper is to develop the JMLSE algorithm for blind source separation of linear mixtures intended specifically for maritime surveillance of the AIS system by a LEO satellite. To that aim, we provide a GMSK modulated signal within the technical characteristics of AIS system (Data encapsulation, NRZI line coding, the I/Q GMSK modulator). Finally, we interest in separation and decollision of the AIS signals (Correct reception of messages from different vessels with following the decollision signals) and this only from observations.

II. TRANSMISSION STEP

A. AIS Message

The AIS message is a length of $N = 256$ bits. It contains a training sequence (TS) consisting 0 and 1 which takes 24 bits. The start flag (SF) and the end flag (EF) for information takes 8 bits. A control sequence (FCS) (16 bits Cyclic Redundancy Code (CRC)) is added to the data information (168 bits). A zero inserted after every five continuous 1 in the information data and FCS as illustrated in Figure 1.

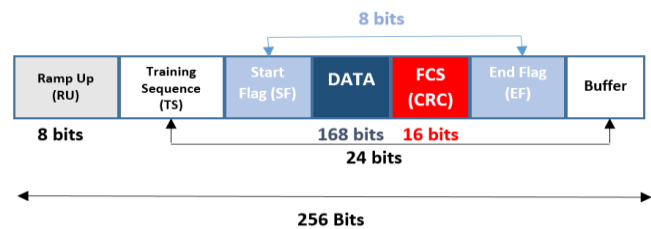


Figure 1: AIS message.

The binary sequence $\{a_n\}_{0 \leq n \leq N}$ of the message AIS takes the values $\{-1, +1\}$ since the line coding constitutes a fundamental block in the AIS system. The NRZI encoding used as $a_n = b_n a_{n-1}, n \geq 0$ where $b_n \in \{-1, 1\}$. The modulation specified by AIS standard is Gaussian Minimum

Shift Keying (GMSK) [1]. The encoded message is modulated and transmitted at 9600 bps on carriers at frequency 162 MHz.

B. GMSK modulator for AIS system

GMSK is a frequency-shift keying modulation [1], which produces constant-envelope and continuous-phase. Like any digital modulation, the GMSK is decomposed into three main steps,

- Mapping: Matching a set of bits to a set of symbols.
- Shaping: Transforming symbol sequence into a signal to be adapted to the transmission channel. It is usually realized by a linear filtering operation. The signal can be written as,

$$s_g[k] = \sum_{n=-\infty}^{+\infty} a_n g[k - nT_s], \quad (1)$$

where a_n are the emitted symbols, T_s is the symbol period and $g[k]$ represents the shaping filter. In the GMSK case, $g[k]$ the response of a Gaussian filter whose expression is given by,

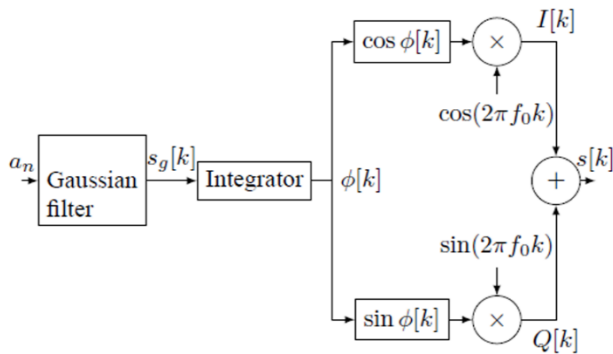
$$g[k] = \sqrt{\frac{2\pi}{\log 2}} B e^{\left(-\frac{2}{\log 2} (\pi B k)^2\right)}, \quad (2)$$

where B is the bandwidth of the Gaussian filter, the GMSK modulation is described by the bandwidth-time product BT , AIS uses $BT = 0.4$ and $T_s = \frac{1}{9600}$ s.

- Modulation: Placing the signal on a frequency f_0 of a carrier. This operation produces a signal of spectral characteristic, which is adapted to the band-pass channel transmission. The GMSK modulated signal is expressed in the following form,

$$\begin{aligned} s[k] &= \text{Re}\{e^{i\phi[k]} e^{i2\pi f_0 k}\} \\ &= I[k] \cos(2\pi f_0 k) - Q[k] \sin(2\pi f_0 k), \end{aligned} \quad (3)$$

where Re is the real part of a complex number, $\phi[k]$ is the instantaneous phase of $s_g[k]$. $I[k]$ (resp. $Q[k]$) modulates the carrier in phase (resp. in phase quadrature). We are talking about I (real) and Q (imaginary) channels. The structure of the GMSK



modulator is illustrated in the Figure 2.

Figure 2: The GMSK modulator scheme.

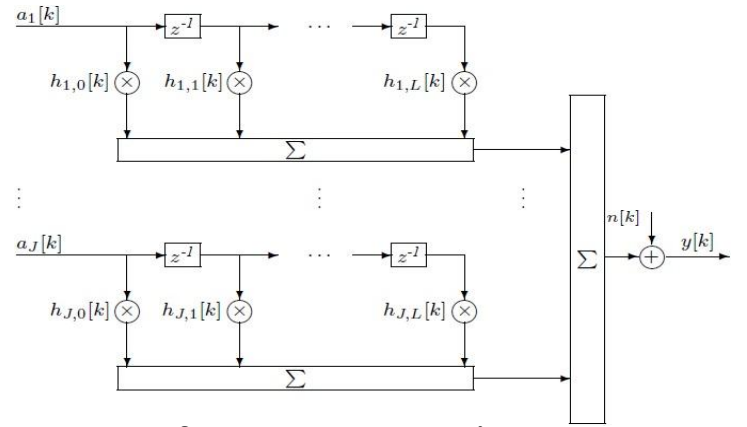
III. RECEPTION STEP

A. Problem statement

The collision of messages caused by the large difference in propagation time from vessels to satellites [2] can be expressed as follows,

$$y[k] = \sum_{j=1}^J \sum_{l=0}^L h_{j,l}[k] s_j[k-l] + n[k], \quad 0 \leq k \leq K-1, \quad (4)$$

where $y[k]$ are the observations received at the satellite, $h_{j,l}[k]$ is the Impulse Response of j channel, $s_j[k]$ is the k -th sample of the j -th co-channel which comes from the GMSK modulation, and these observations are perturbed by an additive gaussian noise $n[k]$, J is the number of co-channel signals, K is the number of symbols per slot and L is the channel length. The structure collision model is presented in Figure 3.



B. Joint Maximum Likelihood Sequence Estimation (JMLSE)

We propose, in this section, to use Joint Multiuser Detection method (JMD) to solve the problem of co-channel user separation. The principle of this algorithm is to estimate jointly the different data sequences. It is based on the Joint Maximum Likelihood Sequence Estimation (JMLSE) [2]. The implementation of this method is based on the Viterbi algorithm [3]. The Viterbi Algorithm can be simply described as a method, which finds the most likely path through a trellis, i.e. shortest path, to separate the AIS signals received by the LEO satellite called observations. Its first application was proposed as a solution to the decoding of convolutional codes [4]. Before that, some assumptions are necessary for the good behavior of the JMLSE algorithm,

- Doppler frequency should not be strong,
- the additive noise is a white Gaussian centered noise,
- the output sequences are independent of the noise,
- to increase the probability estimation, the received signal is re-sampled by a factor, which is equal to four.
- and the state machine of the FIR channel is finite.

C. Description of the JMLSE algorithm

Firstly, to simplify the calculations, we consider the model described by the following expression,

$$y[k] = \sum_{j=1}^J \sum_{l=0}^L h_{j,l}[k] a_j[k-l] + n[k], \quad 0 \leq k \leq K-1 \quad (5)$$

A simple mathematical development of Eq. (5) can reduce the problem to a least square minimization defined by the following equation,

$$(\hat{a}_1, \dots, \hat{a}_J) = \arg \min_{(\hat{a}_1, \dots, \hat{a}_J)} \{\sum_{k=0}^{K-1} \tau[k]\}, \quad (6)$$

where $\hat{a}_j = [\hat{a}_j[0], \dots, \hat{a}_j[K-1]]^T$, $1 \leq j \leq J$, represents the estimated sequence j , $\tilde{a}_j$ is the assumption on the transmitted symbol and $\tau[k]$ represents the branch metric defined by,

$$\tau[k] = |y[k] - \sum_{j=1}^J \sum_{l=0}^L h_{j,l} \tilde{a}_j[k-l]|^2. \quad (7)$$

The interest of JMLSE algorithm is to find the sequence of the lowest metric. The metric of the paths is defined by,

$$\lambda[k+1] = \lambda[k] + \tau[k+1], \quad (8)$$

where $\lambda[k]$ is the metric of the path from the previous step in the lattice.

We propose, here, to use a semi-blind algorithm to estimate the channel. Moreover, this algorithm uses the training sequence to estimate the strongest transmitter. The channel estimation of this transmitter can be performed by the least square method. However, the other channels are initialized as follows,

$$\hat{h}_j = [0, 0, \dots, C, 0, \dots, 0], \text{ for } 2 \leq j \leq J \quad (9)$$

where C is a real constant. The position of this constant depends on the selected channel model. We consider the vector $\tilde{a}_{j,v,\mu}(k+1) = [\tilde{a}_{j,v,\mu}(k+1), \dots, \tilde{a}_{j,v,\mu}(k-L+1)]^T$, for $1 \leq j \leq J$, the transmitted sequences at $k+1$ corresponds to the transition state from v to μ for the transmitter j . To pass from the state v to the state μ , we use the metric function $\tau_{v,\mu}(k+1)$ whose expression is given by the following equation,

$$\tau_{v,\mu}(k+1) = |y[k] - \hat{h}_j^T(k) * \tilde{a}_{j,v,\mu}(k+1)|^2. \quad (10)$$

The first step of the proposed algorithm is to compute the branch metric and then the metric of the paths which is expressed in the following way,

$$\lambda_{v,\mu}[k+1] = \lambda_v[k] + \tau_{v,\mu}[k+1], \quad (11)$$

where $\lambda_v[k+1]$ is obtained by the minimization of $\lambda_{v,\mu}[k+1]$. After that, the sequence is estimated $\hat{a}_j[k-\delta]$, where δ is a decisive constant for performance system. We notice that the vector $\hat{h}_j$ of each channel j at $k+1$ is estimated according to the following optimization scheme,

$$\hat{h}_j[k+1] = \hat{h}_j[k] + \beta[k] e[k] \hat{a}_j[k-\delta], \quad (12)$$

Where $\beta[k]$ is the optimal adaptation step-size defined in LMS algorithm [5] and $e[k] = y[k-\delta] - \sum_{j=0}^J \hat{h}_j^T[k] * a_j[k-\delta]$.

D. Summary

The principle of the JMLSE algorithm summarized in the following table,

Consider the samples of the received signal $y[k]$.
Initialize the channel $\hat{h}_1$ of the strongest transmitter using the least squares method and the other channels initialized by Eq. (9).
Initialize to all states v, μ the assumptions on the transmitted sequence.
For $k = 0, \dots, K-1$
-- Calculate the branch metric whose expression is given by the Eq. (10).
-- Calculate the metric paths metric whose expression is given by the Eq. (11).
-- Determine the best path for each state μ by $\lambda_{\mu}[k+1] = \min\{\lambda_{v,\mu}[k+1]\}$.
-- Trace back to get the estimated $\hat{a}_j[k-\delta]$.
-- Update the coefficients of the channel according to the adaptation rule whose expression is given by Eq. (12).
EndFor

IV. COMPUTER SIMULATIONS

Here, we present simulations to illustrate the effectiveness and the performances of the proposed algorithm. Firstly, we consider $J = 2$ users with instantaneous mixing (no multipath, considered $L = 0$) and $K = 10$. However, we have considered these parameters in order to show the principle of the JMLSE algorithm,

$$\mathbf{a}_1 = [1 \quad -1 \quad 1 \quad 1 \quad 1 \quad -1 \quad 1 \quad -1 \quad -1 \quad -1],$$

$$\mathbf{a}_2 = [1 \quad -1 \quad 1 \quad -1 \quad 1 \quad -1 \quad -1 \quad -1 \quad 1 \quad 1].$$

The channel vectors randomly generated according to a Gaussian distribution,

$$\mathbf{h}_1 = [0.4281, 0.8958, 1.4108, 0.3764, 0.3600, -0.0900, -1.6320, 0.8797, 0.0994, 0.8965],$$

$$\mathbf{h}_2 = [-0.3209, -1.5978, 1.7098, 0.1334, 0.3575, 0.4680, -0.3173, -0.0887, 2.4878, 0.8107].$$

The vector of observations is equal to,

$$\mathbf{y} = [0.1072 \quad 0.7020 \quad 3.1206 \quad 0.2430 \quad 0.7175 \quad -0.3780 \quad -1.3147 \quad -0.7910 \quad 2.3884 \quad -0.0858].$$

The different metric branches are calculated and given in the Tab. (1)

States	(1,1)	(-1,1)	(1,-1)	(-1,-1)
k=0	0	0.7330	0.4119	0.0460
k=1	1.9711	10.2117	3.2099	0
k=2	0	7.9610	11.6931	38.9507
k=3	0.0711	0.2364	0	0.5668

k=4	0	0.5185	0.5111	2.0593
k=5	0.5715	0.8760	0.0324	0
k=6	0.4026	6.9136	0	10.6531
k=7	2.5024	0.0315	3.0951	0
k=8	0.0395	0	22.8182	24.7562
k=9	3.2151	0	0.0295	2.6287

Table 1: Metric branches for two users with instantaneous mixing for the example considered.

The optimal lattice deduced from the Tab. (1) is then presented in the Figure 4. According to the optimal lattice obtained, the sequences are estimated from the observations and found to be the same as transmitted ones.

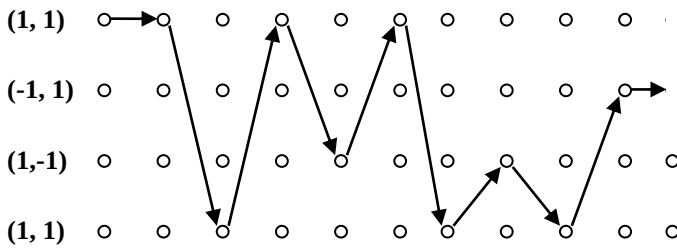


Figure 4: the optimal lattice for two users with instantaneous mixing for the example considered.

The estimated sequences are then,

$$\hat{\mathbf{a}}_1 = [1 \quad -1 \quad 1 \quad 1 \quad 1 \quad -1 \quad 1 \quad -1 \quad -1 \quad -1],$$

$$\hat{\mathbf{a}}_2 = [1 \quad -1 \quad 1 \quad -1 \quad 1 \quad -1 \quad -1 \quad -1 \quad 1 \quad 1].$$

Finally, we consider the following scenario: we provide GMSK baseband signals with a bandwidth-bit-time product parameter $BT = 0.4$, the Sample Rate is $SR = 9600$ bps, the Over Sample Rate is $OSR = 4$ and the Order Gaussian filter is $OF = 21$ within the technical characteristics of an AIS system mentioned in the previous section (the messages are generated according to the AIS recommendation: they are composed of 168 information bits. Then, the Ramp Up, Training Sequence, FCS (CRC), Buffer and the Start/End Flag bits inserted). These signals perturbed by an additive white Gaussian noise. We detect the AIS messages received (called observations) and we separate them using the JMLSE algorithm presented in the section III.

We show the performances obtained when using the JMLSE algorithm proposed for three users, and these sources mixed to create observations. Then, we have to estimate the transmitted sources only from the observations. All displayed results have been average over 10 Monte-Carlo trials.

In the Figure 5, we observe that if we start with a high SNR ($SNR = 0$ dB), the MSE is high. However, when the SNR decreases, the mean square error decreases too for the three sources. We observe also that the MSE equal to zero when $SNR \geq 30$ dB

So all these charts illustrate the good behavior of the proposed algorithm. We observe a good estimation even if the noisy case is considered.

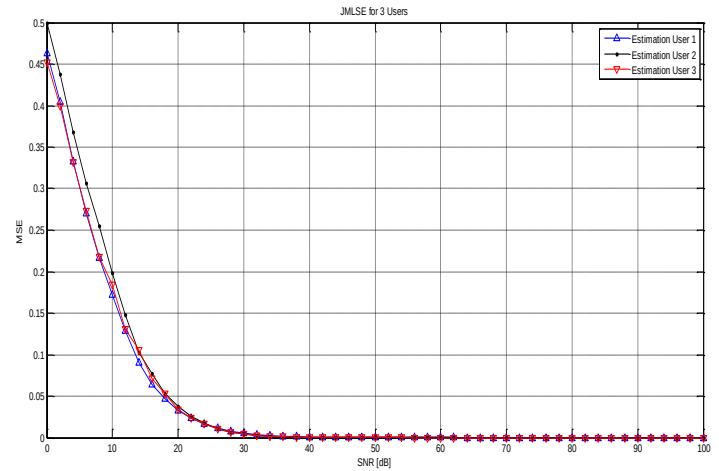


Figure 5: Mean-square-error performance of the JMLSE algorithm versus SNR for $J = 3$ users.

V. DISCUSSION

In this paper, we propose to use the JMLSE algorithm in the order to separate the AIS signals received by a LEO satellite.

The JMLSE algorithm is sensitive to Doppler Effect and the canal estimation. Therefore, it needs a pretreatment in order to estimate the Doppler Effect for each user [6] caused by the speed of the satellite.

VI. CONCLUSION

In this paper, we have presented a new algorithm for decollision of AIS signals. This algorithm uses the Joint Maximum Likelihood Sequence Estimation (JMLSE). Extensions for future research would be to improve the JMLSE algorithm, in order to develop new estimators of Doppler frequency for each user and to propose a new method for the Non-Unitary Joint Block diagonalization that occurs in the blind separation of convolutive mixtures of sources in surveillance maritime context.

VII. REFERENCES

- [1] A. Linz and A. Hendrickson, "Efficient Implementation of an I-Q GMSK Modulator", IEEE Trans. Circuits and Systems-II: Analog and Digital Signal Processing, vol. 43, no.1, January 1996.
- [2] G. D. Forney, "Maximum-Likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference", IEEE Trans. Inform. Theory, 18:363—378, 1972.
- [3] Viterbi, A.J., "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm", IEEE Transactions on Information Theory, April 1967; IT - 13 (2), pp. 260 - 269.
- [4] Viterbi, A.J., "Convolutional Codes and Their Performance in Communication Systems", IEEE Transactions on Communications Technology, October 1971; COM - 19 (5), pp. 751 - 772.
- [5] O. Ali Abu-Ella, B. El-Jabu, "Optimal Robust Adaptive LMS Algorithm without Adaptation Step-Size", GSMM2008 Proceeding, 21-24 April 2008, pp 249 - 251.
- [6] P. Bianchi, P. Loubaton and F. Sirven, "Non data aided estimation of the modulation index of continuous phase modulations", IEEE Trans. Signal Processing, vol. 52, no. 10, pp. 2847-2861, 2004

[7] H. Ghennioui, E.-M. Fadaili, N. Thirion-Moreau, A. Adib and E. Moreau, "A nonunitary joint block diagonalization algorithm for blind separation of convolutive mixtures of sources", in *IEEE Signal Lett.*, vol. 14, no. 11, pp 860-863, November 2007.

Successive Interference Cancellation technique for decollision of AIS signals in maritime surveillance context by a LEO satellite.

Omar Cherrak^{*†}, Hicham Ghennioui^{*}, Nadège Thirion-Moreau[†] and El Hossein Abarkan^{*}

^{*}LSSC, FSTF, B.P. 2202, Route d'Imouzzer, Fès, Maroc

Email: {cherrakomar,ghennioui}@gmail.com

[†]LSIS, UMR CNRS 7296, Université du Sud Toulon Var, F-83957, La Garde Cédex, France

Email: thirion@univ-tln.fr

Abstract—This communication addresses the Successive Interference Cancellation (SIC) technique of decollision of AIS (Automatic Identification System) signals in maritime surveillance context. AIS is a Self-Organized Time Division Multiple Access (SO-TDMA) system, in which vessels periodically transmit information. Researchers are confronted with the increase of maritime traffic, and they seek to increase the security of their maritime zones. The main idea of this communication is decollision between vessels received by a low-orbit satellite lead to a global surveillance of maritime traffic using Serial Interference Cancellation (SIC) technique which is widely studied for other types of co-channel interference applications. These messages are transmitted by vessels using the AIS standards. Simulations are provided in order to evaluate the behavior of the SIC technique for separation and estimation of sources.

Index Terms—AIS signal, GMSK modulation, serial interference cancellation, collision.

I. INTRODUCTION

The Moroccan activities of economic order insured by sea have experienced a highly significant increase, proportionately far greater than that of the world economic development: Maritime transport, Marina, etc. This change imposes an enhanced presence of the state to ensure and enhance the safety and the security at sea. It is necessary to adopt a strategy which emphasizes the monitoring in depth and reactivity. Whenever possible, an action at the nearest source of the threats or dangers is sought to detect the precursors. Maritime protection consists to develop and make available monitoring devices and intervention to guard against certain dangers, to maintain the optimum performance for the control of ships and to prevent or treat a wide range of existing threats at the sea or from the sea which are susceptible undermine our national territory, our interests and the national or international public order. It includes, of course, the maritime defense of the territory. This communication is a part of this problematic.

The major objective of this communication is to develop the SIC technique for blind source separation of linear mixtures intended specifically for maritime surveillance of the AIS system by a LEO satellite (Low Earth Orbit). This work is realized in two steps: the first one provide a GMSK modulated signal within the technical characteristics of AIS system (Data

encapsulation, NRZI line coding, the I/Q GMSK modulator), and the second one is dedicated to the reception and detection of AIS signals (Correct reception of messages from different vessels with following the decollision signals).

II. TRANSMISSION STEP

A. AIS Message

The AIS message is a length of $N = 256$ bits. It contains a training sequence (TS) consisting 0 and 1 which takes 24 bits. The start flag (SF) and the end flag (EF) for information takes 8 bits. A control sequence (FCS) (16 bits Cyclic Redundancy Code (CRC)) is added to the data information (168 bits). A zero is inserted after every five continuous 1 in the information data and FCS as illustrated in Fig. (1).

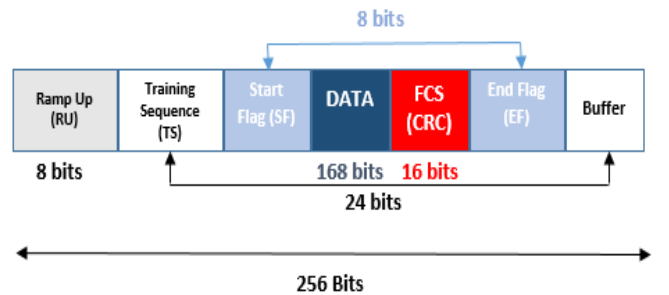


Fig. 1: AIS Message.

The binary sequence $\{a_n\}_{0 \leq n \leq N}$ of the message AIS takes the values $\{-1, +1\}$ since the line coding constitutes a fundamental block in the AIS system. The NRZI encoding is used as $a_n = b_n a_{n-1}, n \geq 0$ where $b_n \in \{-1, 1\}$. The modulation specified by AIS standard is Gaussian Minimum Shift Keying (GMSK) [1]. The encoded message is modulated and transmitted at 9600 bps on carriers at frequency 162MHz. after every five continuous 1s).

B. GMSK modulator for AIS system

GMSK is a frequency-shift keying modulation which produces constant-envelope and continuous-phase. Like any digital modulation, the GMSK is decomposed into three main steps,

- Mapping: Matching a set of bits to a set of symbols.
- Shaping: Transforming symbol sequence into a signal to be adapted to the transmission channel. It is usually realized by a linear filtering operation. The signal can be written as,

$$s_g[k] = \sum_{n=0}^{+\infty} a_n g[k - nT_s], \quad (1)$$

where a_n are the emitted symbols, T_s is the symbol period and $g[k]$ represents the shaping filter. In the GMSK case, $g[k]$ is the response of a Gaussian filter whose expression is given by,

$$g[k] = \sqrt{\frac{2\pi}{\log 2}} B \exp\left(-\frac{2}{\log 2}(\pi B k)^2\right), \quad (2)$$

where B is the bandwidth of the Gaussian filter, the GMSK modulation is described by the bandwidth-time product BT , AIS uses $BT = 0.4$ and $T_s = \frac{1}{9600}s$.

- Modulation: Placing the signal on a frequency f_0 of a carrier. This operation produces a signal of spectral characteristic which is adapted to the band-pass channel transmission. The GMSK modulated signal is expressed in the following form,

$$\begin{aligned} s[k] &= \text{Re}\{\exp(j\phi[k]) \exp(j2\pi f_0 k)\} \\ &= I[k] \cos(2\pi f_0 k) - Q[k] \sin(2\pi f_0 k), \end{aligned} \quad (3)$$

where Re is the real part a complex number, $\phi[k]$ is the instantaneous phase of $s_g[k]$. $I[k]$ (resp. $Q[k]$) modulates the carrier in phase (resp. in phase quadrature). We are talking about I (real) and Q (imaginary) channels. The structure of the GMSK modulator is illustrate in the Fig. (2),

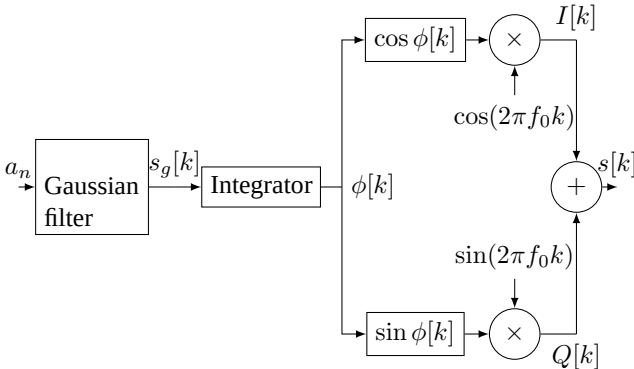


Fig. 2: GMSK modulator scheme.

