Factorisations des tableaux multidimensionnels de données sous contraintes : unicité et applications

Sebastian Miron

UNIVERSITÉ DE LORRAINE

École doctorale IAEM Lorraine

Factorisations des tableaux multidimensionnels de données sous contraintes : unicité et applications

Mémoire de recherche

présenté publiquement le 4 février 2022 pour l'obtention d'une

Habilitation à Diriger des Recherches de l'Université de Lorraine

(mention automatique, traitement du signal et génie informatique)

par

Sebastian MIRON

Composition du jury

Président :	M. Mars Jérôme, Professeur des Universités, Grenoble-INP
Rapporteurs :	M. Jutten Christian, Professeur émérite, Université Grenoble-Alpes
	M. Richard Cédric, Professeur des Universités, Université Côte d'Azur
	Mme. Thirion Moreau Nadège, Professeur des Universités, Université de Toulon
Examinateurs :	M. Boyer Rémy, Professeur des Universités, Université de Lille
	M. Brie David, Professeur des Universités, Université de Lorraine
	Mme. Clausel Marianne, Professeur des Universités, Université de Lorraine
	M. Haardt Martin, Professor, Ilmenau University of Technology (Allemagne)

Centre de Recherche en Automatique de Nancy
CRAN UMR 7039, Université de Lorraine, CNRS
Faculté des Sciences et Technologies, BP 70239, 54506 Vandœuvre Cedex
Tél. : 03 72 74 53 37 • Courriel: sebastian.miron@univ-lorraine.fr

A Gina, Anna et Léonard

Contents

List of Figures

List of Tables

Part I

Présentation générale

Chapitre 1

CV détaillé

1.1 Identification

Nom d'usage :	MIRON	**Prénom :**	Sebastian
Naissance :	11/03/1977	**Lieu :**	Suceava (Roumanie)
Statut familial :	marié, 2 enfants		
Affectation :	Université de Lorraine	**Date :**	1er septembre 2006
Section CNU :	61	**Corps :**	Maître de conférences
Grade :	7ème échelon depuis le 1er août 2020		

1.2 Fonctions principales

2006–21 Maître de conférences à l'Université de Lorraine.
Recherche au département BioSIS du CRAN UMR 7039.
Enseignement au département Réseaux et Télécommunications (R&T) de l'IUT Nancy-Brabois.

2005–06 Attaché Temporaire à l'Enseignement et la Recherche (ATER) à l'IUT 1 de Grenoble.
Enseignement au département Génie Télécoms et Réseaux (GTR) de l'IUT 1 de Grenoble.
Recherche au Laboratoire des Images et des Signaux (LIS), Grenoble, équipe Signaux et Images Naturels.

1.3 Titres universitaires

2002–05 Doctorat en traitement des signaux et images au LIS, soutenue en octobre 2005 (Grenoble).
Sujet : *Méthodes multilinéaires et hypercomplexes en traitement d'antenne multicomposante haute résolution.*

Directeur de thèse : Jérome Mars, co-encadrant : Nicolas Le Bihan.
Jury : Pierre Comon, Jean-Jaques Fuchs (rapporteurs), Pascal Chevalier, Stephen Sangwine (examinateurs), Michel Campillo (président).

2001–02 Diplôme d'études approfondies (DEA) « Signal, Image, Parole, Télécommunications » de l'Institut National Polytechnique de Grenoble (INPG).

1996–01 Diplôme d'ingénieur en Electronique et Télécommunications de l'Université Technique « Gh. Asachi » Iasi, Roumanie.

1.4 Activités d'enseignement et responsabilités pédagogiques

En tant que Maître de conférences, j'enseigne depuis 2006 au département Réseaux et Télécommunications de l'IUT Nancy-Brabois. Mon enseignement est essentiellement dispensé aux étudiants de première et deuxième année du Diplôme Universitaire de Technologie (DUT) et en Licence Professionnelle (LP) Métiers des Réseaux Informatiques et Télécommunications (MRIT), parcours Réseaux Sans Fil et Sécurité (RSFS).