III. RECEPTION STEP

A. Problem statement

The collision of messages caused by the large difference in propagation time from vessels to satellites [2] can be expressed

as follows,

$$y[k] = \sum_{j=1}^J \sum_{l=0}^L h_{j,l}[k] s_j[k-l] + n[k], 0 \leq k \leq K-1, \quad (4)$$

where $y[k]$ are the observations received at the satellite, $h_{j,l}[k]$ is the Impulse Response of j channel, $s_j[k]$ is the k -th sample of the j -th co-channel which comes from the GMSK modulation, and these observations are perturbed by an additive gaussian noise $n[k]$, J is the number of co-channel signals, K is the number of symbols per slot and L is the channel length. The structure collision model is presented in Fig. (3),

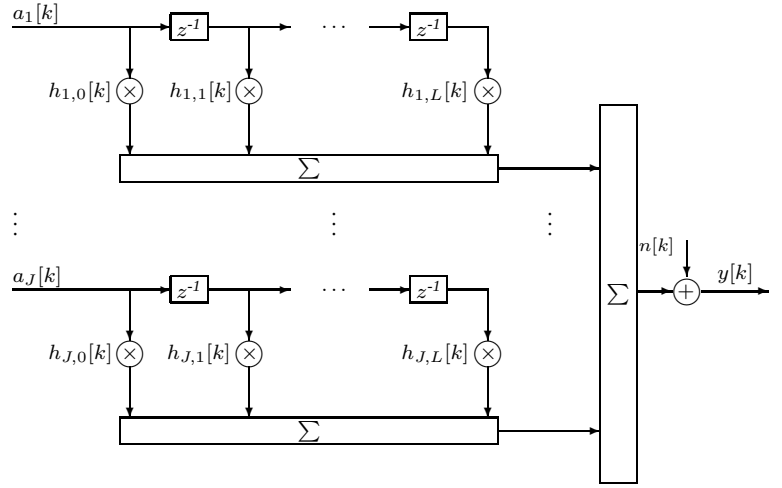


Fig. 3: Flowchart of the collision system.

B. Successive Interference Cancellation (SIC) technique

We propose, here, to use Serial Interference Cancellation technique (SIC) [3] to separate the AIS signals received by the LEO satellite called observations. Before that, some assumptions are necessary for the good behavior of the SIC technique,

- The additive noise is a white gaussian centered noise.
- The output sequences are independent from the noise.
- The channel is Gaussian with a single path but the sources may arrive late.
- To increase the probability estimation, The received signal is re-sampled by a factor which is equal to 4.

The mathematical model considered can then be re-expressed as follows,

$$y[k] = \sum_{j=1}^J s_j[k - \tau_j] + \sqrt{\frac{OSR}{2}} n[k], 0 \leq k \leq K-1, \quad (5)$$

where τ_j , for $j = 1, \dots, J$, is the delay associated with each signal j , $n[k]$ is an additive white Gaussian centered noise and the signal $s_j[k]$ is defined by, $s_j[k] = a_j \exp(-j(\phi_j[k] + 2\pi f_{d_j} \frac{k}{OSR \times SR}))$, where f_{d_j} is the doppler frequency of the j -th source, OSR is the OverSample Rate

which is equal to four (respecting the fourth assumption) and $SR = 9,6 kbps$, $a_j = \sqrt{10^{\frac{SNR_j}{10}}}$ is the amplitude of the j -th source, $\phi_j[k]$ is the instantaneous phase of the modulation defined by, $\phi_j[k] = \phi_j[k-1] + \frac{\pi}{2OSR}s_{gj}$ where (s_{gj} is the output of the Gaussian filter). Three main steps are established in order to apply the SIC technique,

- Estimation of transmission parameters.
- GMSK demodulation adapted to the AIS standard.
- A sequencer for successive demodulation established (Fig. (4)).

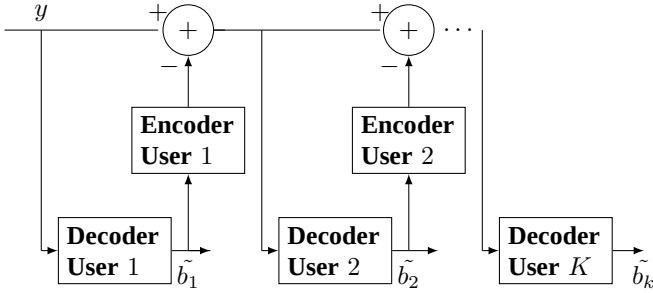


Fig. 4: SIC technique scheme.

C. Parameter estimation

At the initial stage, no prior information is available. In the first step, the received signal is assumed to be without collision. The estimation of the parameters f_{d0} (Doppler effect), τ_0 (Delay), a_0 (Amplitude) and h (modulation index) in Data Aided (DA) is done by a correlation-based technique using the known preamble of AIS signals (32 bits) (See [4] and [5]).

D. Scenario

The SIC technique exploits power level difference between the users and performs very well in scenarios without tight power control. The received signal is processed by estimating the doppler frequency $f_{d(j,max)}$ of the strongest signal $s_{(j,max)}[k]$ using the estimator proposed in [4][5], then we can estimate the delay $\tau_{(j,max)}$ and the amplitude $a_{(j,max)}$ in which are calculated and based on the received reference signals only, and then the estimated signal is subtracted from the received signal to construct a new processed received signal, the doppler frequency $f_{d(j-1,max)}$ of the second strongest signal $s_{(j-1,max)}[k]$ using the same estimator proposed, then one can estimate the delay $\tau_{(j-1,max)}$ and the amplitude $a_{(j-1,max)}$ which are recalculated based on the new processed received signal. The iterations are repeated until the final estimate of the desired signal is available or until a maximum number of iterations is reached.

IV. COMPUTER SIMULATIONS

Here, we present simulations to illustrate the effectiveness and the performances of the suggested separation technique. To that aim, we consider the following scenario: firstly, we

provides GMSK modulated signals with a bandwidth-bit-time product parameter $BT = 0.4$, the Sample Rate is $SR = 9600$ bps, the OverSample Rate is $OSR = 4$ and the Order Gaussian filter is $OF = 21$ within the technical characteristics of an AIS system mentioned in the previous section (the messages are generated according to the AIS recommendation: they are composed of 168 information bits. Then, the Ramp Up, Training Sequence, FCS (CRC), Buffer and the Start/End Flag bits are inserted). These signals are perturbed by an additive white gaussian noise. Finally, we detect AIS messages received by the LEO satellite and we separate the AIS signals using SIC technique presented before.

In the Tab. (I) (resp. the Tab. (II)) lists the different parameters (SNR, delay and doppler) for the two sources (resp. the three sources) to be transmitted, and these sources are mixed to create observations seen by the LEO satellite. Then, we have to estimate the transmitted sources only from the observations. All displayed results have been averaged over 30 Monte-Carlo trials.

User ID	1	2
SNR (dB)	0 ~ 18	0 ~ 18
Delay (ms)	0	0.0260
Doppler (Hz)	0	120

TABLE I: Input parameters for 2 users.

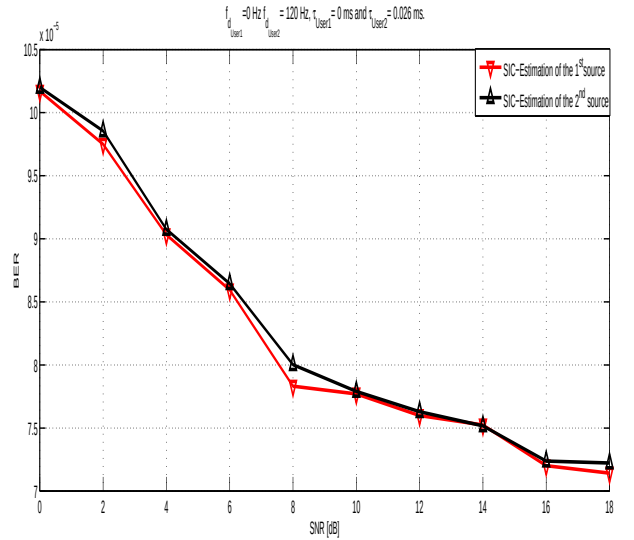


Fig. 5: SIC for 2 users: BER versus SNR for $\tau_{User1} = 0$ ms, $\tau_{User2} = 0.026042$ ms, $f_{dUser1} = 0$ Hz and $f_{dUser2} = 120$ Hz.

User ID	1	2	3
SNR (dB)	0 ~ 18	0 ~ 18	0 ~ 18
Delay (ms)	0	0.0065	0.0260
Doppler (Hz)	0	100	220

TABLE II: Input parameters for the 3 users.

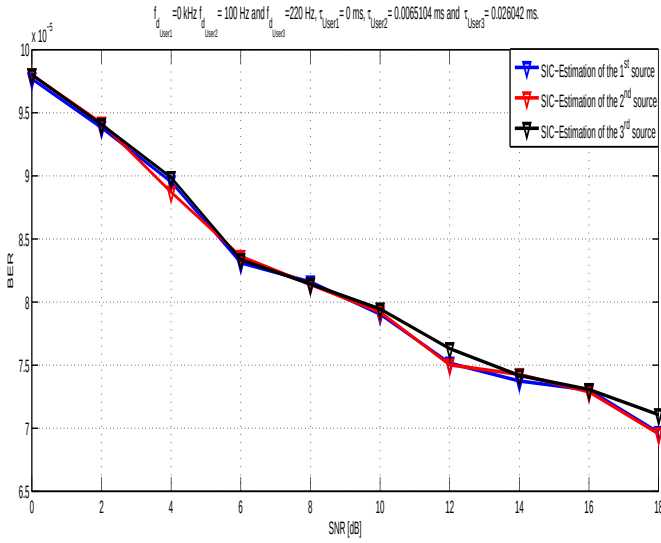


Fig. 6: SIC for 3 users: BER versus SNR for $\tau_{User1} = 0$ ms, $\tau_{User2} = 0.0065104$ ms, $\tau_{User3} = 0.026042$, $f_{dUser1} = 0$ Hz, $f_{dUser2} = 100$ Hz and $f_{dUser3} = 220$ Hz.

In the Fig. (5) (resp. the Fig.(6)), we show obtained performances when using classical GMSK demodulation and the SIC technique proposed for two users (resp. three users). We can observe that if we start with a high SNR (SNR= 0 dB), the BER sensitivity is high. However, when the SNR decreases, the bit error ratio decreases too for the two sources (resp. the three sources). In the Fig. (5), we observe that there is very small difference in BER value between the first and the second source. For example, when SNR = 2 dB (resp. 8 dB and 18 dB) the BER value for the first user is ≈ 9.8 and for the second user is $\approx 9.68 \times 10^{-5}$ (resp. for the first user is $\approx 7.8 \times 10^{-5}$, for the second user is $\approx 8 \times 10^{-5}$ and for the first user is $\approx 7.3 \times 10^{-5}$, for the second user is $\approx 7.19 \times 10^{-5}$). In the Fig.(6), we observe also the same thing that there is very small difference in BER value between the first, the second and the third user. However, when SNR = 4 dB (resp. 12 dB and 18 dB) the BER value for the first user is $\approx 8.85 \times 10^{-5}$, for the second user is $\approx 8.7 \times 10^{-5}$ and the third user is $\approx 9 \times 10^{-5}$ (resp. for the first user is $\approx 7.5 \times 10^{-5}$, for the second user is $\approx 7.5 \times 10^{-5}$, for the third user is $\approx 7.65 \times 10^{-5}$ and for the first user is $\approx 6.9 \times 10^{-5}$, for the second user is $\approx 6.95 \times 10^{-5}$ and for the third user is $\approx 7.2 \times 10^{-5}$). The SIC technique keeps its performance even if the doppler effect and the delay are imposed to users, therefore the SIC technique gives good results, since the BER is presented by small values in decreasing order with the SNR value (from 0 to 18 dB) which explain that the sources estimated from the observation seen by the LEO satellite are almost the same as the transmitted signal from vessels.

A. Discussion

SIC technique has some disadvantages, among them, when more than three AIS signals are mixed on the same time

slot. The main drawback was shown manageable as long as the power control algorithm took the estimation errors in the interference cancellation process into account [6]. This consideration unfortunately appeared to further complicate the required power control distribution, as each user would need to be assigned a specific power level depending on its path loss, estimation error, and the decoding order. Therefore, to avoid these problems we have to develop a new Non-Unitary Joint Block Diagonalization NU – JBD method [7] for blind separation sources with AIS application.

V. CONCLUSION

In this communication, we have presented a new technique for decollision of AIS signals, its first goal is correct reception of messages from different vessels with following the decollision signals from observations received by a LEO satellite using Serial Interference Cancellation (SIC) technique. Extensions for future research would be to develop a new method for the Non-Unitary Joint Block diagonalization NU – JBD which is occurs in the blind separation of convolutive mixtures of sources in order to avoid the problem of the SIC technique mentioned before.

REFERENCES

- [1] A. Linz and A. Hendrickson, "Efficient Implementation of an I-Q GMSK Modulator", IEEE Trans. Circuits and Systems-II: Analog and Digital Signal Processing, vo. 43, no.1, January 1996.
- [2] M. Zhou, A.-J. Van der Veen and R. van Leuken, "Multi-user LEO-satellite receiver for robust space detection of AIS messages", International Conference on Acoustics, Speech, and Signal Processing (ICASSP'2012), Kyoto, Japan, March 2012.
- [3] R.-C. de Lamare, T.-T.-V. Vinhoza and R. Sampaio-Neto, "Blind Adaptive and Iterative Interference Cancellation Receiver Structures based on the Constant Modulus Criterion in Multipath Channels", VI International Telecommunications Symposium (ITS'06), Fortaleza-CE, Brazil, September 2006.
- [4] P. Bianchi, P. Loubaton and F. Sirven, "Non data aided estimation of the modulation index of continuous phase modulations", IEEE Trans. Signal Processing, vol. 52, no. 10, pp. 2847-2861, 2004.
- [5] D. Bonacci and R. Prévost, "Advanced concepts for satellite reception of AIS messages", Toulouse Space Show 2012, Toulouse, France, June 2012.
- [6] A. Agrawal, J. Andrews, J. Cioffi, and T. Meng, "Power Control for Successive Interference Cancellation with Imperfect Cancellation", to appear, IEEE ICC, New York, May 2002.
- [7] H. Ghennoui, E.-M. Fadaili, N. Thirion-Moreau, A. Adib and E. Moreau, "A nonunitary joint block diagonalization algorithm for blind separation of convolutive mixtures of sources", in IEEE Signal Lett., vol. 14, no. 11, pp 860-863, November 2007.

JOINT MAXIMUM LIKELIHOOD SEQUENCE ESTIMATION ALGORITHM FOR DECOLLISION OF AIS SIGNALS IN MARITIME SURVEILLANCE CONTEXT.

O. Cherrak^{1,2}, H. Ghennioui¹, N. Thirion-Moreau² and E.-H. Abarkan¹

¹Université Sidi Mohamed Ben Abdellah, FSTF, LSSC, Faculté des Sciences et Techniques de Fès, Route Immouzer, B.P. 2202, Fès-Maroc.

²Aix-Marseille Université, CNRS, ENSAM, LSIS, 13397 Marseille-France, Université de Toulon, CNRS, LSIS, UMR 7296, 83957 La Garde-France.

Abstract—This communication addresses the Joint Maximum-Likelihood Sequence Estimation (JMLSE) algorithm of decollision of AIS (Automatic Identification System) signals in maritime surveillance context. AIS is a Self-Organized Time Division Multiple Access (SO-TDMA) system, in which vessels periodically transmit information. The main idea of this communication is decollision of signals received by a Low Earth Orbit (LEO) satellite lead to a global surveillance of maritime traffic using JMLSE algorithm, which is widely studied for other types of co-channel interference applications. These messages are transmitted by vessels using the AIS standards. Simulations are provided in order to evaluate the behavior of the JMLSE algorithm for separation and estimation of sources.

Index Terms— AIS signal, Maritime surveillance, GMSK modulation, Joint Maximum Likelihood Sequence Estimation, Collision, Blind Source Separation.

I. INTRODUCTION

The Moroccan activities of economic order insured by sea have experienced a highly significant increase, proportionately far greater than that of the world economic development: Maritime transport, Marina, etc. This change imposes an enhanced presence of the state to ensure and enhance the safety and the security at sea. It is necessary to adopt a strategy which emphasizes the monitoring in depth and reactivity. Whenever possible, an action at the nearest source of the threats or dangers is sought to detect the precursors. Maritime protection consists to develop and make available monitoring devices and intervention to guard against certain dangers, to maintain the optimum performance for the control of ships and to prevent or treat a wide range of existing threats at the sea or from the sea which are susceptible undermine our national territory, our interests and the national or international public order. It includes, of course, the maritime defense of the territory. This communication is a part of this problematic.

The major objective of this communication is to develop the JMLSE algorithm for blind source separation of linear mixtures intended specifically for maritime surveillance of the AIS system by a LEO satellite. To that aim, we provide a GMSK modulated signal within the technical characteristics of AIS system (Data encapsulation, NRZI line coding, the I/Q

GMSK modulator). Finally, we interest in separation and decollision of the AIS signals (Correct reception of messages from different vessels with following the decollision signals) and this only from observations.

II. TRANSMISSION STEP

A. AIS Message

The AIS message is a length of $N = 256$ bits. It contains a training sequence (TS) consisting 0 and 1 which takes 24 bits. The start flag (SF) and the end flag (EF) for information takes 8 bits. A control sequence (FCS) (16 bits Cyclic Redundancy Code (CRC)) is added to the data information (168 bits). A zero is inserted after every five continuous 1 in the information data and FCS as illustrated in Fig. (1).

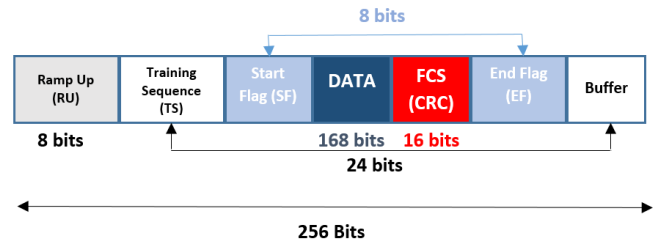


Figure 1: AIS message.

The binary sequence $\{a_n\}_{0 \leq n \leq N}$ of the message AIS takes the values $\{-1, +1\}$ since the line coding constitutes a fundamental block in the AIS system. The NRZI encoding is used as $a_n = b_n a_{n-1}$, $n \geq 0$ where $b_n \in \{-1, 1\}$. The modulation specified by AIS standard is Gaussian Minimum Shift Keying (GMSK) [1]. The encoded message is modulated and transmitted at 9600 bps on carriers at frequency 162 MHz.

B. GMSK modulator for AIS system

GMSK is a frequency-shift keying modulation [1], which produces constant-envelope and continuous-phase. Like any digital modulation, the GMSK is decomposed into three main steps,

- Mapping: Matching a set of bits to a set of symbols.
- Shaping: Transforming symbol sequence into a signal to be adapted to the transmission channel. It is usually realized by a linear filtering operation. The signal can be written as,

$$s_g[k] = \sum_{n=0}^{+\infty} a_n g[k - nT_s] \quad (1)$$

where a_n are the emitted symbols, T_s is the symbol period and $g[k]$ represents the shaping filter. In the GMSK case, $g[k]$ the response of a Gaussian filter whose expression is given by,

$$g[k] = \sqrt{\frac{2\pi}{\log 2}} B e^{\left(-\frac{2}{\log 2} (\pi B k)^2\right)} \quad (2)$$

where B is the bandwidth of the Gaussian filter, the GMSK modulation is described by the bandwidth-time product BT , AIS uses $BT = 0.4$ and $T_s = \frac{1}{9600}$ s.

- Modulation: Placing the signal on a frequency f_0 of a carrier. This operation produces a signal of spectral characteristic, which is adapted to the band-pass channel transmission. The GMSK modulated signal is expressed in the following form,

$$s[k] = \text{Re}\{e^{i\phi[k]} e^{i2\pi f_0 k}\} \quad (3)$$

$$= I[k] \cos(2\pi f_0 k) - Q[k] \sin(2\pi f_0 k)$$

where Re is the real part a complex number, $\phi[k]$ is the instantaneous phase of $s_g[k]$. $I[k]$ (resp. $Q[k]$) modulates the carrier in phase (resp. in phase quadrature). We are talking about I (real) and Q (imaginary) channels. The structure of the GMSK modulator is illustrate in the Fig. (2),

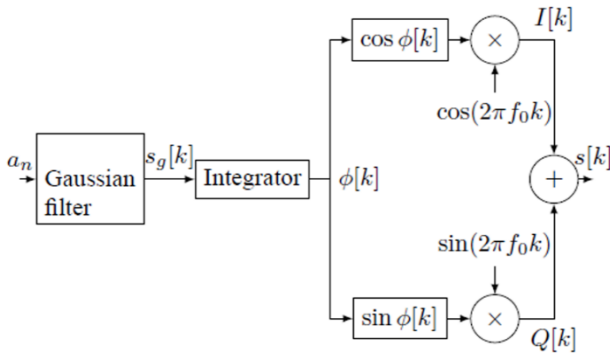


Figure 2: The GMSK modulator scheme.

III. RECEPTION STEP

A. Problem statement

The collision of messages caused by the large difference in propagation time from vessels to satellites [2] can be expressed as follows,

$$y[k] = \sum_{j=1}^J \sum_{l=0}^L h_{j,l}[k] s_j[k-l] + n[k], \quad 0 \leq k \leq K-1 \quad (4)$$

where $y[k]$ are the observations received at the satellite, $h_{j,l}[k]$ is the Impulse Response of j channel, $s_j[k]$ is the k -th sample of the j -th co-channel which comes from the GMSK modulation, and these observations are perturbed by an additive gaussian noise $n[k]$, J is the number of co-channel signals, K is the number of symbols per slot and L is the channel length. The structure collision model is presented in Fig. (3).

B. Joint Maximum Likelihood Sequence Estimation (JMLSE)

We propose, in this section, to use Joint Multiuser Detection method (JMD) to solve the problem of co-channel users separation. The principle of this algorithm is to estimate jointly the different data sequences. It is based on the Joint Maximum Likelihood Sequence Estimation (JMLSE) [2]. The

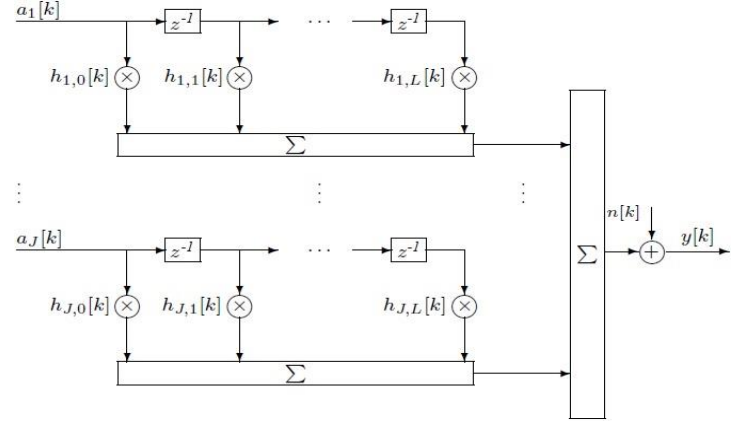


Figure 3: Flowchart of the collision system.

implementation of this algorithm uses the Viterbi algorithm [3]. The Viterbi Algorithm can be simply described as an algorithm, which finds the most likely path through a trellis, i.e. shortest path, to separate the AIS signals received by the LEO satellite called observations. Its first application was proposed as a solution to the decoding of convolutional codes [4]. Before that, some assumptions are necessary for the good behavior of the JMLSE algorithm,

- Doppler should not be strong,
- The additive noise is a white Gaussian centered noise,
- The output sequences are independent of the noise,
- To increase the probability estimation, the received signals are re-sampled by a factor, which is equal to four.
- The state machine of the FIR channel is finite.

C. Description of the JMLSE algorithm

Firstly, to simplify the calculations, we consider the model described in Eq. (6) by the following expression,

$$y[k] = \sum_{j=1}^J \sum_{l=0}^L h_{j,l}[k] a_j[k-l] + n[k], \quad 0 \leq k \leq K-1 \quad (5)$$

A simple mathematical development of Eq. (5) can reduce the problem to a least square minimization defined by the following equation,

$$(\hat{a}_1, \dots, \hat{a}_J) = \arg \min_{(\hat{a}_1, \dots, \hat{a}_J)} \{\sum_{k=0}^{K-1} \tau[k]\}, \quad (6)$$

where $\hat{a}_j = [\hat{a}_j[0], \dots, \hat{a}_j[K-1]]^T$, $1 \leq j \leq J$, represents the estimated sequence j , $\hat{a}_j$ is the assumption on the transmitted symbol and $\tau[k]$ represents the branch metric defined by:

$$\tau[k] = |y[k] - \sum_{j=1}^J \sum_{l=0}^L h_{j,l} \tilde{a}_j[k-l]|^2 \quad (7)$$

The interest of JMLSE algorithm is to find the sequence of the lowest metric. The metric of the paths is defined by:

$$\lambda[k+1] = \lambda[k] + \tau[k+1], \quad (8)$$

where $\lambda[k]$ is the metric of the path from the previous step in the lattice.