Depuis 2019 j'interviens également à PolyTech Nancy dans la formation par apprentissage 4A, spécialité « Informatique industrielle ».

Chaque année, en plus de mon service statutaire, je réalise une cinquantaine d'heures complémentaires. Je participe, comme tout enseignant de l'IUT, à d'autres missions : suivi de stages, forums étudiants, jurys de bac, *etc.*

Mes enseignements sont dispensés principalement en première et deuxième année de DUT (DUT1 et DUT2), à des groupes d'étudiants en formation initiale (FI) et formation en alternance (FA). Concernant mes enseignements en LP, les groupes sont « mélangés » : FI et contrats en alternance (contrat d'apprentissage et/ou de professionnalisation). Je liste ci-dessous les principales matières dans lesquelles je suis intervenu depuis 2006 ainsi que mes principales responsabilités pédagogiques :

- *Télécommunications :* acquisition et codage de l'information (DUT1), principes des transmissions radio (DUT1), chaine de transmission numérique (DUT1),
- *Signaux et systèmes :* initiation à la mesure du signal (DUT1), harmonisation des connaissances et des outils pour le signal (DUT1), mise à niveau en traitement du signal (LP), traitement statistique et big data (M2),
- *Mathématiques :* analyse de Fourier (DUT1), suites et séries numériques (DUT2), matrices et graphes (DUT2), algèbre linéaire en dimension finie (DUT2).

Pour l'ensemble de modules dispensés j'ai participé systématiquement à la mise en place, à la refonte et à l'évolution des cours magistraux, des travaux dirigés et des travaux pratiques.

Le Tableau 1.2 donne un aperçu général sur la répartition (en %) de ma charge annuelle d'enseignement entre les trois disciplines mentionnées ci-dessus au cours de l'année universitaire 2019-2020.

Enseignement	Niveau	Cursus	Heures			
			CM	CM-TD	TD	TP
Télécommunications	DUT1	FI	7%		18 %	8 %
		FA		12 %		
	DUT2	FI		10 %		
Signaux et systèmes	DUT1	FI	2 %		7 %	5 %
		FA			3 %	
	Bac+4	FA	14%			
	LP	FI & FA		4 %		
Mathématiques	DUT2	FI		8 %		2 %

TABLE 1.2 : Répartition approximative par niveau et par type de formation (FI/FA) de ma charge annuelle d'enseignement

Entre 2014 - 2017, j'ai assuré la responsabilité de la licence professionnelle R&T « Réseaux Sans Fil et Sécurité » (RSFS), pour laquelle j'ai porté également le projet de la nouvelle offre de formation 2018-2022. Dans le cadre de mes responsabilités, j'ai participé activement à la recherche des partenaires industriels pour les contrat en alternance et j'ai organisé le recrutement et la venue des intervenants extérieurs.

Depuis 2017 je suis responsable des stages pour les formations DUT et LP de notre département Réseaux et Télécommunications.

1.5 Activités de recherche

J'ai été initié au monde de la recherche lors de ma thèse de doctorat réalisée à Grenoble entre septembre 2002 et octobre 2005 au Laboratoire des Images et des Signaux (LIS), sous la direction de Jérôme Mars et Nicolas Le Bihan. Je me suis intéressé au problème de l'estimation conjointe de la direction d'arrivée (DDA) et des paramètres de polarisation des ondes enregistrées sur un réseau de capteurs vectoriels (multicomposantes[1]). Pendant ma thèse, ainsi que durant mon année d'ATER, j'ai eu la chance de collaborer avec des chercheurs tels que Cécile Cornu (Institut des Sciences de la Terre, Grenoble) ou Christophe Picard (PSA Peugeot Citroën, Vélizy-Villacoublay) et de m'ouvrir ainsi à des domaines différents comme la séismologie, la sismique, ou encore l'acoustique. Cela a été pour moi l'occasion de découvrir le milieu de la recherche industrielle en participant à des contrats de recherche avec les groupes Total (Pau) et PSA (Vélizy-Villacoublay).