We propose, here, to use a semi-blind algorithm to estimate the channel. Moreover, this algorithm uses the training sequence to estimate the strongest transmitter. The channel estimation of this transmitter can be performed by the least square method. However, the other channels are initialized as follows,

$$\hat{\mathbf{h}}_j = [0, 0, \dots, C, 0, \dots, 0], \text{ for } 2 \leq j \leq J \quad (9)$$

where C is a real constant. The position of this constant depends on the selected channel model. We consider the vector, $\tilde{\mathbf{a}}_{j,v,\mu}(k+1) = [\tilde{a}_{j,v,\mu}(k+1), \dots, \tilde{a}_{j,v,\mu}(k-L+1)]^T$, for $1 \leq j \leq J$, the transmitted sequences at $k+1$ corresponds to the transition state from v to μ for the transmitter j . To pass from the state v to the state μ , we use the metric function $\tau_{v,\mu}(k+1)$ whose expression is given by the following equation,

$$\tau_{v,\mu}(k+1) = |y[k] - \hat{\mathbf{h}}_j^T(k) * \tilde{\mathbf{a}}_{j,v,\mu}(k+1)|^2. \quad (10)$$

The first step of the proposed algorithm is to compute the branch metric, then the metric of the paths is expressed in the following way,

$$\lambda_{v,\mu}[k+1] = \lambda_v[k] + \tau_{v,\mu}[k+1], \quad (11)$$

where $\lambda_v[k+1]$ is obtained by the minimization of $\lambda_{v,\mu}[k+1]$. After that, the sequence is estimated $\hat{\mathbf{a}}_j[k-\delta]$, where δ is a decisive constant for performance system. We notice that the vector $\hat{\mathbf{h}}_j$ of each channel j at $k+1$ is estimated according to the following optimization scheme:

$$\hat{\mathbf{h}}_j[k+1] = \hat{\mathbf{h}}_j[k] + \beta[k]e[k]\hat{\mathbf{a}}_j[k-\delta], \quad (12)$$

where $\beta[k]$ is the optimal adaptation step-size defined in LMS algorithm [5] and $e[k] = y[k-\delta] - \sum_{j=0}^J \hat{\mathbf{h}}_j^T[k] * \mathbf{a}_j[k-\delta]$.

D. Summary

The principle of the JMLSE algorithm is summarized in the following table,

Consider the samples of the received signal $y[k]$.
 Initialize the channel $\hat{\mathbf{h}}_1$ of the strongest transmitter using the least squares method and the other channels are initialized by Eq. (9).
 Initialize to all states v, μ the assumptions on the transmitted sequence for $k = 0$.
For $k = 0, \dots, K-1$
 --Calculate the branch metric whose expression is given by the Eq. (10).
 --Calculate the metric paths metric whose expression is given by the Eq. (11).
 --Determine the best path for each state μ by $\lambda_{v,\mu}[k+1] = \min\{\lambda_{v,\mu}[k+1]\}$.
 --Trace back to get the estimated $\hat{\mathbf{a}}_j[k-\delta]$.
 --Update the coefficients of the channel according to the adaptation rule whose expression is given by Eq. (12).
EndFor

IV. COMPUTER SIMULATIONS

Here, we present simulations to illustrate the effectiveness and the performances of the proposed algorithm. Firstly, we consider $J = 2$ users with instantaneous mixing (no multipath considered $L = 0$) and $K = 10$. However, We have considered these parameters in order to show the principle of the JMLSE algorithm,

$$\mathbf{a}_1 = [1 \quad -1 \quad 1 \quad 1 \quad 1 \quad -1 \quad 1 \quad -1 \quad -1 \quad -1],$$

$$\mathbf{a}_2 = [1 \quad -1 \quad 1 \quad -1 \quad 1 \quad -1 \quad -1 \quad -1 \quad 1 \quad 1].$$

The channel vectors are randomly generated according to a Gaussian distribution,

$$\mathbf{h}_1 = [0.4281, 0.8958, 1.4108, 0.3764, 0.3600, -0.0900, -1.6320, 0.8797, 0.0994, 0.8965],$$

$$\mathbf{h}_2 = [-0.3209, -1.5978, 1.7098, 0.1334, 0.3575, 0.4680, -0.3173, -0.0887, 2.4878, 0.8107].$$

The vector of observations is equal to,

$$\mathbf{y} = [0.1072 \quad 0.7020 \quad 3.1206 \quad 0.2430 \quad 0.7175 \quad -0.3780 \quad -1.3147 \quad -0.7910 \quad 2.3884 \quad -0.0858].$$

The different metric branches are calculated and given in the Tab. (1)

States	(1,1)	(-1,1)	(1,-1)	(-1,-1)
k=0	0	0.7330	0.4119	0.0460
k=1	1.9711	10.2117	3.2099	0
k=2	0	7.9610	11.6931	38.9507
k=3	0.0711	0.2364	0	0.5668
k=4	0	0.5185	0.5111	2.0593
k=5	0.5715	0.8760	0.0324	0
k=6	0.4026	6.9136	0	10.6531
k=7	2.5024	0.0315	3.0951	0
k=8	0.0395	0	22.8182	24.7562
k=9	3.2151	0	0.0295	2.6287

Table 1: Metric branches for two users with instantaneous mixing for the example considered.

The optimal lattice deduced from the Tab. (1) is then presented in the Fig. (4). According to the optimal lattice obtained, the sequences are estimated from the observations and found to be the same as transmitted ones.

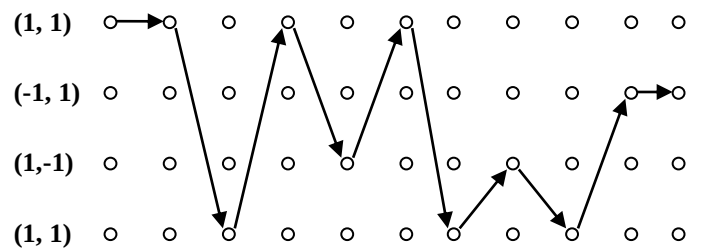


Figure 4: the optimal lattice for two users with instantaneous mixing for the example considered.

The estimated sequences are,

$$\hat{\mathbf{a}}_1 = [1 \quad -1 \quad 1 \quad 1 \quad 1 \quad -1 \quad 1 \quad -1 \quad -1 \quad -1],$$

$$\hat{\mathbf{a}}_2 = [1 \quad -1 \quad 1 \quad -1 \quad 1 \quad -1 \quad -1 \quad -1 \quad 1 \quad 1].$$

Finally, we consider the following scenario: we provide GMSK baseband signals with a bandwidth-bit-time product parameter $BT = 0.4$, the Sample Rate is $SR = 9600$ bps, the OverSample Rate is $OSR = 4$ and the Order Gaussian filter is $OF = 21$ within the technical characteristics of an AIS system mentioned in the previous section (the messages are generated according to the AIS recommendation: they are composed of 168 information bits. Then, the Ramp Up, Training Sequence, FCS (CRC), Buffer and the Start/End Flag bits are inserted). These signals are perturbed by an additive white Gaussian noise. Finally, we detect AIS messages received called observations and we separate them using JMLSE algorithm presented before.

We show the performances obtained when using the JMLSE algorithm proposed for three users, and these sources are mixed to create observations. Then, we have to estimate the transmitted sources only from the observations. All displayed results have been averaged over 10 Monte-Carlo trials.

In the Fig. (4), we observe that if we start with a high SNR ($SNR = 0$ dB), the MSE is high. However, when the SNR decreases, the mean square error decreases too for the three sources. We observe also that the MSE equal to zero when $SNR \geq 54$ dB

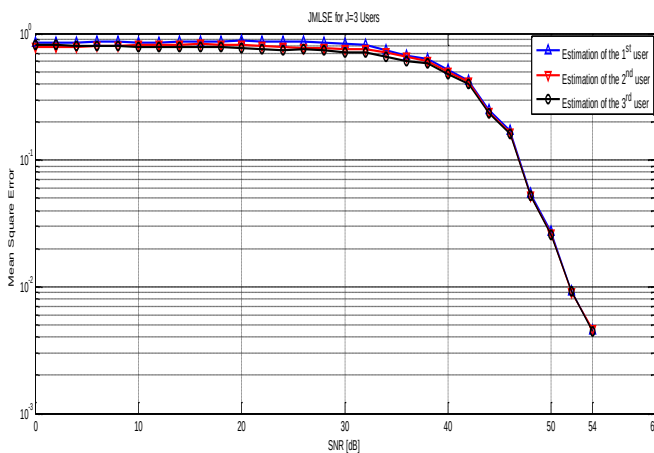


Figure 4: Mean-square-error performance of the JMLSE algorithm versus SNR for $J = 3$ users.

So all these charts illustrate the good behavior of the proposed algorithm. We observe a good estimation even if the noisy case is considered. Therefore, it has a very good resistance to amplitude problems.

E. Discussion

The JMLSE algorithm is sensitive to Doppler effect. Therefore, it needs a pretreatment in order to estimate the Doppler frequency for each user [6], this is why we have neglected the Doppler effect caused by the speed of the satellite. However, to avoid this problems we can develop a new Non-Unitary Joint Block Diagonalization NU-JBD method [7] for blind separation sources with AIS application.

V. CONCLUSION

In this communication, we have presented a new algorithm for decollision of AIS signals, its first goal is correct reception of messages from different vessels with following the decollision signals from observations received by a LEO satellite using Joint Maximum Likelihood Sequence Estimation (JMLSE). Extensions for future research would be to improve the JMLSE algorithm for the estimation of Doppler Effect for each user and to develop a new method for the Non-Unitary Joint Block diagonalization NU-JBD which is occurs in the blind separation of convolutive mixtures of sources in surveillance maritime context.

VI. REFERENCES

- [1] A. Linz and A. Hendrickson, "Efficient Implementation of an I-Q GMSK Modulator", IEEE Trans. Circuits and Systems-II: Analog and Digital Signal Processing, vol. 43, no.1, January 1996.
- [2] G. D. Forney, "Maximum-Likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference", IEEE Trans. Inform. Theory, 18:363–378, 1972.
- [3] Viterbi, A.J., "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm", IEEE Transactions on Information Theory, April 1967; IT - 13 (2), pp. 260 - 269.
- [4] Viterbi, A.J., "Convolutional Codes and Their Performance in Communication Systems", IEEE Transactions on Communications Technology, October 1971; COM - 19 (5), pp. 751 - 772.
- [5] O. Ali Abu-Ella, B. El-Jabu, "Optimal Robust Adaptive LMS Algorithm without Adaptation Step-Size", GSMM2008 Proceeding, 21-24 April 2008, pp 249 – 251.
- [6] P. Bianchi, P. Loubaton and F. Sirven, "Non data aided estimation of the modulation index of continuous phase modulations", IEEE Trans. Signal Processing, vol. 52, no. 10, pp. 2847-2861, 2004
- [7] H. Ghennioui, E.-M. Fadaili, N. Thirion-Moreau, A. Adib and E. Moreau, "A nonunitary joint block diagonalization algorithm for blind separation of convolutive mixtures of sources", in IEEE Signal Lett., vol. 14, no. 11, pp 860-863, November 2007.

Preconditioned optimization algorithms solving the problem of the non unitary joint block diagonalization : Application to blind separation of convolutive mixtures

Omar Cherrak^{a,b}, Hicham Ghennioui^a, Nadège Thirion-Moreau^{*b}, El Hossein Abarkan^a

^aLSSC, Université Sidi Mohamed Ben Abdellah, Faculté des Sciences et Techniques de Fès, Route Immouzer, B.P. 2202, Fès-Maroc, Tel/Fax: +212 5 35 60 80 14/+212 5 35 60 82 14

^bAix-Marseille Université, CNRS, ENSAM, LSIS, UMR 7296, F-13397 Marseille, France
Université de Toulon, CNRS, LSIS, UMR 7296, F-83957 La Garde, France, Tel: +33 4 94 14 24 56/26 71

Abstract

This article addresses the problem of the Non Unitary Joint Block Diagonalization (NU – JBD) of a given set of complex matrices for the blind separation of convolutive mixtures of sources (multidimensionnal deconvolution). We propose, here, to study different iterative optimization schemes based on Conjugate Gradient (CG), Preconditioned Conjugate Gradient (PCG), Levenberg-Marquardt and Quasi-Newton methods to determine which method offers the best compromise between efficiency and convergence speed in the studied context. To be able to derive all these algorithms, a preconditioner has to be computed which requires either the calculation of the complex Hessian matrices or the use of an approximation to these Hessian matrices. Furthermore, the optimal stepsize is also computed algebraically to speed up the convergence of these algorithms. Computer simulations are provided in order to illustrate the behavior of the different algorithms in various contexts: when exactly block-diagonal matrices are considered but also when these matrices are progressively perturbed by an additive gaussian noise. Finally, it is shown that these algorithms enable solving the problem of the blind separation if finite input response of finite impulse response (FIR) convolutive mixtures of sources.

Key words: Joint Block Diagonalization (JBD); Blind Source Separation (BSS); multidimensional deconvolution; optimization; preconditioning; complex Hessian matrices; convolutive mixtures; spatial quadratic time-frequency;

1. Introduction

With the use of more antennas equipped with more receivers, the volume of available data is continually increasing and advanced signal processing techniques are often required to be able to extract relevant information from this huge amount of data. The problem of the joint decomposition of matrices (or tensors) sets can come within this

*Corresponding author.

Email addresses: omar.cherrak@usmba.ac.ma (Omar Cherrak),
hicham.ghennioui@usmba.ac.ma (Hicham Ghennioui), thirion@univ-tln.fr (Nadège Thirion-Moreau), habarkan@yahoo.fr (El Hossein Abarkan)

framework, that is why it has gained a growing attention from the signal processing scientific community over the recent years. Whatever the considered field of application (astronomy, geophysics, remote sensing, biomedical, etc.), algorithms addressing the aforementioned problem are of interest if the problem at hand can finally be formulated as a problem of Array Processing and Direction-Of-Arrival estimation (DOA), Blind Source Separation (BSS), blind multidimensional deconvolution or even data mining or analysis.

Because traditionally a pre-whitening stage was used, one of first problem to have been considered was the so-called Joint Diagonalization (JD) of a given matrix set under the unitary constraint. It led to the nowadays well-known JADE (Joint Approximate Diagonalization of Eigenmatrices) [10] and SOBI (Second Order Blind Identification) [3] algorithms. The following works have addressed either the problem of the JD of tensors [17][39] or the problem of the JD of matrix sets, discarding the unitary constraint [12][20][35][38][45][47][49][55][56]. A fairly exhaustive overview of all the suggested approaches is available in [14]. This first particular type of matrix decomposition proves useful to solve two kinds of problems i) those of sources localization and direction finding and ii) those of blind separation of instantaneous mixtures of sources. At the same time, people have started to consider a second type of matrix decompositions - namely joint zero-diagonalization (JZD) -, since it is not uncommon to encounter this problem in different field of applications such as blind source separation (and more specifically spatial quadratic time-frequency domain methods [2][6]), telecommunications [11] and cryptography [52]. The first proposed algorithms operated under the unitary constraint [2], since once again, they were applied after a classical pre-whitening stage. But such a preliminary pre-whitening stage establishes a bound with regard to the best performances in the context of BSS that is why the unitary constraint was soon discarded, leading to several solutions among which are [23][22][13][58].

Recently, a more general problem has been addressed: that of the joint decomposition of a combination of given sets of complex matrices that can follow potentially different decompositions (for example when noncircular complex valued signals are considered and when people want to exploit additional statistical information by using both Hermitian and complex symmetric matrices simultaneously) [40][50][57]. Such issues can arise in various signal processing problems, among which are blind identification, separation or multidimensional deconvolution.

In this article, we focus on a fourth type of matrix decompositions, namely the Joint Block-Diagonalization (JBD). It is a rather general problem since the aforementioned JD problem is a special case of this one, but it could be seen as a part of the third problem even if - to our knowledge - only JD problems have been considered in that context up to this point. In such a decomposition, the matrices under consideration assume a specific algebraic structure, being block diagonal matrices *i.e.* they are block matrices whose diagonal blocks are square matrices of any (possibly even) size while their off-diagonal blocks are all null. First, this problem was addressed considering positive definite and hermitian block-diagonal matrices and unitary joint-block diagonalizers. Jacobi like algorithms were suggested [5][19]. Then, different alternative methods achieving the same task but for a non unitary joint-block diagonaliser have been proposed [16][25][26][27][32][41][48][54].

In this article, our goal is to investigate the impact/interest of the preconditioning process in the Joint Block-Diagonalization algorithms to be able to determine which method offers the best compromise between efficiency and convergence speed in the studied context (BSS and blind multidimensional deconvolution). But, first, to be able

to derive the different algorithms, the preconditioner has to be determined. It is the reason why, we start with the theoretical calculation of the complex Hessian matrices. But we will also consider approximations to these Hessian matrices. Then the different optimization algorithms can be derived and the optimal stepsize can be calculated. Computer simulations are provided in order to illustrate the behavior of the different algorithms in various contexts: when exactly block-diagonal matrices are considered but also when these matrices are progressively perturbed by an additive gaussian noise. Finally, we show how these algorithms find applications in blind separation of an instantaneous mixture of deterministic or random source signals, based on the use of Spatial Quadratic Time Frequency Spectra (SQTFS) or Spatial Quadratic Time Frequency Distributions (SQTFD).

This article is organized as follows: the general problem of non-unitary joint-block diagonalization is reminded in Section 2. Recalls are also made about the principle of preconditioning in optimization. In Section 3, we introduce the algebraic calculations of the Hessian matrices leading to the proposed possibly preconditioned non-unitary joint-block diagonalization algorithms. In Section 4, we show the different classical deterministic iterative optimization schemes considered and we introduce the calculation of the optimal step-size. Computer simulations are provided in Section 5 in order to illustrate the behavior of the resulting preconditioned JBD (and possibly JD) algorithms. To that aim, the different algorithms are compared in various contexts comparing them with "state-of-the-art approaches". The Section 6 enhance the usefulness of these algorithms through one of their possible applications, namely the blind separation of FIR convolutive mixtures of non-stationary sources. It is recalled that one possible manner to tackle such a problem consists of opting for a (non-unitary) joint-block diagonalization algorithm, but first particular matrices have to be properly selected using an automatic t - f points selection procedure [26]. Computer simulations are performed to illustrate the good performance of the suggested methods. Finally in Section 7, a conclusion is drawn and the perspectives of this work are delineated.

2. Problem statement & some recalls about optimization

2.1. The non unitary joint block-diagonalization problem and its assumptions

We recall that the problem of the non unitary joint block-diagonalization is stated in the following way [25]. Provided that we have a set $\mathcal{M}$ of N_m square matrices $\mathbf{M}_i \in \mathbb{C}^{M \times M}$, for all $i \in \{1, \dots, N_m\}$ that all admit the following decomposition:

$$\mathbf{M}_i = \mathbf{A} \mathbf{D}_i \mathbf{A}^H, \quad (1)$$

where $(\cdot)^H$ stands for the transpose conjugate (or Hermitian) operator and the N_m square $N \times N$ matrices $\mathbf{D}_i$ are block diagonal matrices, *i.e.* they are all in the form:

$$\mathbf{D}_i = \begin{pmatrix} \mathbf{D}_{i,11} & \mathbf{0}_{12} & \dots & \mathbf{0}_{1r} \\ \mathbf{0}_{21} & \mathbf{D}_{i,22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}_{r-1r} \\ \mathbf{0}_{r1} & \dots & \mathbf{0}_{rr-1} & \mathbf{D}_{i,rr} \end{pmatrix} \quad (2)$$

with r the number of considered blocks ($r \in \mathbb{N}^*$), $\mathbf{D}_{i,jj}$, $i \in \{1, \dots, N_m\}$, $j \in \{1, \dots, r\}$ are $n_j \times n_j$ square matrices so that $n_1 + \dots + n_r = N$ where $\mathbf{0}_{ij}$ denotes the $(n_i \times n_j)$ null matrix.

$\mathbf{A}$ is a $M \times N$ ($M \geq N$) full rank matrix while $\mathbf{A}^\dagger$ is a pseudo-inverse of $\mathbf{A}$ (or Moore-Penrose generalized matrix inverse). It is a $N \times M$ matrix denoted by $\mathbf{B}$ ($\mathbf{B} = \mathbf{A}^\dagger$). The set of the N_m square matrices $\mathbf{D}_i \in \mathbb{C}^{N \times N}$ is denoted by $\mathcal{D}$. The block sizes n_j for all $j = 1, \dots, r$ are assumed known. In the context of blind source separation, $\mathbf{A}$ is known as the mixing matrix whereas $\mathbf{B}$ is called the separating matrix, N is the number of sources while M is the number of sensors. Many BSS methods go through a joint matrix decomposition stage in order to estimate either the mixing matrix or the separating matrix even if other approaches have been considered too (direct or deflation methods for example). Moreover, this joint matrix decomposition problem often brings us back to a joint diagonalisation (resp. joint block diagonalisation) problem when instantaneous (resp. convolutive) mixtures of sources are considered. An unitary constraint on the matrices $\mathbf{A}$ or $\mathbf{B}$ is the result of a preprocessing stage called spatial whitening of the observations. But studies have proven that this step establishes a bound with regard to the best reachable performances in the context of BSS. That is why recently, there has been a strong interest in methods discarding this constraint. The general non-unitary JBD problem consists of estimating the matrix $\mathbf{A}$ and the block-diagonal matrices set $\mathcal{D}$ from only the matrix set $\mathcal{M}$. A standard way to determine the unknown matrices, in this case $\mathbf{A}$ and $\{\mathbf{D}_i\}$, consists of minimizing the following quadratic objective function with respect to the different involved matrices:

$$\mathcal{C}_{\text{JBD}}^{(1)}(\mathbf{A}, \{\mathbf{D}_i\}) = \sum_{i=1}^{N_m} \|\mathbf{M}_i - \mathbf{A}\mathbf{D}_i\mathbf{A}^H\|_F^2, \quad (3)$$

where $\|\cdot\|_F$ stands for the Frobenius norm.

But, in [26], it was shown that one can proceed rather differently, using the following objective function instead:

$$\mathcal{C}_{\text{JBD}}^{(2)}(\mathbf{B}, \{\hat{\mathbf{D}}_i\}) = \sum_{i=1}^{N_m} \|\text{OffBdiag}_{(\mathbf{n})}\{\mathbf{B}\mathbf{M}_i\mathbf{B}^H\}\|_F^2 \stackrel{\text{def}}{=} \mathcal{C}_{\text{JBD}}(\mathbf{B}), \quad (4)$$

which is much more simple since it happens that one single matrix is finally involved. We recall that the two matrix operators $\text{Bdiag}_{(\mathbf{n})}\{\cdot\}$ and $\text{OffBdiag}_{(\mathbf{n})}\{\cdot\}$ are respectively defined as:

$$\text{Bdiag}_{(\mathbf{n})}\{\mathbf{M}\} = \begin{pmatrix} \mathbf{M}_{11} & \mathbf{0}_{12} & \dots & \mathbf{0}_{1r} \\ \mathbf{0}_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0}_{r1} & \mathbf{0}_{r2} & \dots & \mathbf{M}_{rr} \end{pmatrix}, \quad (5)$$

and

$$\text{OffBdiag}_{(\mathbf{n})}\{\mathbf{M}\} = \mathbf{M} - \text{Bdiag}_{(\mathbf{n})}\{\mathbf{M}\} = \begin{pmatrix} \mathbf{0}_{11} & \mathbf{M}_{12} & \dots & \mathbf{M}_{1r} \\ \mathbf{M}_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{M}_{r1} & \mathbf{M}_{r2} & \dots & \mathbf{0}_{rr} \end{pmatrix}, \quad (6)$$

if the vector $\mathbf{n} = (n_1, n_2, \dots, n_r)$, and the $(N \times N)$ square matrix $\mathbf{M} = (M_{ij})$ has the

following structure:

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} & \dots & \mathbf{M}_{1r} \\ \mathbf{M}_{21} & \dots & \dots & \vdots \\ \vdots & \dots & \dots & \vdots \\ \mathbf{M}_{r1} & \mathbf{M}_{r2} & \dots & \mathbf{M}_{rr} \end{pmatrix}, \quad (7)$$

where $\mathbf{M}_{ij}$ for all $i, j = 1, \dots, r$ are $(n_i \times n_j)$ matrices (and $n_1 + \dots + n_r = N$).

Several solutions have been suggested to solve the JBD problem. Some of them are based on the optimization of the cost function defined in (3) while others are based on the optimization of the objective function defined in (4) [19][25]. In [19], the authors restrain their attention to the case of an unitary matrix $\mathbf{B}$ and suggest a Jacobi-like method. In [16][25], this assumption is discarded. In [16][25][26], positive definite and hermitian block diagonal matrices are considered. In [25], the criterion (4) is rewritten to perform a line by line iterative algebraic optimization involving a singular values decomposition. In [27], the considered matrices only have to be hermitian ones and the proposed solution is based on a gradient approach, yet, an approximation in the calculation is made to derive the expression of the complex gradient matrix of the cost function (4). The exact calculation of the complex gradient matrix is performed in [26].

Finally, notice that when $\mathbf{D}_{i,jj}, i \in \{1, \dots, N_m\}, j \in \{1, \dots, r\}$ are no more matrices but scalars instead ($n_j = 1$ for all $j \in \{1, \dots, r\}$ which also implies that $r = N$) it brings us back to the classical problem of the non unitary joint diagonalization of a given matrix set [14]. In this case, the two aforementioned matrix operators can be simplified and simply denoted by $\text{OffDiag}\{\cdot\}$ and $\text{Diag}\{\cdot\}$:

$$\text{OffBdiag}_{(\mathbf{1})}\{\mathbf{M}\} = ((1 - \delta_{ij})m_{ij})\mathbf{1}_N = \text{OffDiag}\{\mathbf{M}\},$$

with $\mathbf{M} = (m_{ij})$, $\mathbf{1}_N$ is a $(N \times N)$ matrix whose elements are all equal to 1 and $\delta_{ij} = 1$ if $i = j$ and 0 otherwise and

$$\text{Bdiag}_{(\mathbf{1})}\{\mathbf{M}\} = ((\delta_{ij})m_{ij})\mathbf{1}_N = \text{Diag}\{\mathbf{M}\} = \mathbf{M} - \text{OffDiag}\{\mathbf{M}\}.$$

In this case, it leads to the minimization of the following objective function:

$$\mathcal{C}_D(\mathbf{B}) = \sum_{k=1}^{N_m} \|\text{OffDiag}\{\mathbf{B}\mathbf{M}_k\mathbf{B}^H\}\|_F^2. \quad (8)$$

It has been used considering the constraint of an unitary matrix $\mathbf{B}$ in [10] (*i.e.* $M = N$ and $\mathbf{B}\mathbf{B}^H = \mathbf{B}^H\mathbf{B} = \mathbf{I}_N$, with $\mathbf{I}_N$ the $(N \times N)$ identity matrix) and the first suggested algorithms were based on a generalization of the Jacobi method [29]. Later, this cost function was considered for non unitary matrix [23][43][55].

2.2. Some recalls about optimization and preconditioning

In the previous section, we have recalled that the problem of the non unitary joint block diagonalization which arises in many signal processing applications (array processing and Direction-Of-Arrival localisation and blind sources separation remaining certainly the leading applications) can finally be formulated as a general-purpose unconstrained optimization problem. In the literature, one can find many unconstrained optimization

methods and some of them, namely gradient descent, steepest descent, relative gradient and conjugate gradient methods have already been studied and applied in this specific context of JBD [26][41]. They are known as first-order methods since they only depend on the gradient of the cost function and not on its second derivatives (*i.e.* the Hessian matrix denoted here by $\mathbf{H}_\bullet$). Notice that in the next section we will clarify this notation). To our knowledge, very few, if any, studies have been led on preconditioned optimization methods in such a framework. Yet, when one is designing an algorithm for a specific application, many conflicting requirements have to be taken into account (convergence rate (as few iterations as possible), computation time per iteration (as few floating point operations as possible), stability, sensitivity to numerical errors, robustness versus noise and/or model errors, problems of initialization, storage requirements, potential parallelization, ease of implementation, real time implementation, just to name a few), and at the end, for one given application, one must choose the algorithm which offers the best compromise between all those requirements. That is why we have chosen, here, to look specifically at preconditioned optimization methods to determine whether they are interesting or not for joint block diagonalization problems.

The main purposes for which preconditioning is generally used, are:

- to further accelerate the convergence since gradient descent or steepest descent algorithms are well known for their sometimes slow convergence,
- to increase stability and robustness versus noise or model errors,
- to cope with the fact that variables have wildly different magnitudes (since the most simple form of preconditioning is a scaling of the variables (thanks to a diagonal preconditioner) with well chosen coefficients),
- to cope with the fact that the cost function can rapidly change in some directions and slowly in other ones,
- to be able to tackle ill-posed problems,
- to ensure certain constraints such as non-negativity (see for example the non-negative matrix factorization algorithm (NMF) based on multiplicative updates suggested by Lee and Seung [33] which can be interpreted as a preconditioned gradient algorithm).

Sometimes the preconditioning process just speed up the convergence, but there also exist some difficult cases which cannot be solved without good preconditioning. That is why several works have addressed the problem of the design of effective preconditioners. First, the preconditioning has been used in direct methods and, latter this concept has been extended to the case of iterative processes (Chebyshev acceleration) in [21][51]. The key issue remains obviously the choice of the preconditioner denoted by $\mathbf{P}$. In the literature, one can find many works on different types of preconditioners, among which are the Jacobi and Gauss-Seidel preconditioners [53]. The first one consists of using the diagonal of the desired matrix while the second one consists of decomposing the preconditioner matrix into a lower triangular matrix, a diagonal matrix, and an upper triangular matrix. Yet, for non-symmetric problems, such preconditioners are limited in their effectiveness and robustness [1].

The analysis of the convergence rate of the preconditioned gradient descent method provides a substantial help in the choice of $\mathbf{P}$ and directions to follow. If the cost function is twice differentiable in the neighborhood of a local minimizer, *i.e.* with Hessian matrix $\mathbf{H}_\bullet \stackrel{\text{def}}{=} \nabla_{\mathbf{B}}^2 \mathcal{C}_{\text{JBD}} \stackrel{\text{def}}{=} \frac{\partial^2 \mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B} \partial \mathbf{B}^*}$ then the fastest convergence is obtained when the preconditioner $\mathbf{P}$ minimizes the condition number, namely κ , of the product $\mathbf{P}\mathbf{H}_\bullet$. The condition number stands for the ratio of the largest singular value of $\mathbf{P}\mathbf{H}_\bullet$ divided by the smallest singular value of this matrix *i.e.* $\kappa \stackrel{\text{def}}{=} \frac{\lambda_{\max}(\mathbf{P}\mathbf{H}_\bullet)}{\lambda_{\min}(\mathbf{P}\mathbf{H}_\bullet)}$. For quadratic cost functions, the ideal preconditioner would be $\mathbf{P} = (\mathbf{H}_\bullet)^{-1}$ so that $\mathbf{P}\mathbf{H}_\bullet = \mathbf{I}$ since the identity matrix $\mathbf{I}$ possesses the minimal (unity) condition number. But since it is not always easy to calculate or to compute $(\mathbf{H}_\bullet)^{-1}$ (especially when N or M become large in our case), it remains interesting to develop preconditioners that approximate $(\mathbf{H}_\bullet)^{-1}$ and that are cheaper to construct and to apply [15][7].

To be able to design a suitable preconditioning matrix $\mathbf{P}$ in the context of JBD, we will start with the calculation of the four exact complex Hessian matrices. Then, we will be able to derive several new preconditioned algorithms among which is a preconditioned Conjugate Gradient algorithm. The conjugate gradient method is an iterative optimization method which has become really popular in nonlinear optimization [30][42] and its efficiency may be improved by a proper choice of the preconditioning matrix [1].

3. New preconditioned joint block-diagonalization algorithms

To estimate the complex matrix $\mathbf{B} \in \mathbb{C}^{N \times M}$, the cost function given in (4) has to be minimized. To that aim, the differential $d\mathcal{C}_{\text{JBD}}$ of $\mathcal{C}_{\text{JBD}}$ has to be derived, involving the calculation of the partial derivatives $\frac{\partial}{\partial \cdot}$ of the objective function $\mathcal{C}_{\text{JBD}}$ with respect to $\mathbf{B}$ and $\mathbf{B}^*$ (it is the reason why, from now on, the objective function is denoted by $\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)$ where $\mathbf{B}^*$ stands for the complex conjugate of the complex matrix $\mathbf{B}$, except in subsection (4.3)). Then, the complex gradient matrix can be derived. Eventually, the second order differential $d^2\mathcal{C}_{\text{JBD}}$ can be calculated too in order to derive the complex Hessian matrices.

3.1. Complex gradient matrix

The $N \times M$ complex gradient matrix of our real-valued scalar cost function given in (4), which is denoted by $\nabla_{\mathbf{B}} \mathcal{C}_{\text{JBD}}$, is defined as in [31] (p. 63 and p. 77) by:

$$\nabla_{\mathbf{B}} \mathcal{C}_{\text{JBD}} \stackrel{\text{def}}{=} \frac{\partial \mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}^*} \quad \text{and} \quad \mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)) = \text{vec}^T \left(\frac{\partial \mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}^*} \right), \quad (9)$$

where $\text{vec}^T(\mathbf{A}) = (\text{vec}(\mathbf{A}))^T$, operator $\text{vec}(\cdot)$ is stacking the columns of a matrix into a column vector and $(\cdot)^T$ stands for the transpose operator. This quantity has been calculated in [26] and found to be equal to:

$$\frac{\partial \mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}^*} = \sum_{i=1}^{N_m} \left[\text{OffBdiag}_{(n)} \{ \mathbf{B} \mathbf{M}_i \mathbf{B}^H \} \mathbf{B} \mathbf{M}_i^H + \left(\text{OffBdiag}_{(n)} \{ \mathbf{B} \mathbf{M}_i \mathbf{B}^H \} \right)^H \mathbf{B} \mathbf{M}_i \right]. \quad (10)$$

3.2. Complex Hessian matrices

The complex matrix $\mathbf{H}$ is a square $2NM \times 2NM$ matrix constituted of four $NM \times NM$ square blocks which are the Hessian matrices:

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{\mathbf{B},\mathbf{B}^*} & \mathbf{H}_{\mathbf{B}^*,\mathbf{B}^*} \\ \mathbf{H}_{\mathbf{B},\mathbf{B}} & \mathbf{H}_{\mathbf{B}^*,\mathbf{B}} \end{bmatrix}. \quad (11)$$

To calculate the aforementioned matrix $\mathbf{H}$, we follow the same path as in [31] (details of this calculation can be found in the Appendix A). In short, we have to derive the expression of the second-order complex differential of $\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)$ which is then written as:

$$d^2\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*) = \begin{bmatrix} d\text{vec}^T(\mathbf{B}^*) & d\text{vec}^T(\mathbf{B}) \end{bmatrix} \begin{bmatrix} \mathbf{A}_{00} & \mathbf{A}_{01} \\ \mathbf{A}_{10} & \mathbf{A}_{11} \end{bmatrix} \begin{bmatrix} d\text{vec}(\mathbf{B}) \\ d\text{vec}(\mathbf{B}^*) \end{bmatrix}. \quad (12)$$

As shown in [31], the four $NM \times NM$ complex matrices $\mathbf{A}_{00}$, $\mathbf{A}_{01}$, $\mathbf{A}_{10}$ and $\mathbf{A}_{11}$ involved in the above expression are linked to the Hessian matrices by:

$$\mathbf{H}_{\mathbf{B},\mathbf{B}^*}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)) = \frac{\mathbf{A}_{00} + \mathbf{A}_{11}^T}{2} = (\mathbf{H}_{\mathbf{B}^*,\mathbf{B}}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)))^T, \quad (13)$$

$$\mathbf{H}_{\mathbf{B}^*,\mathbf{B}^*}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)) = \frac{\mathbf{A}_{01} + \mathbf{A}_{01}^T}{2}, \quad (14)$$

$$\mathbf{H}_{\mathbf{B},\mathbf{B}}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)) = \frac{\mathbf{A}_{10} + \mathbf{A}_{10}^T}{2}. \quad (15)$$

There are shown to be equal to (see the Appendix A.1 for the detailed calculations):

$$\begin{aligned} \mathbf{A}_{00} &= (\mathbf{M}_i^T \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{\text{Boff}}^T (\mathbf{B}^* \mathbf{M}_i^* \otimes \mathbf{I}_N) + (\mathbf{M}_i^* \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{\text{Boff}}^T (\mathbf{B}^* \mathbf{M}_i^T \otimes \mathbf{I}_N) \\ &\quad + \mathbf{M}_i^* \otimes \text{OffBdiag}_{(n)}\{\mathbf{B} \mathbf{M}_i \mathbf{B}^H\} + \mathbf{M}_i^T \otimes \text{OffBdiag}_{(n)}\{\mathbf{B} \mathbf{M}_i^H \mathbf{B}^H\} \\ &= \mathbf{A}_{11}^*, \end{aligned} \quad (16)$$

$$\begin{aligned} \mathbf{A}_{10} &= \mathbf{K}_{N,M}^T (\mathbf{I}_N \otimes \mathbf{M}_i \mathbf{B}^H) \mathbf{T}_{\text{Boff}}^T (\mathbf{B}^* \mathbf{M}_i^* \otimes \mathbf{I}_N) \\ &\quad + \mathbf{K}_{N,M}^T (\mathbf{I}_N \otimes \mathbf{M}_i^H \mathbf{B}^H) \mathbf{T}_{\text{Boff}}^T (\mathbf{B}^* \mathbf{M}_i^T \otimes \mathbf{I}_N) = \mathbf{A}_{01}^*, \end{aligned} \quad (17)$$

where the operator $\otimes$ denotes the Kronecker product, $\mathbf{K}_{N,M}$ is a square commutation matrix of size $NM \times NM$ and $\mathbf{T}_{\text{Boff}} = \mathbf{I}_{N^2} - \mathbf{T}_{\text{Diag}}$, is the $N^2 \times N^2$ “transformation” matrix, with $\mathbf{T}_{\text{Diag}} = \text{diag}\{\text{vec}(\text{BDiag}\{\mathbf{1}_N\})\}$, $\mathbf{1}_N$ is the $N \times N$ matrix whose components are all ones, $\text{diag}\{\mathbf{a}\}$ is a square diagonal matrix whose diagonal elements are the elements of the vector $\mathbf{a}$, $\mathbf{I}_{N^2} = \text{Diag}\{\mathbf{1}_{N^2}\}$ is the $N^2 \times N^2$ identity matrix, and $\text{Diag}\{\mathbf{A}\}$ is the square diagonal matrix with the same diagonal elements as $\mathbf{A}$.

One can check that the bigger complex matrix $\mathbf{H}$ defined in (11) is hermitian and that two parts can be distinguished in this matrix $\mathbf{H}$: its block-diagonal part and its off-block diagonal part.

According to [46], it is advisable to consider only the diagonal (real part) of the Hessian matrix *i.e.* (13) and (16) to ensure a non increase of the cost function, and to easily obtain its inverse. The $NM \times NM$ preconditioning matrix, denoted by $\mathbf{P}$, that we suggest thus, reads:

$$\mathbf{P} = (\hat{\mathbf{H}})^{-1}, \quad (18)$$

$$\hat{\mathbf{H}} = \text{Diag}\{\mathbf{H}_{\mathbf{B},\mathbf{B}^*}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*))\} = \text{Diag}\{\mathbf{H}_{\mathbf{B}^*,\mathbf{B}}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*))\}. \quad (19)$$

Hence in the next sections, we can focus on the different possibly preconditioned algorithms that can be derived, and try to better assess their behavior in different contexts.

4. Classical deterministic iterative optimization schemes

To estimate the joint block diagonalizer $\mathbf{B}$, the cost function given in (4) has to be minimized. To that aim different deterministic iterative optimization schemes can be considered. It means that $\mathbf{B}$ has to be re-estimated at each iteration m , so from now on, it will be denoted by $\mathbf{B}^{(m)}$. This matrix can also be stored in a vector $\mathbf{b}^{(m)}$ if the $\text{vec}(\cdot)$ operator is applied *i.e.* $\mathbf{b}^{(m)} = \text{vec}(\mathbf{B}^{(m)})$. Depending on the considered optimization scheme, $\mathbf{B}^{(m)}$ for all m is updated according to one of the following adaptation rules. For example, when a classical steepest descent approach is considered (gradient approach if the step size is a fixed positive scalar *i.e.* $\mu^{(m)} = \mu$ for all m), $\mathbf{B}$ is updated at each iteration m ($m = 1, 2, \dots$) according to:

$$\mathbf{B}^{(m+1)} = \mathbf{B}^{(m)} - \mu^{(m)} \nabla_{\mathbf{B}} \mathcal{C}_{\text{JBD}}(\mathbf{B}^{(m)}, \mathbf{B}^{*(m)}), \quad (20)$$

or equivalently

$$\mathbf{b}^{(m+1)} = \mathbf{b}^{(m)} - \mu^{(m)} \mathbf{g}_{\mathbf{B}^*}^{(m)}, \quad (21)$$

where $\nabla_{\mathbf{B}} \mathcal{C}_{\text{JBD}}(\mathbf{B}^{(m)}, \mathbf{B}^{*(m)})$ is the gradient matrix given in (10) or $\mathbf{g}_{\mathbf{B}^*}^{(m)}$ is its vectorized form *i.e.* $\mathbf{g}_{\mathbf{B}^*}^{(m)} = \left(\mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{JBD}}(\mathbf{B}^{(m)}, \mathbf{B}^{*(m)})) \right)^T$ and $\mu^{(m)}$ is a real positive constant called the step size (the problem of its choice is treated in the subsection 4.3). This algorithm may converge slowly, that is why we are interested in preconditioned algorithms in order to speed up the convergence.

4.1. Preconditioned Non Linear Conjugate Gradient

To further accelerate convergence, conjugate gradient methods modify the search directions to ensure that they are approximately mutually conjugate *i.e.* orthogonal with respect to an inner product related to the Hessian of the cost function. Thus, when the preconditioned conjugate gradient algorithm is considered, the search direction, $\mathbf{d}$, is given at the first iteration by:

$$\mathbf{d}^{(1)} = -\mathbf{P}^{(1)} \mathbf{g}_{\mathbf{B}^*}^{(1)}. \quad (22)$$

Then, $\forall m = 2, \dots$, the following adaptation rule is used:

$$\begin{cases} \mathbf{b}^{(m+1)} = \mathbf{b}^{(m)} + \mu^{(m)} \mathbf{d}^{(m)}, \\ \mathbf{d}^{(m+1)} = -\mathbf{P}^{(m+1)} \mathbf{g}_{\mathbf{B}^*}^{(m+1)} + \beta^{(m)} \mathbf{d}^{(m)}, \end{cases} \quad (23)$$

where the expression of the preconditioner $\mathbf{P}$ is given in (18)-(19). It should be a positive-definite matrix, so non positive diagonal elements are prohibited: if it occurs, the algorithm is no more preconditioned ($\mathbf{P}$ is then set to the identity matrix $\mathbf{I}_{NM}$ and we come back to the classical non linear conjugate gradient algorithm). In exact line search methods, several expressions for β are classically used: the Fletcher-Reeves (β_{FR}), the Polak-Ribière (β_{PR}) and the Dai-Yuan (β_{DY}) formula [18][44][46]:

$$\beta_{\text{FR}}^{(m+1)} = \frac{(\mathbf{g}_{\mathbf{B}^*}^{(m+1)})^H \mathbf{P}^{(m+1)} \mathbf{g}_{\mathbf{B}^*}^{(m+1)}}{(\mathbf{g}_{\mathbf{B}^*}^{(m)})^H \mathbf{P}^{(m)} \mathbf{g}_{\mathbf{B}^*}^{(m)}}, \quad (24)$$

$$\beta_{\text{PR}}^{(m+1)} = \frac{(\mathbf{g}_{\mathbf{B}^*}^{(m+1)} - \mathbf{g}_{\mathbf{B}^*}^{(m)})^H \mathbf{P}^{(m+1)} \mathbf{g}_{\mathbf{B}^*}^{(m+1)}}{(\mathbf{g}_{\mathbf{B}^*}^{(m)})^H \mathbf{P}^{(m)} \mathbf{g}_{\mathbf{B}^*}^{(m)}}, \quad (25)$$

$$\beta_{\text{DY}}^{(m+1)} = -\frac{(\mathbf{g}_{\mathbf{B}^*}^{(m+1)})^H \mathbf{P}^{(m+1)} \mathbf{g}_{\mathbf{B}^*}^{(m+1)}}{(\mathbf{d}^{(m)})^H \mathbf{P}^{(m)} (\mathbf{g}_{\mathbf{B}^*}^{(m+1)} - \mathbf{g}_{\mathbf{B}^*}^{(m)})}. \quad (26)$$

The resulting algorithm is denoted by $\text{JBD}_{\text{PCGEH}}$ (EH stands for Exact Hessian).

4.1.1. Levenberg-Marquardt algorithm

When the Hessian matrix $\hat{\mathbf{H}}$ tends to lose its positive-definiteness property through the iterations, and hence may fail to construct descent directions, it is better to stabilize it using trust region techniques that modify $\hat{\mathbf{H}}$ by adding a multiple of the identity matrix. It is the principle of the Levenberg-Marquardt approach [36]:

$$\mathbf{b}^{(m+1)} = \mathbf{b}^{(m)} - \mu^{(m)} \left(\hat{\mathbf{H}}^{(m)} + \alpha \mathbf{I}_{NM} \right)^{-1} \mathbf{g}_{\mathbf{B}^*}^{(m)}, \quad (27)$$

where α is a relaxation coefficient. We notice that by setting $\hat{\mathbf{H}} = \mathbf{I}_{NM}$ in (27) (or by considering that α is chosen high enough), the gradient algorithm in (21) may be obtained. It can also be seen as a variant of the preconditioned conjugate gradient algorithm given in (23) where $\beta = 0$ and where the preconditioner is not directly the inverse of the Hessian but is slightly modified. With regards to $\hat{\mathbf{H}}$, we can use either our solution given in (18)-(19). The resulting algorithm is denoted by JBD_{LM_{EH}} (again EH stands for Exact Hessian).

4.2. Quasi-Newton algorithm

In (27), by setting $\alpha = 0$ and considering that the preconditioner $\mathbf{P}$ is a $NM \times NM$ calculation of the inverse of the Hessian matrix given by (18) and (19), we obtain the Quasi-Newton adaptation rule. In this case, the algorithm is initialized using for $\mathbf{P}^{(1)}$ (or $\hat{\mathbf{H}}^{(1)}$), a symmetric, $NM \times NM$ positive-definite matrix. The resulting algorithm is denoted by JBD_{QNEH} (QN standing for Quasi-Newton).

4.3. Choice of the step size: Enhanced Line Search or seek of the optimal step-size

It remains essential to find a good stepsize μ in the given search direction $\mathbf{d}^{(m)}$ at each iteration m . This stepsize $\mu^{(m)}$ can be found in different ways: it can be fixed or decreased versus the iterations, or we can opt for a global search with Enhanced Line Search (ELS) or for an approximation by a line search method like backtracking [8] (which is a locally optimal stepsize method). This subsection is dedicated to the calculation of the optimal stepsize which is denoted by μ_{opt} . When the preconditioned conjugate gradient algorithm is considered for example, it implies the algebraical calculation of the following quantity: $\mathcal{C}_{\text{JBD}}(\mathbf{B}^{(m+1)}) = \mathcal{C}_{\text{JBD}}(\mathbf{B}^{(m)} - \mu \mathbf{D}^{(m)})$, where $\mathbf{D}^{(m)} = \text{unvec}(\mathbf{d}^{(m)})$ or equivalently $\mathbf{d}^{(m)} = \text{vec}(\mathbf{D}^{(m)})$ where the vector $\mathbf{d}^{(m)}$ is given in (23) (operator unvec enables to come back to a matrix). It is a 4th-degree polynomial whose expression is given by (to simplify the different expressions, the dependency upon the iteration m is omitted):

$$\mathcal{C}_{\text{JBD}}(\mathbf{B} - \mu \mathbf{D}) = a_0 + a_1\mu + a_2\mu^2 + a_3\mu^3 + a_4\mu^4, \quad (28)$$

where the five coefficients a_0, a_1, a_2, a_3 and a_4 are found equal to (see Appendix A.2 for the details of the calculations),

$$a_0 = \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_0^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_0 \} \right\}, \quad (29)$$

$$a_1 = - \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_1^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_0 \} + \mathbf{C}_0^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_1 \} \right\}, \quad (30)$$

$$a_2 = \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_2^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_0 \} + \mathbf{C}_1^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_1 \} + \mathbf{C}_0^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_2 \} \right\}, \quad (31)$$

$$a_3 = - \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_2^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_1 \} + \mathbf{C}_1^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_2 \} \right\}, \quad (32)$$

$$a_4 = \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_2^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_2 \} \right\}. \quad (33)$$

with

$$\mathbf{C}_0 = \mathbf{B} \mathbf{M}_i \mathbf{B}^H, \quad (34)$$

$$\mathbf{C}_1 = \mathbf{B} \mathbf{M}_i \mathbf{D}^H + \mathbf{D} \mathbf{M}_i \mathbf{B}^H, \quad (35)$$

$$\mathbf{C}_2 = \mathbf{D} \mathbf{M}_i \mathbf{D}^H. \quad (36)$$

Then, the third order polynomial (37) corresponding to the derivative with respect to μ of the 4th-degree polynomial given in (28) is calculated and its three roots are derived

$$\frac{\partial \mathcal{C}_{\text{JBD}}(\mathbf{B} - \mu \mathbf{D})}{\partial \mu} = 4a_4\mu^3 + 3a_3\mu^2 + 2a_2\mu + a_1. \quad (37)$$

Finally the optimal stepsize, μ_{opt} , corresponds to the root that leads to the smallest value of (28).