[1]capteurs capables de décrire les oscillations d'un milieu/grandeur physique dans les trois directions de l'espace physique

J'ai intégré le Centre de Recherche en Automatique de Nancy (CRAN) le 1er septembre 2006 et je fais actuellement partie de l'équipe SiMul (Signaux Multidimensionnels) du département BioSiS (Biologie, Signaux et Systèmes en Cancérologie et Neurosciences). Mon activité de recherche au CRAN est axée sur le traitement des signaux multidimensionnels en utilisant principalement des méthodes d'algèbre multilinéaire (tensorielle). Cette activité couvre à la fois des aspects méthodologiques (modèles et algorithmes de décomposition tensoriels, unicité) et appliqués (traitement d'antenne, imagerie hyperspectrale, traitement de données de spectroscopie, etc.).

Les sections suivantes présentent brièvement mes principaux thèmes de recherche et résultats obtenus pendant et après ma thèse.

1.5.1 Travaux de thèse

J'ai soutenu ma thèse de doctorat intitulée : *Méthodes multilinéaires et hypercomplexes en traitement d'antenne multicomposante haute résolution*, le 25 octobre 2005 devant le jury suivant :

Président du jury	M. Michel CAMPILLO	- Pr. Université Grenoble 1
Rapporteur	M. Pierre COMON	- D.R. CNRS Valbonne
Rapporteur	M. Jean-Jacques FUCHS	- Pr. Université Rennes 1
Directeur de thèse	M. Jérôme MARS	- Pr. INPG
Co-encadrant	M. Nicolas LE BIHAN	- C.R. CNRS Grenoble
Examinateur	M. Stephen SANGWINE	- Pr. Univ. of Essex (U.K.)
Examinateur	M. Pascal CHEVALIER	- Expert Thales

Les travaux de recherche menés pendant ma thèse traitent du problème de l'estimation conjointe de la direction d'arrivée (DDA) et des paramètres de polarisation des ondes enregistrées sur un réseau de capteurs vectoriels (multicomposantes). Je me suis focalisé sur les méthodes de traitement d'antenne « à haute résolution », de type MUSIC, fondées sur la décomposition en sous-espaces propres de l'espace vectoriel engendré par les observations sur l'antenne.

L'originalité des méthodes proposées tient à l'utilisation des modèles mathématiques sortant du cadre de l'algèbre vectorielle classique, et qui se trouvent particulièrement bien adaptés à la nature des signaux multicomposantes. Deux approches différentes ont été proposées :

1.5.1.1 Approche tensorielle

Une première approche est fondée sur un modèle multilinéaire d'une onde polarisée enregistrée sur une antenne multicomposantes. À partir de ce modèle, nous avons construit une représentation multilinéaire des statistiques d'ordre deux des données multicomposantes, sous la forme d'un *tenseur interspectral (ou de covariance)*. Le tenseur interspectral conserve l'information sur la structure intrinsèquement multimodale de l'acquisition. Deux algorithmes

(*Vector-MUSIC* et *Higher Order-MUSIC*) ont été proposés, basés sur des décompositions orthogonales du tenseur interspectral. *Higher Order-MUSIC* s'est montré plus résolutif que *Vector-MUSIC* car il impose une contrainte d'orthogonalité plus forte entre les sous-espaces signal et bruit issus de la décomposition. Cependant, pour *Higher Order-MUSIC* le nombre maximal de sources détectables est fortement limité, car il est nécessairement inférieur au nombre des composantes de l'antenne.