4.4. Algorithms

The non-unitary JBD algorithms (denoted by JBD_{EH}) are summarized below:

Data: N_m square matrices $\mathbf{M}_1, \mathbf{M}_2, \dots, \mathbf{M}_{N_m}$, stopping criterion γ , step-size μ , max. number of iterations N_{max} , number used for restart N_0

Result: Estimation of the joint block diagonalizer $\mathbf{B}$

initialize: $\mathbf{B}^{(0)}; \mathbf{D}^{(0)}; m = 0;$

repeat

if (for (P)CG algorithms only) $m \bmod N_0 = 0$ **then**

 restart

else

 Calculate optimal step-size $\mu_{\text{opt}}^{(m)}$ (root of (37) leading to the smallest value of (28))

 Compute gradient matrix $\nabla_{\mathbf{B}} \mathcal{C}_{\text{JBD}}(\mathbf{B}^{(m+1)}, (\mathbf{B}^{(m+1)})^*)$ given in (10)

 Compute exact preconditioner $\mathbf{P}^{(m+1)}$ given in (18) and (19).

 Compute matrix $\mathbf{B}^{(m+1)}$ thanks to (23)

 For (P)CG algorithms only, compute $\beta_{\text{FR}}^{(m)}$ or $\beta_{\text{PR}}^{(m)}$ given in (24) or (25)

 Compute the search direction $\mathbf{D}^{(m+1)}$ thanks to (23)

$m = m + 1;$

end

until ($(\|\mathbf{B}^{(m+1)} - \mathbf{B}^{(m)}\|_F^2 \leq \gamma)$ or $(m \geq N_{\text{max}})$);

4.5. Algorithmic complexity

To compute the gradient matrix $\nabla_{\mathbf{B}} \mathcal{C}_{\text{JBD}}(\mathbf{B}^{(m+1)}, (\mathbf{B}^{(m+1)})^*)$ whose expression is given in (10), the algorithmic complexity C_1 is:

$M \neq N$	$4N_m N M(M+N) + 2N^2 N_m$
$M \gg N$	$4N_m N M^2$
$M = N$	$8N_m N^3 + 2N_m N^2$

To compute the preconditioning matrix according to $\mathbf{P} = \text{Diag}\{\mathbf{H}_{\mathbf{B}, \mathbf{B}^*}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*))\}$, the complexity C_2 is:

$M \neq N$	$2N_m(MN(4M+3N) + MN^4(M+N))$
$M \gg N$	$8N_m M^2 N + 2N_m M^2 N^4$
$M = N$	$8N_m N^3 + 2N^6$

To update the search direction given in (23), using the Polak-Ribière formula $\beta^{(m)}$ given in (25), the complexity is:

$M \neq N$	$2M^3 N^3 + 4N_m(MN(4M+3N) + 2MN^4(M+N)) + 2NM(M+N)$
$M \gg N$	$2M^3 N^3 + 16N_m M^2 N + 8N_m M^2 N^4 + 2NM^2$
$M = N$	$8N^6 + 28N_m N^3 + 4N^4$

To compute the coefficients $a_0^{(m)}, \dots, a_4^{(m)}$ thanks to (29)–(33), and then to obtain the optimal step-size $\mu_{\text{opt}}^{(m)}$ by the search of the root of the polynomial given in (37) attaining the minimum in the polynomial given in (28), the complexity is:

$M \neq N$	$24MNN_m(M+N) + 9N^2 N_m(1+N)$
$M \gg N$	$24M^2 N N_m$
$M = N$	$57N^3 N_m + 9N^2 N_m$

Finally, the total complexity for the different algorithms that we were able to derive is summed up in the following table:

Algorithm	Algorithmic cost	
JBD _{CG}	$M \neq N$	$28N_m N M(M+N) + 11N^2 N_m + 2N^3 M^3 + NM(M+N) + 9N^3 N_m$
	$M \gg N$	$11N^2 N_m + 28N_m N M^2 + 2N^3 M^3 + 9N^3 N_m + NM^2$
	$M = N$	$17N_m N^3 + 11N^2 N_m + 2N^6 + 2N^3 + 48N_m N^3$
JBD _{PCC_{EH}}	$M \neq N$	$2N_m N M(20M + 17N + 5N^3(M+N)) + 2N^3 M^3 + 19N^2 N_m + 9N^3 N_m$
	$M \gg N$	$40N_m N M^2 + 10M^2 N^4 + 2N^3 M^3 + 19N^2 N_m + 9N^3 N_m$
	$M = N$	$83N_m N^3 + 20N^4 + 2N^6 + 19N^2 N_m$
JBD _{LM_{EH}}	$M \neq N$	$2N_m N M(18M + 17N + N^3(M+N)) + 11N^2 N_m + 2N^3 M^3 + 9N^3 N_m$
	$M \gg N$	$36N_m M^2 N + 11N^2 N_m + 2N_m M^2 N^4 + 2N^3 M^3 + 9N^3 N_m$
	$M = N$	$88N_m N^3 + 4N_m N^6 + 2N^6 + 2N^2 N_m$
JBD _{QN_{EH}}	$M \neq N$	$2N_m N M(18M + 17N) + 11N^2 N_m + 2N_m M N^4(M+N) + 9N^3 N_m$
	$M \gg N$	$36N_m N M^2 + 11N^2 N_m + 2N_m M^2 N^4 + 9N^3 N_m$
	$M = N$	$69N_m N^3 + 11N^2 N_m + 4N_m N^6$

Finally, notice that the global complexity of the different algorithms has to be multiplied by the total number of iterations N_i needed to reach the convergence. In the case of practical applications, the computational time necessary to build the set of the N_m matrices should be counted too.

5. Computer simulations & comparisons

In this section, we present simulations to illustrate the behavior of the suggested preconditioned JBD algorithms ($\text{JBD}_{\text{PCGEH}}$, JBD_{LMEH} and JBD_{QNEH} but also of the JBD_{CG}). All these algorithms are compared with the conjugate gradient $\text{JBD}_{\text{CG}_{\text{NION}}}$ proposed by D. Nion [41] and another algorithm proposed by H. Ghennioui [26] based on a relative gradient approach denoted $\text{JBD}_{\text{RGrad}}$. To that aim, we build a set $\mathcal{D}$ of N_m complex block-diagonal matrices $\mathbf{D}_i$ (for all $i = 1, \dots, N_m$), randomly drawn according to a Gaussian law with zero mean and unit variance. Initially, these matrices are considered as exactly block-diagonal, then, a random noise matrix of mean 0 and variance σ_N^2 is added on the off-diagonal blocks of the matrices $\mathbf{D}_i$, for all $i = 1, \dots, N_m$. The signal to noise ratio is defined as $\text{SNR} = 10 \log(\frac{1}{\sigma_N^2})$. This study includes several cases. Firstly, all algorithms are initialized using the same initialization provided by the generalized eigenvalue decomposition (GEVD) as suggested in [41]. Then we go further from the solution while these algorithms are initialized randomly to show the behavior of these algorithms. It should also be mentioned that the case of rectangular matrix $\mathbf{A}$ (or $\mathbf{B}$) is considered too with γ is fixed to 1×10^{-13} in all simulations. To better assess the quality of the estimation and to be able to compare the different algorithms, an error index is required. We are using the one, $I(\mathbf{G})$, that we introduced in [25]. It reads:

$$\frac{1}{r(r-1)} \left[\sum_{i=1}^r \left(\sum_{j=1}^r \frac{\|(\mathbf{G})_{i,j}\|_F^2}{\max_{\ell} \|(\mathbf{G})_{i,\ell}\|_F^2} - 1 \right) + \sum_{j=1}^r \left(\sum_{i=1}^r \frac{\|(\mathbf{G})_{i,j}\|_F^2}{\max_{\ell} \|(\mathbf{G})_{\ell,j}\|_F^2} - 1 \right) \right],$$

where $(\mathbf{G})_{i,j}$ for all $i, j \in \{1, \dots, r\}$ is the (i, j) -th (square) block matrix of $\mathbf{G} = \hat{\mathbf{B}}\mathbf{A}$. The best results are obtained when the error index $I(\cdot)$ is found to be close to 0 in linear scale ($-\infty$ in logarithmic scale). Regarding to the charts, $I(\cdot)$ is given in dB and is then defined by $I(\cdot) \text{ dB} = 10 \log(I(\cdot))$. The matrix $\mathbf{A}$ has been randomly chosen in all simulations. Moreover, all displayed results have been averaged over 30 Monte-Carlo trials.

We begin with $M = N = 8$ ($\mathbf{A}$ is square) and $r = 4$ (with $r = 4$ for all $j = 1, \dots, 4$). The set $\mathcal{M}$ of N_m square matrices $\mathbf{M}_i, i = 1, \dots, N_m$ consists of $N_m = 100$ (resp. $N_m = 5$) $M \times M$ complex matrices. All algorithms are initialized using the initialization provided by the GEVD as suggested in [41].

In the Fig. 1, the evolution of the error index versus the number of iterations in the noiseless case ($\text{SNR} = 100$ dB) is displayed. In addition, two cases are considered: i) on the left $N_m=100$ ii) and on the right $N_m = 5$ for the considered matrices set. In the Fig. 2 we illustrate the evolution of the error index versus the number of iterations in the noisy case ($\text{SNR} = 30$ dB). Again, the left chart consists a set of $N_m = 100$ matrices while the right chart consider $N_m = 5$ matrices only. Notice that the maximal number of iterations allowed is set at $N_{\text{max}} = 120$ (the default value when GEVD is used).

From these curves, several observations can be drawn. First we observe that the convergence of algorithms based on an exact preconditioning namely $\text{JBD}_{\text{PCGEH}}$, JBD_{LMEH} and JBD_{QNEH} are the fastest algorithms. Moreover, comparing the convergence of different gradient approaches namely $\text{JBD}_{\text{CG}_{\text{NION}}}$ and $\text{JBD}_{\text{RGrad}}$, we check the interest of our proposed preconditioner based on the exact Hessian matrices. Furthermore, we deduce that the convergence speed of the algorithm JBD_{CG} is higher than $\text{JBD}_{\text{RGrad}}$ and especially the algorithm $\text{JBD}_{\text{CG}_{\text{NION}}}$ which follows the same adaptation rule but

operates on a different cost function. However, the best performances are generally achieved using the $\text{JBD}_{\text{PCGEH}}$ algorithm in difficult contexts : with noise and a set containing a small number of matrices to be joint block diagonalized. In addition, we note that $\text{JBD}_{\text{PCGEH}}$ and JBD_{QNEH} converge to the same solution approximately. However, the others require more iterations to reach the same performances.

[Figure 1 about here.]

We study, now, the behavior of these algorithms by initializing with a random matrix (away from the solution). For this reason, we trace in the Figures 3 and 4 the evolution of the error index in dB versus the number of iterations in the noiseless and noisy cases ($\text{SNR} = 100$ and 30 dB). Two cases are studied: on the left the set to be joint block diagonalized consists of $N_m = 100$ matrices, while on the right we have $N_m = 10$ matrices only.

These charts draw almost the same conclusions as those presented in the GEVD initialization context especially in a noisy context. We find, logically, all algorithms take more iterations to converge due to the initial point which starts far from the solution ($\simeq 0$). This why set the maximal number of iterations at $N_{\text{max}} = 200$ (the default value when we start far from the solution).

[Figure 2 about here.]

Secondly, we consider now the $N \times M$ where $M = 12$, $N = 8$ ($\mathbf{A}$ is rectangular) and the block size is $r = 4$. In the Fig. 5 (resp. Fig. 6), we compare $\text{JBD}_{\text{PCGEH}}$, JBD_{LMEH} , JBD_{QNEH} , $\text{JBD}_{\text{CGNION}}$ and $\text{JBD}_{\text{RGrad}}$ algorithms using GEVD initialization. To that aim, we display again the error index $I(G)$ in dB versus the number of iterations in the noiseless case ($\text{SNR} = 100$ dB) (resp. noisy case $\text{SNR} = 30$ dB). The left chart consists a set of $N_m = 100$ matrices while the right chart works on $N_m = 5$ matrices.

[Figure 3 about here.]

These charts verify that the algorithms ($\text{JBD}_{\text{PCGEH}}$, JBD_{LMEH} and JBD_{QNEH}) based on exact preconditioner lead to better performances. We check again that the preconditioned conjugate gradient algorithm ($\text{JBD}_{\text{PCGEH}}$) is more robust in difficult contexts (noise and/or very few matrices to be joint block diagonalized). We observe also a gain in performance and stability in our proposed conjugate gradient algorithm (JBD_{CG}) comparing with Nion's one ($\text{JBD}_{\text{CGNION}}$) and which illustrates a bad behavior in the noisy context. As a conclusion, the performance depends on the size N_m of the set of matrices (increasing N_m , we improve the performances of proposed algorithms).

Then, we emphasize the influence of i) the number N_m of matrices to be block diagonalized, ii) and SNR. We start first with the number N_m of matrices to be joint block diagonalized, then we finish with the SNR at a fixed number of block ($r = 4$). All algorithms are initialized using GEVD. In the Fig. 7 (resp. Fig. 8), we plot the error index $I(G)$ in dB versus N_m (resp. versus SNR) for $\text{SNR} = 25$ dB (resp. for $N_m = 5$) on the left and for $\text{SNR} = 100$ dB (resp. for $N_m = 100$) on the right.

[Figure 4 about here.]

We observe that the more matrices to be joint block-diagonalized we have, the best the obtained results are. In addition, whatever the considered algorithm, they all reach the same level of performance (which not too surprising either), yet, the $\text{JBD}_{\text{CGNION}}$

and $\text{JBD}_{\text{RGrad}}$ algorithms need more iterations. Even though when smaller subsets of matrices and the SNR value are considered, the results remain relatively good. Not too surprisingly, the noise affects the performances.

In the Figures 9 and 10, we study the behavior of the proposed NU – JBD algorithms in the restricted context of the non unitary joint-diagonalization (NU – JD). The different algorithms are denoted by JD_{PCGEH} , JD_{CG} , JD_{QNEH} , JD_{LMEH} , JD_{RGrad} and $\text{JD}_{\text{CGNION}}$. In the Fig. 9, we give the error index $I(\mathbf{G})$ versus the size of the matrix set to be joint-diagonalized for different values of the Signal to Noise Ratio (SNR = 30 dB and 100 dB). In the Fig. 10, we display the evolution of the error index $I(\mathbf{G})$ versus the SNR for $N_m = 5$ and 100 matrices to be joint-diagonalized. From these figures, we could actually say that JD_{PCGEH} , JD_{QNEH} , JD_{LMEH} and JD_{CG} constitute four new solutions to the NU – JD problem.

[Figure 5 about here.]

Finally, in the Fig. 11, we display the evolution of the error index versus the number of blocks r that are considered for a fixed size of the matrix sets ($N_m = 100$) and for two different values of the Signal to Noise Ratio (SNR = 30 dB and 100 dB). Again, the best performances are obtained with algorithms using the proposed preconditioner namely $\text{JBD}_{\text{PCGEH}}$, JBD_{QNEH} and JBD_{LMEH} (it is particularly true when the number of blocks r is equal to 1). We find that the performance of the BDC proposed algorithms depends on the size of a block considered r or the number of blocks $n_j = \frac{N}{r}$ ¹. The performances are better when decreasing the size of a block r . We observe also a decrease of the performances when the SNR decreases, even though the results remain relatively good. Note again that in the particular case where $r = 1$, we get the joint diagonalization problem.

[Figure 6 about here.]

6. Application of blind separation of convolutive mixtures of non-stationary sources using SQTFD

In the BSS problem, the goal is to recover multiple sources mixed through an unknown mixing system from the system outputs only (namely the observations). In this section, our aim is to provide an application example of the proposed algorithms showing how they enable to address the problem of the separation of convolutive FIR mixtures of nonstationary sources. We focus on BSS methods that are based on spatial time-frequency distributions or spectra applied to deterministic or stochastic, correlated or uncorrelated, stationary, cyclo-stationary or non-stationary signals. This kind of methods were first introduced in [4] in an instantaneous/whitened mixture of deterministic source signals and same authors were the first to bridge time-frequency to array processing and BSS. It is well known that this kind of BSS methods generally require four main steps to guarantee successful signal separations.

- The convolutive mixture is rewritten as an instantaneous one.
- Regions of signal power concentration and localization should be properly determined and time-frequency points of peaky values must be selected.

¹ In BSS, the number of blocks n_j represents the number of sources n .

- The set of time-frequency matrices of the observations across the array at these time-frequency points must be constructed.
- The mixing matrix should be properly estimated so as to undo the mixing of signals at the multi-sensor receiver.

6.1. Step 1: from a convolutive mixture to an instantaneous one

We assume that we have m ($m \in \mathbb{N}^*$) observation signals $x_i(t)$, $i = 1, \dots, m$, $t \in \mathbb{Z}$, and that the n ($n \in \mathbb{N}^*$) sources $s_j(t)$, for all $j = 1, \dots, n$ are mixed through a linear Finite Impulse Response (FIR) multichannel system, denoted by $\mathbf{H}(t) = (H_{ij}(t))$. Thus, the considered system is described by the following input-output relation:

$$x_i(t) = \sum_{j=1}^n \sum_{\ell=0}^L H_{ij}(\ell) s_j(t-\ell) + n_j(t), \quad \forall i = 1, \dots, m \quad (38)$$

where $H_{ij}(t)$ is the impulse response function between the j -th source and the i -th sensor with an overall extent of $L+1$ taps and $n_i(t)$, for all $i = 1, \dots, m$ are noises. Denoting by $M = mL'$, $N = n(L+L') = nQ$ (with $Q = L+L'$ and $L' \in \mathbb{N}^*$), we consider the $(N \times 1)$ vector $\mathbf{S}(t) = [\mathbf{s}_1(t)^T, \mathbf{s}_2(t)^T, \dots, \mathbf{s}_n(t)^T]^T$ and the $(M \times 1)$ vectors $\mathbf{X}(t) = [\mathbf{x}_1(t)^T, \mathbf{x}_2(t)^T, \dots, \mathbf{x}_m(t)^T]^T$ and $\mathbf{N}(t) = [\mathbf{n}_1(t)^T, \mathbf{n}_2(t)^T, \dots, \mathbf{n}_m(t)^T]^T$, where the $(Q \times 1)$ vectors $\mathbf{s}_i(t)$ (for all $i = 1, \dots, n$) read $\mathbf{s}_i(t) = [s_i(t), s_i(t-1), \dots, s_i(t-Q+1)]^T$, the $(L' \times 1)$ vectors $\mathbf{x}_i(t)$ (for all $i = 1, \dots, m$) read $\mathbf{x}_i(t) = [x_i(t), x_i(t-1), \dots, x_i(t-L'+1)]^T$ and the $(L' \times 1)$ vectors $\mathbf{n}_i(t)$ (for all $i = 1, \dots, m$) read $\mathbf{n}_i(t) = [n_i(t), n_i(t-1), \dots, n_i(t-L'+1)]^T$. The model described by Eq. (38) can be rewritten in the ensuing matrix form:

$$\mathbf{X}(t) = \mathbf{A}\mathbf{S}(t) + \mathbf{N}(t), \quad (39)$$

where the $(M \times N)$ mixing matrix $\mathbf{A}$ is a block-matrix defined as $\mathbf{A} = (\mathbf{A}_{ij})$ for all $i = 1, \dots, m$ and $j = 1, \dots, n$ whose blocks $\mathbf{A}_{ij}$ are $(L' \times Q)$ Toeplitz matrices:

$$\mathbf{A}_{ij} = \begin{pmatrix} H_{ij}(0) & \dots & H_{ij}(L) & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & H_{ij}(0) & \dots & H_{ij}(L) \end{pmatrix}. \quad (40)$$

To still deal with an over-determined model, L' is chosen such that $M \geq N$. It is further assumed that:

Assumption A: The noises $n_j(t)$ for all $j = 1, \dots, m$ are stationary, white, zero-mean random signals, mutually uncorrelated and independent from the sources.

Assumption B: For each of the n sources s_i , there exist time-frequency points where only this source and its delayed versions are present in the time-frequency plane.

Assumption C: The number of sensors m and the number of sources n are both known and $m \geq n$ (i.e. the over-determined case is considered). The FIR filter order L is known too.

Finally, it is important to notice that the assumption B is a key assumption since it replaces the classical “random independent sources” assumption in the one side. It makes it possible to tackle the nonstationary correlated sources separation problem in the other side.

6.2. Step 2: selection of “useful” time-frequency points

The Spatial Quadratic Time-Frequency Spectrum (SQTFS)² of the observations across the array at a given t - f point is a $(M \times M)$ matrix, admits the following decomposition (since the noises are centered and independent from the source signals (Assumption A)):

$$\begin{aligned}\mathbf{D}_{\mathbf{X}}(t, \nu) &= \mathbf{A}\mathbf{D}_{\mathbf{S}}(t, \nu)\mathbf{A}^H + \mathbf{D}_{\mathbf{N}}(t, \nu) + \mathbf{A}\mathbf{D}_{\mathbf{SN}}(t, \nu) + \mathbf{D}_{\mathbf{NS}}(t, \nu)\mathbf{A}^H \\ &= \mathbf{A}\mathbf{D}_{\mathbf{S}}(t, \nu)\mathbf{A}^H + \mathbf{D}_{\mathbf{N}}(t, \nu),\end{aligned}\quad (44)$$

where, respectively, $\mathbf{D}_{\mathbf{S}}(t, \nu)$ represents the $(N \times N)$ source SQTFS, $\mathbf{D}_{\mathbf{N}}(t, \nu)$ is the $(M \times M)$ noise SQTFS and finally $\mathbf{D}_{\mathbf{SN}}(t, \nu)$ and $\mathbf{D}_{\mathbf{NS}}(t, \nu)$ are the Spatial Bilinear Time-Frequency Spectra (SBTFS) between the sources and the noises.

To reduce the noise, as mentinned in [26], the weak energy t - f points are eliminated. Only those such that: $\min(|\mathbf{D}_{x_i x_j}(t, \nu)|^2) > \epsilon$ $i, j \in \{1, \dots, M\}$, are kept, where ϵ is a (sufficiently) small constant (for example $\epsilon \approx 0.01$), $\min(\cdot)$ is the minimum value and $\mathbf{D}_{x_i x_j}(t, \nu) = (\mathbf{D}_{\mathbf{X}}(t, \nu))_{ij}$. Thus, Eq. (44) finally reads

$$\mathbf{D}_{\mathbf{X}}(t, \nu) \simeq \mathbf{A}\mathbf{D}_{\mathbf{S}}(t, \nu)\mathbf{A}^H. \quad (45)$$

When t - f points (t, ν) satisfying Assumption B are considered, the matrix $\mathbf{D}_{\mathbf{S}}(t, \nu)$ possesses a very specific algebraic structure since it is block-diagonal with one single non null $(Q \times Q)$ block on the block-diagonal (the same property holds in the pre-whitened case and has already been used by the authors of [24]). This algebraic property is general since in the case of matrices stemming from spatial time-frequency distributions or spectra (SQTFD(or S)), the block-diagonal matrix with one single non null block in the block-diagonal is the only possibility of block-diagonal matrix.

Nonstationary signal separation techniques based on SQTFD(or S) require a preliminary stage of selecting, automatically or manually, particular t - f points: those whose existence is ensured from Assumption B. This step is important to construct the set of matrices denoted $\mathcal{M}_{\text{JBD}}$ for joint block-diagonalization. All the matrices belonging

²We recall that considering two vectors, namely $\mathbf{X}(t)$ and $\mathbf{Y}(t)$, the SBTFS is defined as

$$\mathbf{D}_{\mathbf{XY}}(t, \nu) = \int_{\mathbb{R}^2} \mathbf{R}_{\mathbf{XY}}(\theta, \theta') \mathbf{K}(\theta, \theta'; t, \nu) d\theta d\theta', \quad (41)$$

where the function $\mathbf{K}(\theta, \theta'; t, \nu)$ is the kernel of the transformation and $\mathbf{R}_{\mathbf{XY}}(t, \tau)$ is the cross-correlation matrix of signals $\mathbf{X}(t)$ and $\mathbf{Y}(t)$ i.e. $\mathbf{R}_{\mathbf{XY}}(t, \tau) = \mathbf{E}\{\mathbf{X}(t + \frac{\tau}{2})\mathbf{Y}^H(t - \frac{\tau}{2})\}$ if $\mathbf{E}\{\cdot\}$ is the mathematical expectation operator. The SQTFS is a case of the SBTFS with $\mathbf{Y}(t) = \mathbf{X}(t)$:

$$\mathbf{D}_{\mathbf{X}}(t, \nu) \stackrel{\text{def}}{=} \mathbf{D}_{\mathbf{XX}}(t, \nu) = \int_{\mathbb{R}^2} \mathbf{R}_{\mathbf{X}}(\theta, \theta') \mathbf{K}(\theta, \theta'; t, \nu) d\theta d\theta', \quad (42)$$

where $\mathbf{R}_{\mathbf{X}}(\theta, \theta') \stackrel{\text{def}}{=} \mathbf{R}_{\mathbf{XX}}(\theta, \theta')$ is the correlation matrix of signal $\mathbf{X}(t)$. To illustrate, consider two important examples: first the Spatial Pseudo Wigner Spectrum (SPWS), whose expression is given by:

$$\mathbf{D}_{\text{SPWS}, \mathbf{X}}(t, \nu) = \int_{\mathbb{R}} \mathbf{R}_{\mathbf{X}}(t, \tau) h(\tau) \exp(-2i\pi\nu\tau) d\tau, \quad (43)$$

where $\nu^2 = -1$, $h(\cdot)$ stands for the smoothing (short-time) window (rectangular, Kaiser-Bessel, Hamming, Bartlett, etc). When deterministic instead of random signals are considered, one simply has to discard the mathematical expectation in Eq. (43) and the Spatial Pseudo-Wigner Spectrum (SPWD), $\mathbf{D}_{\text{SPWD}, \mathbf{X}}(t, \nu)$ is recovered. It has to be noted too that the Spatial Wigner Spectrum (SWS), $\mathbf{D}_{\text{SWS}, \mathbf{X}}(t, \nu)$, is obtained by considering $h(\tau) = 1$ for all τ in Eq. (43), while the Spatial Wigner distribution (SWD), $\mathbf{D}_{\text{SWD}, \mathbf{X}}(t, \nu)$, is recovered by discarding the mathematical expectation.

to $\mathcal{M}_{\text{JBD}}$ admit the decomposition given by Eq. (45) (or equally Eq. (1)) where the source SQTFD(or S) matrices $\mathbf{D}_{\mathbf{S}}(t, \nu)$ assume a specific algebraic structure, being block diagonal matrices. As a consequence, a natural manner to tackle the separation problem will consist of opting for a JBD algorithm (see the Step 4 detailed Section 6.4).

6.3. Step 3: construction of $\mathcal{M}_{\text{JBD}}$ (set of t - f matrices of the observations across the array at the chosen t - f points)

In order to select the t - f matrices whose rank is equal to Q , we follow the same procedures for the detector proposed in [26] denoted $\mathbf{C}_{\text{ConvGH}}$. In fact, it can be seen either as a generalization of the detector proposed in [28] in the instantaneous context for non-whitened signals (it was based on a rank one property) or as a generalization to non-whitened signals of the detector suggested in [24] for the convolutive case for whitened observations. However, the idea of this detector is as follows :

- Due to previous assumptions, $\mathbf{D}_{\mathbf{X}}(t, \nu)$ and $\mathbf{D}_{\mathbf{S}}(t, \nu)$ have the same rank.
- A singular value decomposition is used (SVD) leading to $\mathbf{D}_{\mathbf{X}}(t, \nu) = \mathbf{U}(t, \nu) \Delta(t, \nu) \mathbf{V}(t, \nu)^H$ with $\mathbf{V}(t, \nu)$ and $\mathbf{U}(t, \nu)$ are $(M \times M)$ unitary matrices and $\Delta(t, \nu) = \text{diag}\{\delta(t, \nu)\}$ is a diagonal matrix whose diagonal elements are positive.
- Denoting by $\delta(t, \nu) = (\delta_1(t, \nu), \dots, \delta_M(t, \nu))^T$ and assuming that the singular values are sorted in a decreasing order: $\delta_1(t, \nu) \geq \delta_2(t, \nu) \geq \dots \geq \delta_M(t, \nu) \geq 0$.
- A way to check whether a matrix is rank Q or not is required and given by: for $\mathcal{M}_{\text{JBD}}$ choose SQTFD(or S) corresponding to t - f points (t, ν) such that:

$$\mathcal{R}(t, \nu) = \frac{\sum_{i=1}^Q \delta_i^2(t, \nu)}{\|\mathbf{D}_{\mathbf{X}}(t, \nu)\|_F^2} \geq 1 - \epsilon \quad (46)$$

where ϵ is a (sufficiently) small positive constant.

6.4. Step 4: non-whitened based separation algorithm

The matrices belonging to the set $\mathcal{M}_{\text{JBD}}$ (whose size is denoted by N_m ($N_m \in \mathbb{N}^*$)) all admit a particular structure since they can be decomposed into $\mathbf{A} \mathbf{D}_{\mathbf{S}}(t, \nu) \mathbf{A}^H$ with $\mathbf{D}_{\mathbf{S}}(t, \nu)$ a block-diagonal matrix with only one non null ($Q \times Q$) block on its block-diagonal. One possible way to recover the mixing matrix $\mathbf{A}$ (or its pseudo-inverse: the separation matrix $\mathbf{B}$) is to directly joint block diagonalize the matrix set $\mathcal{M}_{\text{JBD}}$. However, the sources recovered after inversion of the system are obtained up to a permutation and up to a filter which are the classical indeterminations of the BSS in the convolutive case.

Four BSS methods can be derived: the first one is called $\text{JBD}_{\text{PCGEH-TF}}$ since it combines the JBD algorithm based on a preconditioned conjugate gradient approach $\text{JBD}_{\text{PCGEH}}$ together with the automatic time-frequency points detector $\mathbf{C}_{\text{ConvGH}}$. The second method called $\text{JBD}_{\text{QNEH-TF}}$ consists of replacing the preconditioned conjugate gradient-based JBD algorithm by the Quasi-Newton algorithm JBD_{QNEH} . The third method called $\text{JBD}_{\text{LMEH-TF}}$ consists in replacing by the Levenberg-Marquardt-based JBD algorithm ($\text{JBD}_{\text{LMEH-TF}}$) and the fourth is called JBD_{CGTF} in which the conjugate gradient without preconditioning is used (JBD_{CG}).

6.5. Computer simulations

In this section, computer simulations are performed to illustrate the good performance of the four suggested methods and to compare them again with the same kind of existing approaches: The first one is denoted by $\text{JBD}_{\text{CGNION-TF}}$ and combines the non-unitary JBD algorithm proposed in [41] with the t - f point detector C_{ConvGH} . The other algorithm denoted $\text{JBD}_{\text{RGrad-TF}}$ combine the non-unitary JBD algorithm proposed in [26] with the same detector.

We consider $m = 4$ mixtures of $n = 2$ sources of 128 time samples. The first source (resp. the second source) is a sinusoidal frequency modulation (resp. a linear frequency modulation), $L = 3$ and $L' = 2$. These sources are mixed according to a mixture matrix $\mathbf{A}(t)$ whose components were randomly generated and whose z -transform $\mathbf{A}[z]$ is given by:

$$\begin{pmatrix} -0.0134 + 0.2221z^{-1} + 0.9749z^{-2} & -0.6579 + 0.7521z^{-1} - 0.0382z^{-2} \\ -0.7425 - 0.6639z^{-1} - 0.0891z^{-2} & 0.7287 + 0.6049z^{-1} - 0.3210z^{-2} \\ 0.6173 + 0.3921z^{-1} - 0.6820z^{-2} & 0.9477 - 0.2211z^{-1} - 0.2301z^{-2} \\ 0.7119 - 0.6532z^{-1} - 0.2578z^{-2} & 0.2533 - 0.7511z^{-1} - 0.6096z^{-2} \end{pmatrix},$$

We use the SPWVS with a Hamming smoothing window of size 128 and 64 frequency bins. In this example, 30 realizations of signals are computed. Then, the Spatial Pseudo-Wigner distribution of each realization is calculated. Finally, all 30 resulting SPWD have been averaged. Moreover, we have chosen $\epsilon = 0.0050$ for the detector C_{ConvGH} .

In the Figure 13, we have plotted the error index $I(G)$ versus the SNR (≈ 1410 time-frequency matrices were selected when $\text{SNR} \geq 20$ dB). However, the more the SNR decreases (for $\text{SNR} \leq 20$ dB), the more this number decreases too (see the Figure 12). For the noises, we use random entries chosen from a Gaussian distribution with zero mean and variance σ_N^2 and the same definition of the SNR as mentioned in Section 5.

[Figure 7 about here.]

The resulting error index shows that the four proposed methods show the best performances. We observe again that the performances of proposed algorithms based on an exact preconditioning namely $\text{JBD}_{\text{PCGEH-TF}}$, $\text{JBD}_{\text{QNEH-TF}}$ and $\text{JBD}_{\text{LMEH-TF}}$ are the fastest algorithms. The JBD_{CGTF} method still performs better than $\text{JBD}_{\text{CGNION-TF}}$ based on conjugate gradient and which exhibits the worst performances especially in the noisy context. However, the $\text{JBD}_{\text{PCGEH-TF}}$ and $\text{JBD}_{\text{QNEH-TF}}$ methods that provide the best results in the same context. Finally, all t - f methods reach the same performances especially when $\text{SNR} \geq 40$ dB except, logically, $\text{JBD}_{\text{CGNION-TF}}$ which operates on a different cost function.

6.6. Discussion & conclusion

In this article, we are interested in non unitary joint block diagonalization algorithms which are of great interest in different fields of application among which are array processing for wideband signals and blind sources separation when convolutive mixtures are considered. Our aim, here, was to investigate the relevance/interest of a preconditioning for those algorithms. Such a procedure usually requires either the calculation of the exact complex Hessian matrix at each iteration. In this article, the exact calculation has been performed. Then, different iterative preconditioned joint block diagonalization algorithms have been derived and compared. For instance, we have focused

our attention on Quasi-Newton, Preconditioned Conjugate Gradient and Levenberg-Marquardt algorithms based on the use of the diagonal of the true Hessian matrix. Computer simulations performed in difficult contexts (noise and/or very few matrices to be joint block diagonalized) have emphasized the good behavior of the preconditioned conjugate gradient algorithm (based on the use of the diagonal of the true Hessian matrix) and denoted by $\text{JBD}_{\text{PCG-EH}}$. This algorithm offers the best compromise between good performance and fast convergence speed even in complicated scenarios. The main advantage of the approach suggested here remains its really general aspect. In fact, it does not rely on restrictive assumptions neither about the considered matrix set (they only have to be complex but not necessarily hermitian matrices) nor about the joint block-diagonalizer (it is not necessarily an unitary matrix). Besides, since the joint block-diagonalization problem encompasses that of joint diagonalization, new JD algorithms can also be derived at the same time. The usefulness of the proposed JBD algorithms is emphasized in the BSS context. Combined with a t - f points detector C_{ConvGH} operating in a non-whitened context, these algorithms allow to obtain four alternative time-frequency based BSS methods (denoted by $\text{JBD}_{\text{PCGEH-TF}}$, $\text{JBD}_{\text{QNEH-TF}}$, $\text{JBD}_{\text{LMEH-TF}}$ and JBD_{CGTF}) devoted to the separation of deterministic signals or mixtures of (correlated) nonstationary stochastic processes.

Appendix A

We consider three square matrices, $\mathbf{D}_1$, $\mathbf{D}_2$ and $\mathbf{D}_3$ in $\mathbb{C}^{M \times M}$ and five rectangular matrices $\mathbf{D}_4$, $\mathbf{D}_5$, $\mathbf{D}_6$ in $\mathbb{C}^{N \times M}$, $\mathbf{D}_7$ in $\mathbb{C}^{M \times N}$ and $\mathbf{D}_8$ in $\mathbb{C}^{F \times G}$. Let $\text{tr}\{\cdot\}$, $d(\cdot)$ (or sometimes simply d), $\text{vec}(\cdot)$, $\text{OffBdiag}_{(n)}\{\cdot\}$ and $\mathbf{T}_{\text{Boff}}$ respectively denote the trace operator of a square matrix, the differential operator, the vectorization operator that stacks the columns of a matrix in a long column vector, the zero-block-diagonal operator defined in (5) and the $N^2 \times N^2$ “transformation” matrix defined before. Our developments are based on the following properties (they can be found in [9][26][29][31][34] or [37]):

- P_0 . $\|\text{OffBdiag}_{(n)}\{\mathbf{D}_1\}\|_F^2 = \text{tr}\{\mathbf{D}_1^H \text{OffBdiag}_{(n)}\{\mathbf{D}_1\}\}$.
- P_1 . $(\mathbf{D}_4 + \mathbf{D}_5)^T = \mathbf{D}_4^T + \mathbf{D}_5^T$.
- P_2 . $(\mathbf{D}_4 \mathbf{D}_7)^T = \mathbf{D}_7^T \mathbf{D}_4^T$.
- P_3 . $(\mathbf{D}_4 \mathbf{D}_7)^* = \mathbf{D}_4^* \mathbf{D}_7^*$.
- P_4 . $\text{tr}\{\mathbf{D}_1 + \mathbf{D}_2\} = \text{tr}\{\mathbf{D}_1\} + \text{tr}\{\mathbf{D}_2\}$.
- P_5 . $\text{tr}\{\mathbf{D}_4^H \mathbf{D}_5\} = (\text{vec}(\mathbf{D}_4))^H \text{vec}(\mathbf{D}_5)$.
- P_6 . $d(\text{tr}\{\mathbf{D}_1\}) = \text{tr}\{d(\mathbf{D}_1)\}$.
- P_7 . $\text{vec}(\mathbf{D}_1 \mathbf{D}_2 \mathbf{D}_3) = (\mathbf{D}_3^T \otimes \mathbf{I}_1) \text{vec}(\mathbf{D}_2) \Rightarrow \text{vec}(\mathbf{D}_1 \mathbf{D}_2) = (\mathbf{I}_N \otimes \mathbf{D}_1) \text{vec}(\mathbf{D}_2) = (\mathbf{D}_2^T \otimes \mathbf{I}_N) \text{vec}\{\mathbf{D}_1\}$
- P_{7b} . $\text{vec}(\mathbf{D}_4 \mathbf{D}_1 \mathbf{D}_7) = (\mathbf{D}_7^T \otimes \mathbf{D}_4) \text{vec}(\mathbf{D}_1) \Rightarrow \text{vec}(\mathbf{D}_4 \mathbf{D}_1) = (\mathbf{I}_M \otimes \mathbf{D}_4) \text{vec}(\mathbf{D}_1) = (\mathbf{D}_1^T \otimes \mathbf{I}_N) \text{vec}\{\mathbf{D}_4\}$
- P_8 . $\text{vec}(\mathbf{D}_4^T) = \mathbf{K}_{N,M} \text{vec}(\mathbf{D}_4)$ with $\mathbf{K}_{N,M}$ the unique square $NM \times NM$ commutation matrix defined as: $\mathbf{K}_{N,M} = \sum_n \sum_m \mathbf{E}_{nm} \otimes \mathbf{E}_{nm}^T$ where $\mathbf{E}_{nm}$ are $N \times M$ matrices with a 1 in the (n, m) position and zeros elsewhere: $(\mathbf{E}_{nm}) = (\delta_{nm})$ (δ_{nm} standing for the Kronecker Delta). As a consequence, the permutation matrix $\mathbf{K}_{N,M}$ has one single “1” in each row and in each column.
- P_{8b} . $\mathbf{K}_{N,M} \mathbf{K}_{M,N} = \mathbf{I}_{MN}$ and $\mathbf{K}_{N,M} = \mathbf{K}_{M,N}^{-1} = \mathbf{K}_{M,N}^T$.
- P_9 . $\text{vec}(\text{OffBdiag}\{\mathbf{D}_1\}) = \mathbf{T}_{\text{Boff}} \text{vec}(\mathbf{D}_1)$.

- $\mathbf{P}_{10}.$ $d(\text{vec}(\mathbf{D}_4)) = \text{vec}(d(\mathbf{D}_4)).$
 $\mathbf{P}_{11}.$ $d(\mathbf{D}_4 + \mathbf{D}_5) = d(\mathbf{D}_4) + d(\mathbf{D}_5).$
 $\mathbf{P}_{12}.$ $d(\mathbf{D}_4 \mathbf{D}_7) = d(\mathbf{D}_4) \mathbf{D}_7 + \mathbf{D}_4 d(\mathbf{D}_7).$
 $\mathbf{P}_{13}.$ $d(\mathbf{D}_4^*) = d(\mathbf{D}_4)^*.$
 $\mathbf{P}_{14}.$ $\text{tr}\{\mathbf{D}_1\} = \text{tr}\{\mathbf{D}_1^T\}.$
 $\mathbf{P}_{15}.$ $(\mathbf{D}_4 \otimes \mathbf{D}_8)^T = \mathbf{D}_4^T \otimes \mathbf{D}_8^T.$
 $\mathbf{P}_{16}.$ $(\mathbf{D}_4 \otimes \mathbf{D}_8)^H = \mathbf{D}_4^H \otimes \mathbf{D}_8^H.$
 $\mathbf{P}_{17}.$ $\text{vec}(\mathbf{D}_4 + \mathbf{D}_5) = \text{vec}(\mathbf{D}_4) + \text{vec}(\mathbf{D}_5).$
 $\mathbf{P}_{18}.$ $\mathbf{D}_4^H = (\mathbf{D}_4^*)^T.$
 $\mathbf{P}_{19}.$ $(\mathbf{D}_4 \mathbf{D}_7)^H = \mathbf{D}_7^H \mathbf{D}_4^H.$
 $\mathbf{P}_{20}.$ $(\mathbf{D}_1 \otimes \mathbf{D}_2)(\mathbf{D}_3 \otimes \mathbf{D}_7) = (\mathbf{D}_1 \mathbf{D}_3 \otimes \mathbf{D}_2 \mathbf{D}_7).$
 $\mathbf{P}_{21}.$ $(\mathbf{D}_4 + \mathbf{D}_5)^H = (\mathbf{D}_4^H + \mathbf{D}_5^H).$
 $\mathbf{P}_{22}.$ $\text{OffBdiag}_{(\mathbf{n})}\{\mathbf{D}_1 + \mathbf{D}_2\} = \text{OffBdiag}_{(\mathbf{n})}\{\mathbf{D}_1\} + \text{OffBdiag}_{(\mathbf{n})}\{\mathbf{D}_2\}.$
 $\mathbf{P}_{23}.$ $d\mathbf{D}_4^* = (d\mathbf{D}_4)^* \text{ ([31] p. 53)}.$

A.1- Finding the complex Hessian matrices of the scalar cost function $\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)$

According to [31], the first-order differential of the cost function $\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)$ defined in (4) is given by:

$$d\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*) = \mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)) d\text{vec}(\mathbf{B}) + \mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)) d\text{vec}(\mathbf{B}^*), \quad (47)$$

if the $1 \times NM$ complex vectors $\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*))$ and $\mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*))$ are defined as:

$$\begin{aligned} \mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)) &= \text{vec}^T \left(\frac{\partial \mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}} \right), \\ \mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)) &= \text{vec}^T \left(\frac{\partial \mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}^*} \right). \end{aligned} \quad (48)$$

We also recall that the two partial derivatives that are involved *i.e.* $\frac{\partial \mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}}$ and $\frac{\partial \mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}^*}$ were calculated in [26] using some of the previous properties $\mathbf{P}_0$ - $\mathbf{P}_{20}$. They were found to be equal to:

$$\frac{\partial \mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}} = \sum_{i=1}^{N_m} \left[\left(\text{OffBdiag}_{(\mathbf{n})}\{\mathbf{B} \mathbf{M}_i \mathbf{B}^H\} \right)^T \mathbf{B}^* \mathbf{M}_i^* + \left(\text{OffBdiag}_{(\mathbf{n})}\{\mathbf{B} \mathbf{M}_i \mathbf{B}^H\} \right)^* \mathbf{B}^* \mathbf{M}_i^T \right], \quad (49)$$

$$\begin{aligned} \frac{\partial \mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}^*} &= \sum_{i=1}^{N_m} \left[\text{OffBdiag}_{(\mathbf{n})}\{\mathbf{B} \mathbf{M}_i \mathbf{B}^H\} \mathbf{B} \mathbf{M}_i^H + \left(\text{OffBdiag}_{(\mathbf{n})}\{\mathbf{B} \mathbf{M}_i \mathbf{B}^H\} \right)^H \mathbf{B} \mathbf{M}_i \right] \\ &= \left(\frac{\partial \mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)}{\partial \mathbf{B}} \right)^*. \end{aligned} \quad (50)$$

$$\mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)) = (\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)))^*. \quad (51)$$

As shown in [31], the second order differential is then given by:

$$d^2 \mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*) = d\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)) d\text{vec}(\mathbf{B}) + d\mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)) d\text{vec}(\mathbf{B}^*). \quad (52)$$

Thus, we have to evaluate the differential of the two derivatives $\mathcal{D}_{\mathbf{B}}$ and $\mathcal{D}_{\mathbf{B}^*}$, i.e. $d\mathcal{D}_{\mathbf{B}}$ and $d\mathcal{D}_{\mathbf{B}^*}$. We start with $d\mathcal{D}_{\mathbf{B}}$. Using the properties $\mathbf{P}_2, \mathbf{P}_3, \mathbf{P}_{10}\text{-}\mathbf{P}_{14}, \mathbf{P}_{17}\text{-}\mathbf{P}_{19}$ in (49), we have:

$$\begin{aligned} d\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)) &= \sum_{i=1}^{N_m} \left[\text{vec}^T \left(\left(\text{OffBdiag}_{(\mathbf{n})} \{ d\mathbf{B}^* \mathbf{M}_i^T \mathbf{B}^T \} \right) \mathbf{B}^* \mathbf{M}_i^* \right. \right. \\ &\quad + \left(\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B}^* \mathbf{M}_i^T d\mathbf{B}^T \} \right) \mathbf{B}^* \mathbf{M}_i^* \\ &\quad + \text{vec}^T \left(\left(\text{OffBdiag}_{(\mathbf{n})} \{ d\mathbf{B}^* \mathbf{M}_i^* \mathbf{B}^T \} \right) \mathbf{B}^* \mathbf{M}_i^T \right. \\ &\quad + \left(\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B}^* \mathbf{M}_i^* d\mathbf{B}^T \} \right) \mathbf{B}^* \mathbf{M}_i^T \\ &\quad + \text{vec}^T \left(\left(\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{M}_i \mathbf{B}^H \} \right)^T d\mathbf{B}^* \mathbf{M}_i^* \right. \\ &\quad \left. \left. + \left(\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{M}_i^H \mathbf{B}^H \} \right)^T d\mathbf{B}^* \mathbf{M}_i^T \right) \right]. \end{aligned}$$

Then, using the properties $\mathbf{P}_1, \mathbf{P}_7$, and $\mathbf{P}_{17}$ we obtain:

$$\begin{aligned} d\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)) &= \sum_{i=1}^{N_m} \left[\left((\mathbf{B}^* \mathbf{M}_i^*)^T \otimes \mathbf{I}_N \right) \text{vec} \left(\text{OffBdiag}_{(\mathbf{n})} \{ d\mathbf{B}^* \mathbf{M}_i^T \mathbf{B}^T \} \right) \right. \\ &\quad + \left((\mathbf{B}^* \mathbf{M}_i^*)^T \otimes \mathbf{I}_N \right) \text{vec} \left(\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B}^* \mathbf{M}_i^T d\mathbf{B}^T \} \right) \\ &\quad + \left((\mathbf{B}^* \mathbf{M}_i^T)^T \otimes \mathbf{I}_N \right) \text{vec} \left(\text{OffBdiag}_{(\mathbf{n})} \{ d\mathbf{B}^* \mathbf{M}_i^* \mathbf{B}^T \} \right) \\ &\quad + \left((\mathbf{B}^* \mathbf{M}_i^T)^T \otimes \mathbf{I}_N \right) \text{vec} \left(\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B}^* \mathbf{M}_i^* d\mathbf{B}^T \} \right) \\ &\quad + \left(\mathbf{I}_N \otimes \left(\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{M}_i \mathbf{B}^H \} \right)^T \right) (\mathbf{M}_i^H \otimes \mathbf{I}_N) \text{vec}(d\mathbf{B}^*) \\ &\quad \left. + \left(\mathbf{I}_N \otimes \left(\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{M}_i^H \mathbf{B}^H \} \right)^T \right) (\mathbf{M}_i \otimes \mathbf{I}_N) \text{vec}(d\mathbf{B}^*) \right]^T. \end{aligned}$$

While properties $\mathbf{P}_2, \mathbf{P}_3, \mathbf{P}_7, \mathbf{P}_8, \mathbf{P}_9$, and $\mathbf{P}_{18}$ involve:

$$\begin{aligned} d\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)) &= \sum_{i=1}^{N_m} \left[\left(\mathbf{M}_i^H \mathbf{B}^H \otimes \mathbf{I}_N \right) \mathbf{T}_{\text{Boff}}(\mathbf{B} \mathbf{M}_i \otimes \mathbf{I}_N) \text{vec}(d\mathbf{B}^*) \right. \\ &\quad + \left(\mathbf{M}_i^H \mathbf{B}^H \otimes \mathbf{I}_N \right) \mathbf{T}_{\text{Boff}}(\mathbf{I}_N \otimes \mathbf{B}^* \mathbf{M}_i^T) \mathbf{K}_{N,M} \text{vec}(d\mathbf{B}) \\ &\quad + \left(\mathbf{M}_i \mathbf{B}^H \otimes \mathbf{I}_N \right) \mathbf{T}_{\text{Boff}}(\mathbf{B} \mathbf{M}_i^H \otimes \mathbf{I}_N) \text{vec}(d\mathbf{B}^*) \\ &\quad + \left(\mathbf{M}_i \mathbf{B}^H \otimes \mathbf{I}_N \right) \mathbf{T}_{\text{Boff}}(\mathbf{I}_N \otimes \mathbf{B}^* \mathbf{M}_i^*) \mathbf{K}_{N,M} \text{vec}(d\mathbf{B}) \\ &\quad + \left(\mathbf{I}_N \otimes \left(\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{M}_i \mathbf{B}^H \} \right)^T \right) (\mathbf{M}_i^H \otimes \mathbf{I}_N) \text{vec}(d\mathbf{B}^*) \\ &\quad \left. + \left(\mathbf{I}_N \otimes \left(\text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{B} \mathbf{M}_i^H \mathbf{B}^H \} \right)^T \right) (\mathbf{M}_i \otimes \mathbf{I}_N) \text{vec}(d\mathbf{B}^*) \right]^T. \end{aligned}$$

The properties $\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3, \mathbf{P}_{15}, \mathbf{P}_{16}, \mathbf{P}_{17}, \mathbf{P}_{18}, \mathbf{P}_{19}$, and $\mathbf{P}_{20}$ imply the following

result:

$$\begin{aligned}
d\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)) &= \sum_{i=1}^{N_m} \text{vec}^T(d\mathbf{B}) \left[(\mathbf{K}_{N,M})^T (\mathbf{I}_N \otimes \mathbf{M}_i \mathbf{B}^H) \mathbf{T}_{\text{Boff}}^T(\mathbf{B}^* \mathbf{M}_i^* \otimes \mathbf{I}_N) \right. \\
&\quad + (\mathbf{K}_{N,M})^T (\mathbf{I}_N \otimes \mathbf{M}_i^H \mathbf{B}^H) \mathbf{T}_{\text{Boff}}^T(\mathbf{B}^* \mathbf{M}_i^T \otimes \mathbf{I}_N) \left. \right] \\
&\quad + \text{vec}^T(d\mathbf{B}^*) \left[(\mathbf{M}_i^T \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{\text{Boff}}^T(\mathbf{B}^* \mathbf{M}_i^* \otimes \mathbf{I}_N) \right. \\
&\quad + (\mathbf{M}_i^* \mathbf{B}^T \otimes \mathbf{I}_N) \mathbf{T}_{\text{Boff}}^T(\mathbf{B}^* \mathbf{M}_i^T \otimes \mathbf{I}_N) \\
&\quad + (\mathbf{M}_i^* \otimes (\text{OffBdiag}_{(n)}\{\mathbf{B} \mathbf{M}_i \mathbf{B}^H\})) \\
&\quad \left. + (\mathbf{M}_i^T \otimes (\text{OffBdiag}_{(n)}\{\mathbf{B} \mathbf{M}_i^H \mathbf{B}^H\})) \right]. \tag{53}
\end{aligned}$$

Finally this $1 \times NM$ vector can be rewritten as follows:

$$d\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)) = \begin{bmatrix} d\text{vec}^T(\mathbf{B}^*) & d\text{vec}^T(\mathbf{B}) \end{bmatrix} \begin{bmatrix} \mathbf{A}_{00} \\ \mathbf{A}_{10} \end{bmatrix}, \tag{54}$$

which leads to the expressions of $\mathbf{A}_{00}$ and $\mathbf{A}_{10}$ given in (16)-(17).

Concerning the second differential $d\mathcal{D}_{\mathbf{B}^*}$, we are taking advantage of (51). Using property $\mathbf{P}_{23}$, we finally have $d\mathcal{D}_{\mathbf{B}^*} = (d\mathcal{D}_{\mathbf{B}}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)))^*$, which leads to:

$$d\mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)) = \begin{bmatrix} d\text{vec}^T(\mathbf{B}) & d\text{vec}^T(\mathbf{B}^*) \end{bmatrix} \begin{bmatrix} \mathbf{A}_{00}^* \\ \mathbf{A}_{10}^* \end{bmatrix}. \tag{55}$$

But this $1 \times NM$ vector was supposed to be written as:

$$d\mathcal{D}_{\mathbf{B}^*}(\mathcal{C}_{\text{JBD}}(\mathbf{B}, \mathbf{B}^*)) = \begin{bmatrix} d\text{vec}^T(\mathbf{B}^*) & d\text{vec}^T(\mathbf{B}) \end{bmatrix} \begin{bmatrix} \mathbf{A}_{01} \\ \mathbf{A}_{11} \end{bmatrix}, \tag{56}$$

implying that: $\mathbf{A}_{11} = \mathbf{A}_{00}^*$ and that $\mathbf{A}_{10}^* = \mathbf{A}_{01}$.

By incorporating (54) and (56) in (52), the results stated by (12) can be found, where $\mathbf{A}_{ij}$ for $i, j = 1, \dots, 2$ are given by (16)-(17). The exact expression of the four complex Hessian matrices given in (13), (14) and (15) can be derived.

A.2- Finding the coefficients of the 4th-degree polynomial involved in the calculation of the optimal stepsize

To determine at each iteration the optimal stepsize, we have to evaluate $\mathcal{C}_{\text{JBD}}(\mathbf{B}^{(m+1)}) = \mathcal{C}_{\text{JBD}}(\mathbf{B}^{(m)} - \mu \mathbf{D}^{(m)})$. In the following, to simplify the different expressions, the dependency upon the iteration m will be omitted. Using the properties $\mathbf{P}_{19}$ and $\mathbf{P}_{21}$, the cost function $\mathcal{C}_{\text{JBD}}(\mathbf{B} - \mu \mathbf{D})$ can be expressed as:

$$\begin{aligned}
&\mathcal{C}_{\text{JBD}}(\mathbf{B} - \mu \mathbf{D}) \\
&= \sum_{i=1}^{N_m} \left\| \text{OffBdiag}_{(n)} \left\{ \mu^2 \mathbf{D} \mathbf{M}_i \mathbf{D}^H - \mu (\mathbf{B} \mathbf{M}_i \mathbf{D}^H + \mathbf{D} \mathbf{M}_i \mathbf{B}^H) + \mathbf{B} \mathbf{M}_i \mathbf{B}^H \right\} \right\|_F^2 \\
&= \sum_{i=1}^{N_m} \left\| \text{OffBdiag}_{(n)} \left\{ \mu^2 \mathbf{C}_2 - \mu \mathbf{C}_1 + \mathbf{C}_0 \right\} \right\|_F^2 \\
&= \sum_{i=1}^{N_m} \text{tr} \left\{ (\mu^2 \mathbf{C}_2 - \mu \mathbf{C}_1 + \mathbf{C}_0)^H \text{OffBdiag}_{(n)} \left\{ \mu^2 \mathbf{C}_2 - \mu \mathbf{C}_1 + \mathbf{C}_0 \right\} \right\} \tag{57}
\end{aligned}$$

where

$$\begin{aligned}\mathbf{C}_2 &= \mathbf{D}\mathbf{M}_i\mathbf{D}^H \\ \mathbf{C}_1 &= (\mathbf{B}\mathbf{M}_i\mathbf{D}^H + \mathbf{D}\mathbf{M}_i\mathbf{B}^H), \\ \mathbf{C}_0 &= \mathbf{B}\mathbf{M}_i\mathbf{B}^H.\end{aligned}\tag{58}$$

Using properties $\mathbf{P}_0, \mathbf{P}_4, \mathbf{P}_{19}, \mathbf{P}_{21}$ and $\mathbf{P}_{22}$, we find that:

$$\begin{aligned}\mathcal{C}_{\text{JBD}}(\mathbf{B} - \mu\mathbf{D}) &= \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_0^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_0 \} \right\} \\ &\quad - \mu \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_1^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_0 \} + \mathbf{C}_0^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_1 \} \right\} \\ &\quad + \mu^2 \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_2^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_0 \} + \mathbf{C}_1^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_1 \} + \mathbf{C}_0^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_2 \} \right\} \\ &\quad - \mu^3 \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_2^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_1 \} + \mathbf{C}_1^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_2 \} \right\} \\ &\quad + \mu^4 \sum_{i=1}^{N_m} \text{tr} \left\{ \mathbf{C}_2^H \text{OffBdiag}_{(\mathbf{n})} \{ \mathbf{C}_2 \} \right\} = a_0 + a_1\mu + a_2\mu^2 + a_3\mu^3 + a_4\mu^4.\end{aligned}\tag{59}$$

It finally leads to the results stated by Eq. (29)–(33).

References

- [1] O. Axelsson, “A survey of preconditioned iterative methods for linear systems of algebraic equations,” *BIT*, vol. 25, no. 1, pp. 166–187, 1985.
- [2] A. Belouchrani, K. Abed-Meraïm, M. G. Amin, and A. Zoubir, “Joint anti-diagonalization for blind source separation,” in *Proc. IEEE ICASSP*, Salt Lake City, USA, May 2001, pp. 2196–2199.
- [3] A. Belouchrani, K. Abed-Meraïm, J.-F. Cardoso, and E. Moulines, “A blind source separation technique using second order statistics,” *IEEE Trans. Signal Proc.*, vol. 45, pp. 434–444, Feb. 1997.
- [4] A. Belouchrani and M. G. Amin, “Blind source separation based on time-frequency signal representations,” *IEEE Trans. Signal Proc.*, vol. 46, no. 11, pp. 2888–2897, Nov. 1998.
- [5] A. Belouchrani, M. G. Amin, and K. Abed-Meraïm, “Direction finding in correlated noise fields based on joint block-diagonalization of spatio-temporal correlation matrices,” *IEEE Signal Process. Lett.*, vol. 4, no. 9, pp. 266–268, Sept. 1997.
- [6] A. Belouchrani, M. G. Amin, N. Thirion-Moreau, and Y. Zang, “Source separation and localization using time-frequency distributions,” *IEEE Signal Process. Mag.*, pp. 97–107, Nov. 2013.

- [7] M. Benzi, “Preconditioning techniques for large linear systems: A survey,” *J. COMPUT. PHYS.*, vol. 182, pp. 418–477, 2002.
- [8] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, March 2004.
- [9] J. Brewer, “Kronecker products and matrix calculus in system theory,” *IEEE Trans. on Circuits and Systems*, vol. Cas-25, no. 9, pp. 772–781, September 1978.
- [10] J.-F. Cardoso and A. Souloumiac, “Blind beamforming for non-gaussian signals,” *IEE Proc.-F*, vol. 140, no. 6, pp. 362–370, Dec. 1993.
- [11] G. Chabriel and J. Barrère, “Non-symmetrical joint zero-diagonalization and mimo zero-division multiple access,” *IEEE Trans. Signal Proc.*, vol. 59, no. 5, pp. 2296–2307, 2011.
- [12] —, “A direct algorithm for nonorthogonal approximate joint diagonalization,” *IEEE Trans. Signal Proc.*, vol. 60, no. 1, pp. 39–47, 2012.
- [13] G. Chabriel, J. Barrère, N. Thirion-Moreau, and E. Moreau, “Algebraic joint zero-diagonalization and blind source separation,” *IEEE Trans. Signal Proc.*, vol. 56, no. 3, pp. 980–989, March 2008.
- [14] G. Chabriel, M. Kleinstuber, E. Moreau, H. Shen, P. Tichavsky, and A. Yeredor, “Joint matrices decompositions and blind source separation: A survey of methods, identification, and applications,” *IEEE Signal Process. Mag.*, vol. 31, no. 3, pp. 34–43, 2014.
- [15] K. Chen, *Matrix Preconditioning Techniques and Applications (Cambridge Monographs on Applied and Computational Mathematics)*, 2005.
- [16] O. Cherrak, H. Ghennioui, N. Thirion-Moreau, and E.-H. Abarkan, “Non-unitary joint block diagonalization of matrices using a levenberg-marquardt algorithm,” in *Proc. EUSIPCO*, Marrakech, Morocco, Sept. 2013, pp. 1–5.
- [17] P. Comon, “Independant component analysis, a new concept ?” *Signal Process.*, vol. 36, pp. 287–314, 1994.
- [18] Y. H. Dai and Y. Yuan, “A nonlinear conjugate gradient method with a strong global convergence property,” *SIAM J. Optim*, vol. 10, pp. 177–182, 1999.
- [19] L. De Lathauwer, C. Févotte, B. De Moor, and J. Vandewalle, “Jacobi algorithm for block diagonalization in blind identification,” in *Proc. Symp. Inform. Theory in the Benelux*, Louvain-la-Neuve, Belgium, May 2002, pp. 155–162.
- [20] S. Dégerine and E. Kane, “A comparative study of approximate joint diagonalization algorithms for blind source separation in presence of additive noise,” *IEEE Trans. Signal Proc.*, vol. 55, no. 6, pp. 3022–3031, June 2007.
- [21] D. J. Evans, “The use of pre-conditioning in iterative methods for solving linear systems with symmetric positive definite matrices,” *J. Inst. Math. Its Appl.*, vol. 4, pp. 295–314, 1968.
- [22] E. M. Fadaili, N. Thirion-Moreau, and E. Moreau, “Algorithme de zéro-diagonalisation conjointe pour la séparation de sources déterministes,” in *20 Colloque GRETSI*, Louvain-La-Neuve, Belgique, Septembre 2005.

- [23] ———, “Non-orthogonal joint diagonalization/zero-diagonalization for source separation based on time-frequency distributions,” *IEEE Trans. Signal Proc.*, vol. 55, no. 5, May 2007.
- [24] C. Févotte and C. Doncarli, “A unified presentation of blind separation methods for convolutive mixtures using block-diagonalization,” Nara, Japan, April 2003, pp. 349–354.
- [25] H. Ghennioui, E.-M. Fadaili, N. Thirion-Moreau, A. Adib, and E. Moreau, “A nonunitary joint block diagonalization algorithm for blind separation of convolutive mixtures of sources,” *IEEE Signal Process. Lett.*, vol. 14, no. 11, pp. 860–863, No. 2007.
- [26] H. Ghennioui, N. Thirion-Moreau, E. Moreau, and D. Aboutajdine, “Gradient-based joint block diagonalization algorithms: Application to blind separation of fir convolutive mixtures,” *Signal Process.*, vol. 90, no. 6, pp. 1836–1849, June 2010.
- [27] H. Ghennioui, N. Thirion-Moreau, E. Moreau, A. Adib, and D. Aboutajdine, “Non unitary joint-block diagonalization of complex matrices using a gradient approach,” in *7th International Conference on Independent Component Analysis and Signal Separation*, M. E. D. et al., Ed. London, UK: Springer-Verlag Berlin Heidelberg, LNCS 4666, Sept. 2007, pp. 201–208.
- [28] L. Giulieri, H. Ghennioui, N. Thirion-Moreau, and E. Moreau, “Non orthogonal joint diagonalization of spatial quadratic time-frequency matrices for sources separation,” *IEEE Signal Process. Lett.*, vol. 12, no. 5, pp. 415–418, May 2005.
- [29] G.-H. Golub and C.-F. Van Loan, *Matrix computation*, M. J. Hopkins, Ed. Univ. Press Baltimore, 1989.
- [30] W. Hager and H. Zhang, “A survey of nonlinear conjugate gradient methods,” *Pacific J. Optim.*, vol. 2, pp. 35–58, 2006.
- [31] A. Hjørungnes, *Complex-Valued Matrix Derivatives With Applications in Signal Processing and Communications*. Cambridge, University Press, Feb. 2011.
- [32] D. Lahat, J.-F. Cardoso, and H. Messer, “Joint block diagonalization algorithms for optimal separation of multidimensional components,” in *J. In Latent Variable Analysis and Signal Separation, ser. LNCS*, ser. Heidelberg: Springer, F. Theis, A. Cichocki, A. Yeredor, and M. E. Zibulevsky, Eds., vol. 7191, 2012, pp. 155–162.
- [33] D. D. Lee and H. S. Seung, “Algorithms for non-negative matrix factorization,” in *In NIPS*. MIT Press, 2000, pp. 556–562.
- [34] H. Lev-Ari, “Efficient solution of linear matrix equations with application to multistatic antenna array processing,” *Communications in Information and Systems*, vol. 5, no. 1, pp. 123–130, 2005.
- [35] X.-L. Li and X. D. Zhang, “Nonorthogonal joint diagonalization free of degenerate solution,” *IEEE Trans. Signal Proc.*, vol. 55, no. 5, pp. 1803–1814, May 2007.

- [36] D. G. Luenberger, *Optimization by vector space methods*. Wiley, 1969.
- [37] J. R. Magnus and H. Neudecker, *Matrix differential calculus with applications in statistics and econometrics*. John Wiley and sons, Inc, Third Edition, 2007.
- [38] V. Maurandi, C. De Luigi, and E. Moreau, “Fast jacobi like algorithms for joint diagonalization of complex symmetric matrices,” in *Signal Processing Conference (EUSIPCO), 2013 Proceedings of the 21st European*, Sept 2013, pp. 1–5.
- [39] E. Moreau, “A generalization of joint-diagonalization criteria for source separation,” *IEEE Trans. Signal Proc.*, vol. 49, no. 3, pp. 530–541, March 2001.
- [40] E. Moreau and T. Adali, *Blind identification and separation of complex valued signals*, ser. Focus Digital Signal and Image Processing Series, I. Ltd and I. Wiley & Sons, Eds., Sept. 2013.
- [41] D. Nion, “A tensor framework for non-unitary joint block diagonalization,” *IEEE Trans. Signal Proc.*, vol. 59, no. 10, pp. 4585–4594, Oct. 2011.
- [42] P. Paatero, “The multilinear engine—a table-driven, least-squares program for solving multi-linear problems, including the n-way parallel factor analysis mode,” *J. Comput. Graph. Stat.*, vol. 8, no. 4, pp. 854–888, Dec. 1999.
- [43] D.-T. Pham, “Joint approximate diagonalization of positive definite matrices,” *SIAM J. Matrix Analysis and Applications*, vol. 22, no. 4, pp. 1136–1152, 2001.
- [44] E. Polak, *Optimization algorithms and consistent approximations*. Springer, 1997.
- [45] H. Shen and M. Kleinstuber, “Non-unitary matrix joint diagonalization for complex independent vector analysis,” *EURASIP Journal on Advances in Signal Processing*, vol. 2012, no. 1, 2012.
- [46] J. R. Shewchuk, “Introduction to the conjugate gradient method without the agonizing pain,” Carnegie Mellon University, Pittsburgh, Pennsylvania, USA, Technical report, 1994.
- [47] A. Souloumiac, “Nonorthogonal joint diagonalization by combining givens and hyperbolic rotations,” *IEEE Trans. Signal Proc.*, vol. 57, no. 6, pp. 2222–2231, June 2009.
- [48] P. Tichavsky and Z. Koldovsky, “Algorithms for nonorthogonal approximate joint block-diagonalization,” in *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*, Aug 2012, pp. 2094–2098.
- [49] P. Tichavsky and A. Yeredor, “Fast approximate joint diagonalization incorporating weight matrices,” *IEEE Trans. Signal Proc.*, vol. 57, no. 3, pp. 878–891, March 2009.
- [50] T. Trainini and E. Moreau, “A coordinate descent algorithm for complex joint diagonalization under hermitian and transpose congruences,” *IEEE Trans. Signal Proc.*, vol. 62, no. 19, pp. 4974–4983, Oct. 2014.
- [51] A. M. Turing, “Rounding-off errors in matrix processes,” *Quarterly J. Mech. Appl. Math.*, vol. 1, pp. 287–308, 1948.

- [52] J. Walgate, A. J. Short, L. Hardy, and V. Vedral, "Local distinguishability of multipartite orthogonal quantum states," *Physical Review Letter*, vol. 85, pp. 4972–4975, 2000.
- [53] J. R. Westlake, *A Handbook of Numerical Matrix Inversion and Solution of Linear Equations*. New York: Wiley, 1968.
- [54] X.-F. Xu, D.-Z. Feng, W. X. Zheng, and H. Zhang, "Convulsive blind source separation based on joint block toeplitzization and block-inner diagonalization," *Signal Processing*, vol. 90, no. 1, pp. 119–133, 2010.
- [55] A. Yeredor, "Non-orthogonal joint-diagonalization in the least squares sense with application in blind sources separation," *IEEE Trans. Signal Proc.*, vol. 50, no. 7, pp. 1545–1553, July 2002.
- [56] A. Yeredor, A. Ziehe, and K.-R. Müller, "Approximate joint diagonalization using natural gradient approach," in *Lecture Notes in Computer Science, in Proceedings ICA*, Granada, Espagne, May 2004, pp. 89–96.
- [57] W.-J. Zeng, X.-L. Li, and X. Zhang, X.-D. and Jiang, "An improved signal-selective direction finding algorithm using second-order cyclic statistics," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Taipei, Taiwan, April 2009, pp. 2141–2144.
- [58] W. T. Zhang and S. T. Lou, "A low complexity iterative algorithm for joint zero-diagonalization," *IEEE Signal Process. Lett.*, vol. 19, no. 2, pp. 115–118, Jan. 2012.

List of Figures

1	Comparison of the different algorithms : evolution of the error index $I(\mathbf{G})$ versus the number of iterations in the noiseless case (SNR =100 dB) for different sizes of the matrix sets. Left : $N_m =100$. Right: $N_m =5$	30
2	Comparison of the different algorithms : evolution of the error index $I(\mathbf{G})$ versus the number of iterations in the noisy case (SNR =30 dB) for different sizes of the matrix sets. Left : $N_m =100$. Right: $N_m =5$	30
3	Random initialization : Comparison of the different algorithms : evolution of the error index $I(\mathbf{G})$ versus the number of iterations in the noiseless case (SNR =100 dB) for different sizes of the matrix sets. Left : $N_m =100$. Right: $N_m =10$	31
4	Random initialization : Comparison of the different algorithms : evolution of the error index $I(\mathbf{G})$ versus the number of iterations in the noisy case (SNR =30 dB) for different sizes of the matrix sets. Left : $N_m =100$. Right: $N_m =10$	31
5	Comparison of the different algorithms: evolution of the error index $I(\mathbf{G})$ versus the number of iterations in the noiseless case (SNR =100 dB) for different sizes of the matrix sets where the matrix $\mathbf{A}$ is rectangular. Left : $N_m =100$. Right: $N_m =5$	32
6	Comparison of the different algorithms: evolution of the error index $I(\mathbf{G})$ versus the number of iterations in the noiseless case (SNR =100 dB) for different sizes of the matrix sets where the matrix $\mathbf{A}$ is rectangular. Left : $N_m =100$. Right: $N_m =5$	32
7	Comparison of the different algorithms : evolution of the error index $I(\mathbf{G})$ versus the number, N_m , of matrices that are used. Left : SNR = 25 dB. Right: SNR = 100 dB.	33
8	Comparison of the different algorithms : evolution of the error index $I(\mathbf{G})$ versus the SNR for different sizes of the matrix sets. Left : $N_m =5$. Right: $N_m =100$	33
9	Comparison of the different algorithms in the joint diagonalization context. Evolution of the error index $I(\mathbf{G})$ versus the number, N_m , of matrices that are used. Left : SNR = 30 dB. Right: SNR = 100 dB.	34
10	Comparison of the different algorithms in the joint diagonalization context. Evolution of the error index $I(\mathbf{G})$ versus the SNR for different sizes of the matrix sets. Left : $N_m =5$. Right: $N_m =100$	34
11	Comparison of the different algorithms: evolution of the error index $I(\mathbf{G})$ versus the number r of blocks that are considered. The size N_m of the matrix sets is equal to 100. Two different values of the SNR are considered. Left : SNR = 100 dB. Right : SNR = 30 dB.	35
12	t - f points selected with automatic points detector C_{ConvGH} . Left: SNR =20 dB. Right: SNR =100 dB.	36
13	Comparison of the different algorithms in the BSS context: evolution of the error index $I(\mathbf{G})$ versus the SNR.	36

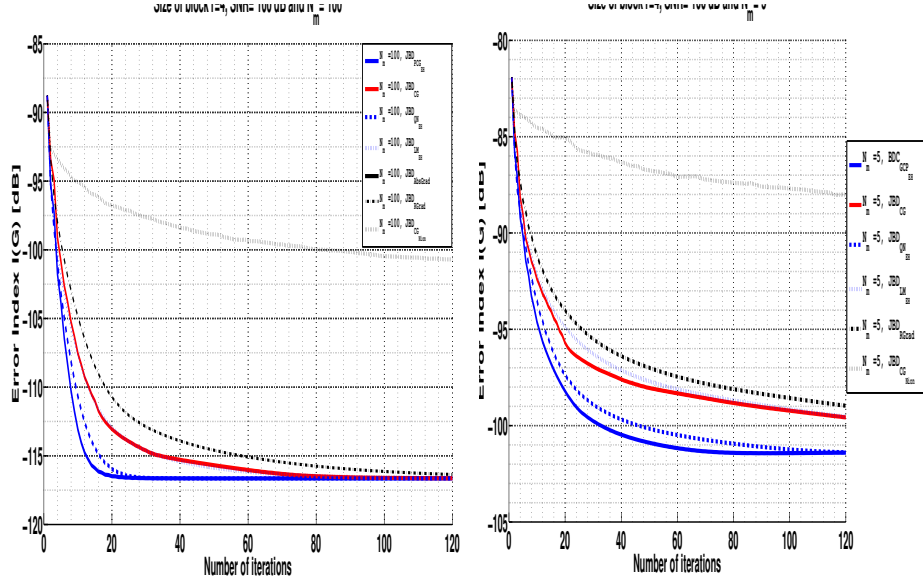


Figure 1: Comparison of the different algorithms : evolution of the error index $I(\mathbf{G})$ versus the number of iterations in the noiseless case (SNR = 100 dB) for different sizes of the matrix sets. Left : $N_m = 100$. Right: $N_m = 5$.

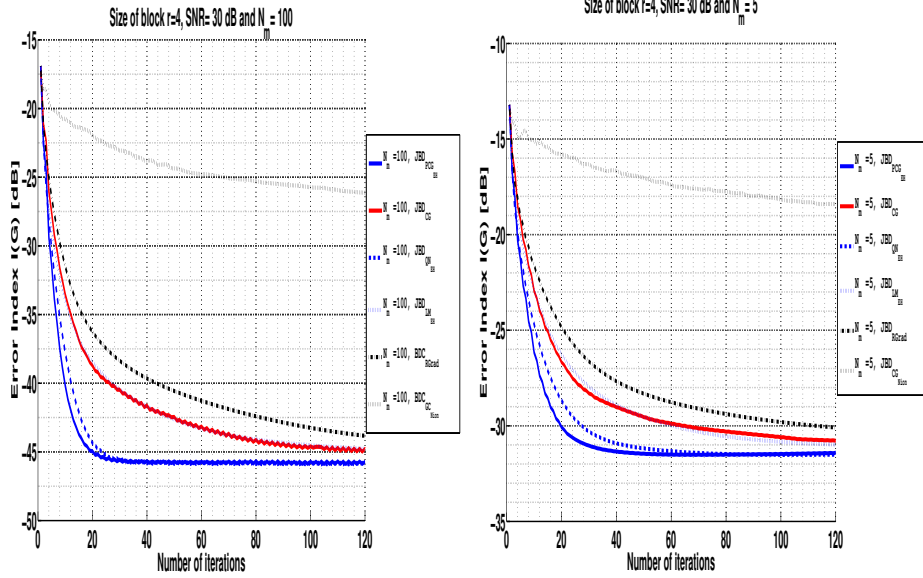


Figure 2: Comparison of the different algorithms : evolution of the error index $I(\mathbf{G})$ versus the number of iterations in the noisy case (SNR = 30 dB) for different sizes of the matrix sets. Left : $N_m = 100$. Right: $N_m = 5$.

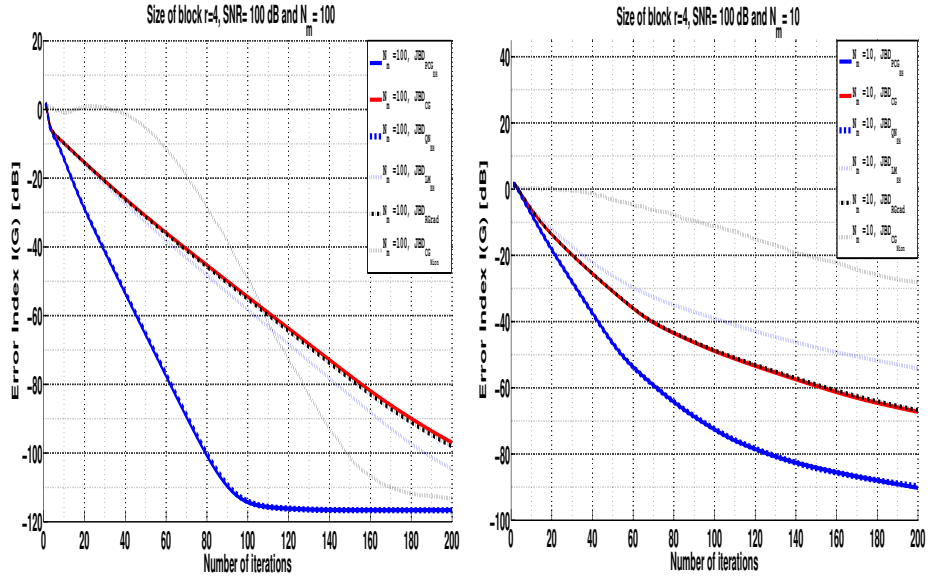


Figure 3: Random initialization : Comparison of the different algorithms : evolution of the error index $I(G)$ versus the number of iterations in the noiseless case (SNR = 100 dB) for different sizes of the matrix sets. Left : $N_m = 100$. Right: $N_m = 10$.

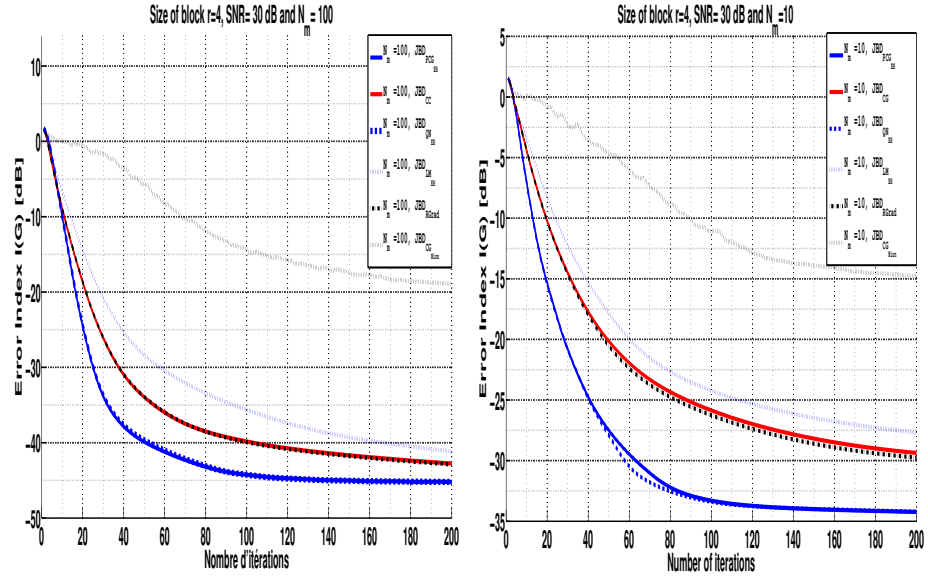


Figure 4: Random initialization : Comparison of the different algorithms : evolution of the error index $I(G)$ versus the number of iterations in the noisy case (SNR = 30 dB) for different sizes of the matrix sets. Left : $N_m = 100$. Right: $N_m = 10$.

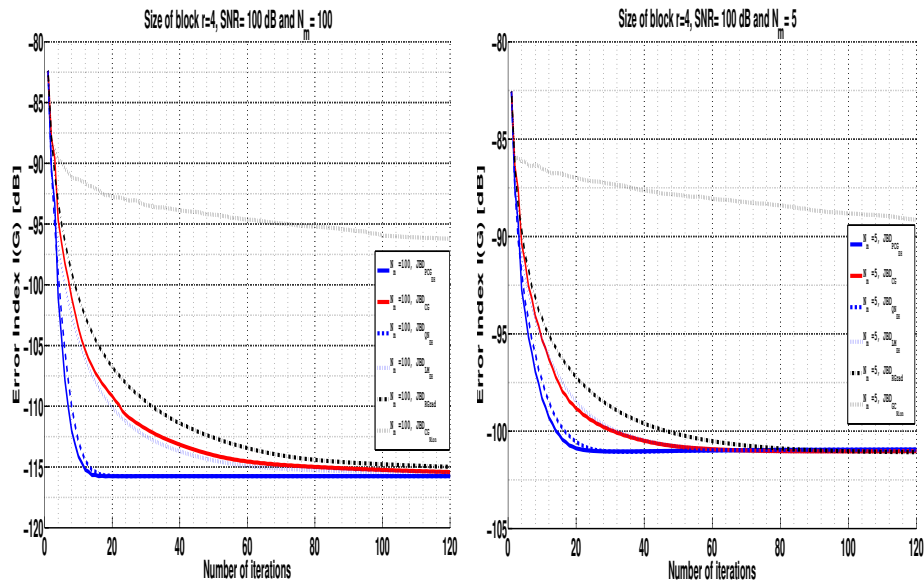


Figure 5: Comparison of the different algorithms: evolution of the error index $I(\mathbf{G})$ versus the number of iterations in the noiseless case (SNR = 100 dB) for different sizes of the matrix sets where the matrix $\mathbf{A}$ is rectangular. Left : $N_m = 100$. Right: $N_m = 5$.

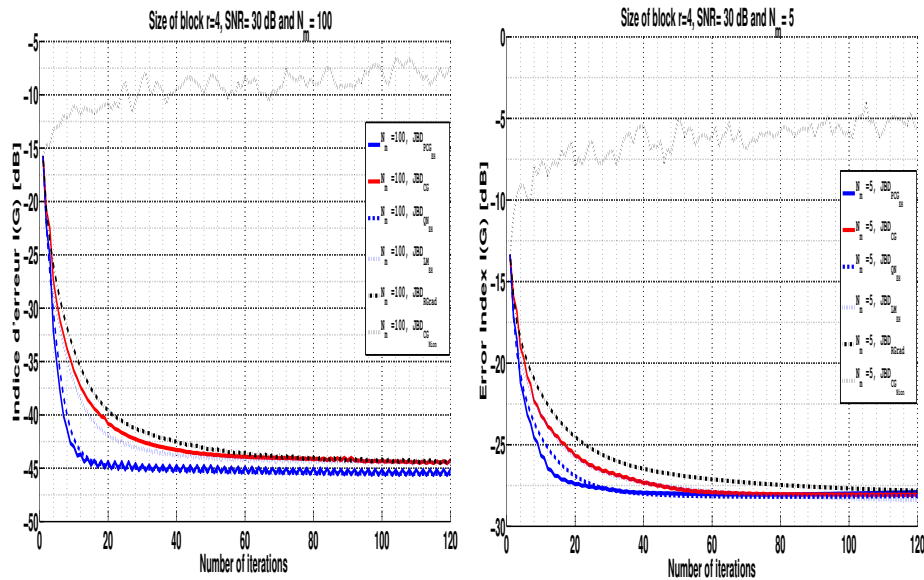


Figure 6: Comparison of the different algorithms: evolution of the error index $I(\mathbf{G})$ versus the number of iterations in the noiseless case (SNR = 100 dB) for different sizes of the matrix sets where the matrix $\mathbf{A}$ is rectangular. Left : $N_m = 100$. Right: $N_m = 5$.

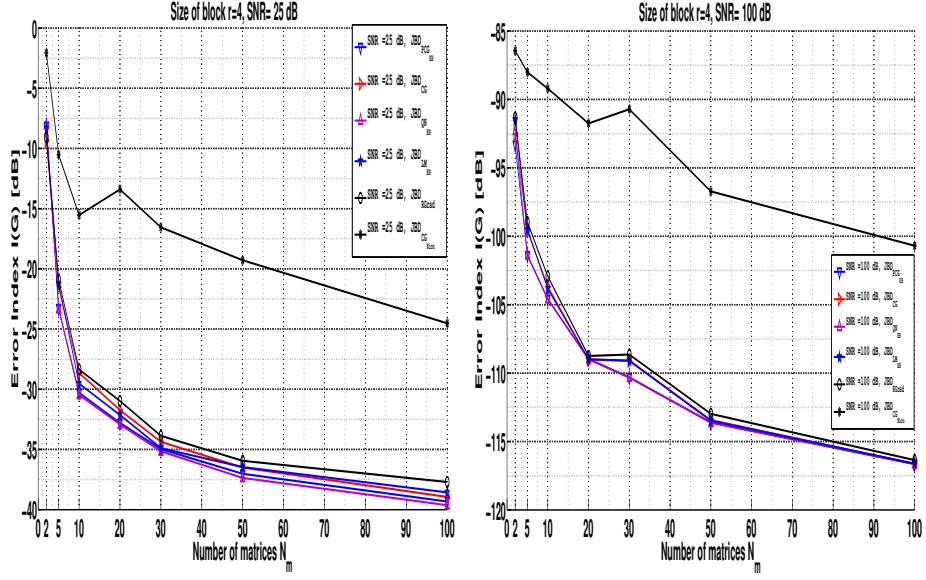


Figure 7: Comparison of the different algorithms : evolution of the error index $I(\mathbf{G})$ versus the number, N_m , of matrices that are used. Left : SNR = 25 dB. Right: SNR = 100 dB.

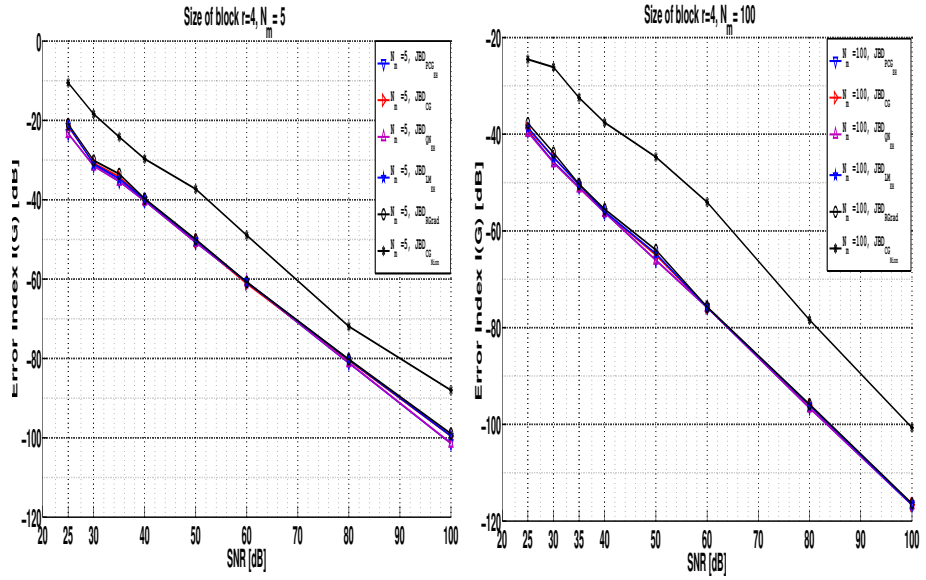


Figure 8: Comparison of the different algorithms : evolution of the error index $I(\mathbf{G})$ versus the SNR for different sizes of the matrix sets. Left : $N_m = 5$. Right: $N_m = 100$.

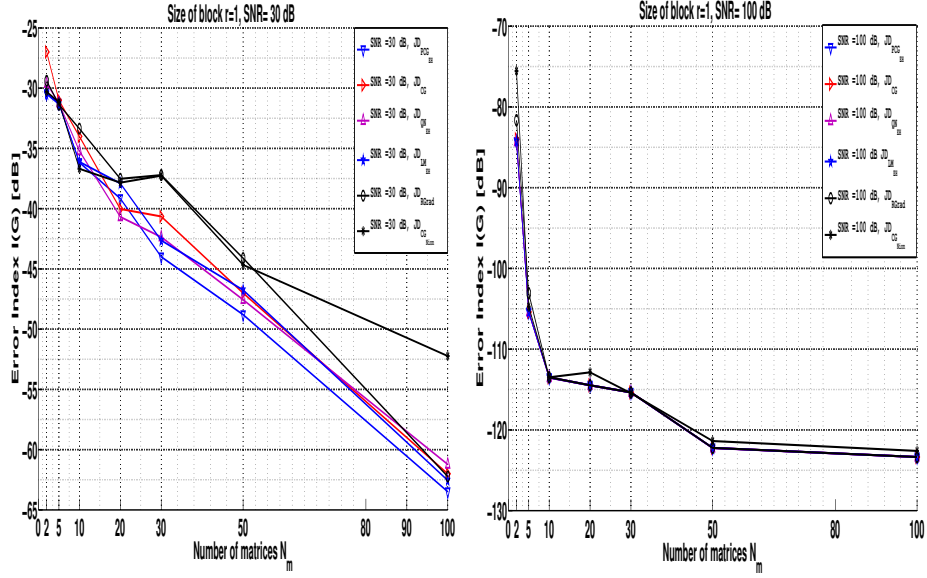


Figure 9: Comparison of the different algorithms in the joint diagonalization context. Evolution of the error index $I(G)$ versus the number, N_m , of matrices that are used. Left : SNR = 30 dB. Right: SNR = 100 dB.

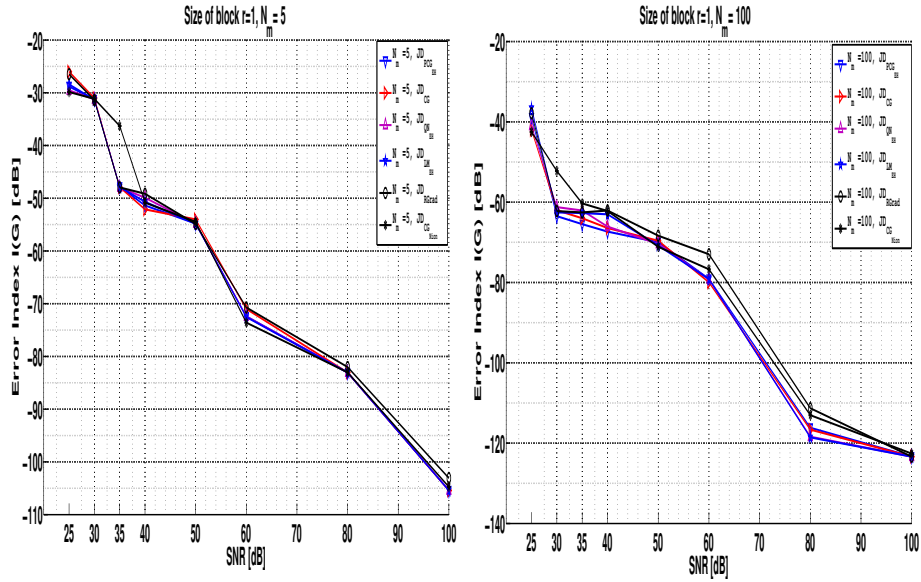


Figure 10: Comparison of the different algorithms in the joint diagonalization context. Evolution of the error index $I(G)$ versus the SNR for different sizes of the matrix sets. Left : $N_m = 5$. Right: $N_m = 100$.

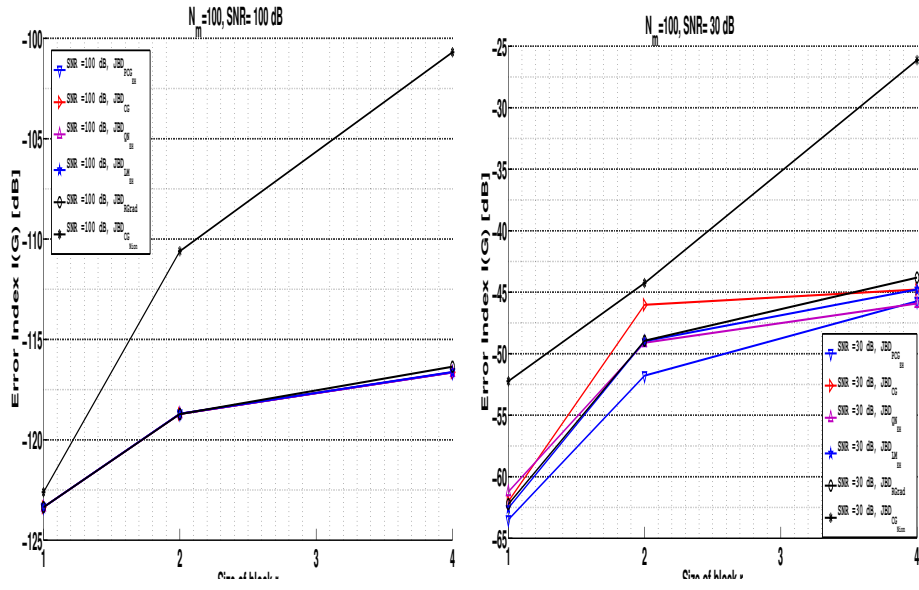


Figure 11: Comparison of the different algorithms: evolution of the error index $I(\mathbf{G})$ versus the number r of blocks that are considered. The size N_m of the matrix sets is equal to 100. Two different values of the SNR are considered. Left : SNR = 100 dB. Right : SNR = 30 dB.

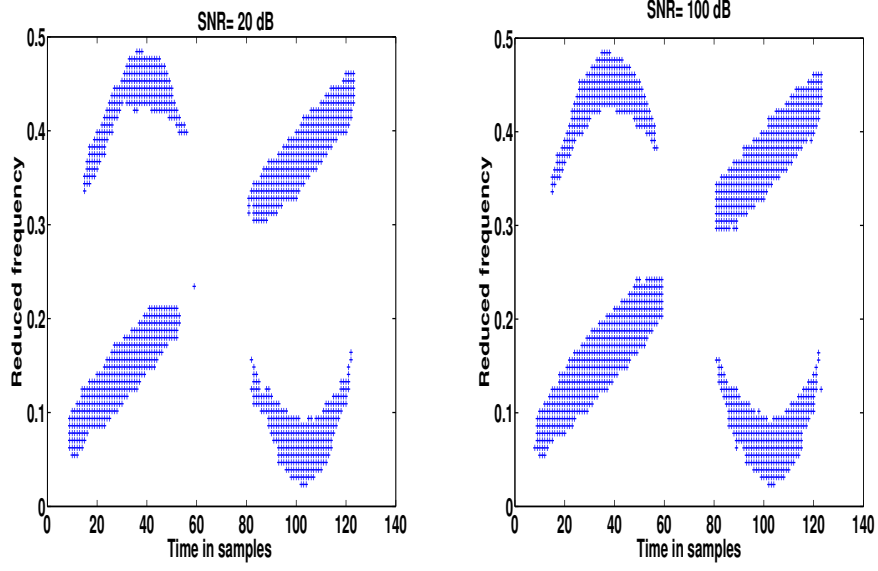


Figure 12: t - f points selected with automatic points detector C_{ConvGH} . Left: SNR = 20 dB. Right: SNR = 100 dB.

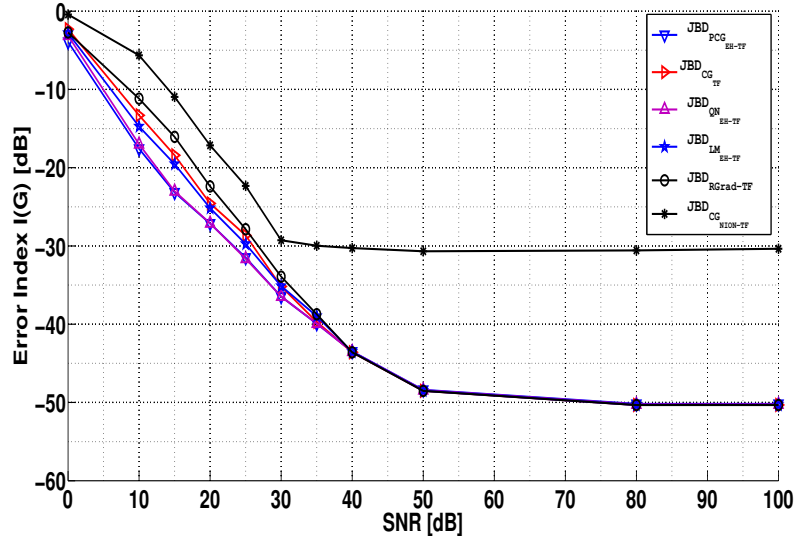


Figure 13: Comparison of the different algorithms in the BSS context: evolution of the error index $I(G)$ versus the SNR.



Résumé de la thèse

Dans cette thèse, nous nous intéressons au système d'identification automatique spatial lequel est dédié à la surveillance maritime par satellite. Ce système couvre une zone bien plus large que le système standard à terre correspondant à plusieurs cellules traditionnelles ce qui peut entraîner des risques de collision des données envoyées par des navires localisés dans des cellules différentes et reçues au niveau de l'antenne du satellite. Nous présentons différentes approches afin de répondre au problème de collision considéré. Elles ne reposent pas toujours sur les mêmes hypothèses en ce qui concerne les signaux reçus, et ne s'appliquent donc pas toutes dans les mêmes contextes (nombre de capteurs utilisés, mode semi-supervisé avec utilisation de trames d'apprentissage et information a priori ou mode aveugle, problèmes liés à la synchronisation des signaux, etc...).

Dans un premier temps, nous proposons des méthodes permettant la séparation/dé-collision des messages en modèle sur-déterminé (plus de capteurs que de messages). Elles sont fondées sur des algorithmes de décompositions matricielles conjointes combinés à des détecteurs de points temps-fréquence (retard-fréquence Doppler) particuliers permettant la construction d'ensembles de matrices devant être (bloc) ou zéro (bloc) diagonalisées conjointement. En ce qui concerne les algorithmes de décompositions matricielles conjointes, nous proposons quatre nouveaux algorithmes de bloc diagonalisation conjointe (de même que leur version à pas optimal) fondés respectivement sur des algorithmes d'optimisation de type gradient conjugué, gradient conjugué pré-conditionné, Levenberg-Marquardt et Quasi-Newton. Le calcul exact du gradient matriciel complexe et des matrices Hessiennes complexes est mené. Nous introduisons également un nouveau problème dénommé zéro-bloc diagonalisation conjointe non-unitaire lequel généralise le problème désormais classique de la zéro-diagonalisation conjointe non-unitaire. Il implique le choix d'une fonction de coût adaptée et à nouveau le calcul de quantités telles que gradient matriciel complexe et les matrices Hessiennes complexes. Nous proposons ensuite trois nouveaux algorithmes à pas optimal fondés sur des algorithmes d'optimisation de type gradient conjugué, gradient conjugué pré-conditionné et Levenberg-Marquardt.

Finalement, nous terminons par des approches à base de techniques de détection multi-utilisateurs conjointe susceptibles de fonctionner en contexte sous-déterminé dans lequel nous ne disposons plus que d'un seul capteur recevant simultanément plusieurs signaux sources. Nous commençons par développer une première approche par déflation consistant à supprimer successivement les interférences. Nous proposons ensuite un deuxième mode opératoire fondé quant à lui sur l'estimateur du maximum de vraisemblance conjoint qui est une variante de l'algorithme de VITERBI.

Mots clés :

Système d'identification automatique spatial, surveillance maritime, (dé)-collision, décompositions matricielles conjointes, algorithmes de bloc et zéro-bloc diagonalisation conjointe, séparation aveugle de sources, détection automatique, spectres et distributions spatiaux (spatiales) bilinéaires et quadratiques, représentation temps-fréquence, détection multi-utilisateurs conjointe, détection par suppression successive des interférences, estimateur de séquence à maximum de vraisemblance conjoint.