Les méthodes proposées ont été validées par une publication dans une revue internationale ([A1]) et deux conférences internationales ([Ci2],[Ci3]) (voir Chapitre 2 : « Production scientifique »).

1.5.1.2 Approche hypercomplexe

La deuxième approche consiste à utiliser les nombres hypercomplexes pour modéliser les signaux multicomposantes. Nous avons proposé un modèle pour les signaux à deux composantes (2C) s'appuyant sur les *quaternions* et une modélisation des signaux 3C, 4C fondée sur l'utilisation des *biquaternions*. À partir de ces modèles, deux algorithmes hypercomplexes (*Quaternion-MUSIC* et *Biquaternion-MUSIC*) pour l'estimation des paramètres d'une onde polarisée, fondés sur la décomposition en sous-espaces orthogonaux, ont été proposés. J'ai développé également une méthode de décomposition en valeurs propres d'une matrice de biquaternions basée sur un nouvel objet mathématique : *la matrice quaternionique adjointe*.

Les résultats obtenus par cette approche ont également donné lieu à plusieurs publications ([A2], [A3], [Ci4], [Ci5], [Cn1]).

1.5.2 Travaux de recherche après la thèse

Dans le cadre de mes travaux de recherche au CRAN j'ai gardé comme fil conducteur, les *approches tensorielles* et les thématiques liées au *traitement d'antenne* et des *signaux polarisés*, auxquels j'ai commencé à m'intéresser pendant ma thèse. En parallèle, les interactions avec les chercheurs du CRAN et avec l'environnement scientifique nancéien de façon générale, m'ont permis d'élargir mes centres d'intérêt et de travailler sur des problématiques telles que les *décompositions des signaux multidimensionnels (N-D) sous contraintes*. Dans ce cadre, je me suis intéressé notamment aux modèles de mélange de signaux à valeurs *non-négatives* et *binaires*, et aux algorithmes de démélange d'images hyperspectrales avec des contraintes de type *temps-réel*. Ma récente collaboration avec Julien Flamant (CR recruté au CRAN en 2019) a été l'occasion de reprendre la thématique « signaux à valeurs quaternionique » abordée pendant ma thèse.

1.5.2.1 Traitement d'antenne multidimensionnelle/ multicomposante

Ces travaux s'inscrivent dans la continuité de mon travail de thèse et s'articulent principalement autour de la thèse de Xijing Guo ([D.1], voir Section 1.6 « Activités d'encadrement »), de la collaboration avec l'université de Xi'an, ainsi que de la collaboration mise en place avec

Kainam Thomas Wong de Polytechnic University of Hong-Kong (PUHK). Les contributions portent sur plusieurs volets, présentés succinctement ci-dessous.

Méthodes et modèles tensoriels en traitement d'antenne Ce travail regroupe un ensemble d'algorithmes et de modèles fondés sur des approches tensorielles pour l'estimation des directions d'arrivées (DDAs) des sources incidentes sur une antenne. L'idée derrière ces méthodes est d'exploiter la structure multilinéaire des données enregistrées sur certains types d'antennes, ainsi que les propriétés d'unicité des décompositions tensorielles, afin de proposer des algorithmes pour l'estimation des DDAs, plus efficients et précis par rapport aux algorithmes classiques de type MUSIC/ ESPRIT.

Pour les antennes vectorielles (permettant d'enregistrer les oscillations de la grandeur d'intérêt suivant plusieurs directions de l'espace), nous avons montré que les données enregistrées présentent une structure de type *Canonical Polyadic Decomposition (CPD)* [Ci6, Ci8]. Nous avons introduit un algorithme d'estimation basé sur ce modèle qui, contrairement aux approches de type MUSIC, ne nécessite pas une recherche exhaustive des DDAs et permet d'estimer conjointement les DDAs et les polarisations des ondes. Nous avons également proposé des bornes sur les performances de ce type de méthodes, en termes d'erreur d'estimation et de nombre maximum de sources détectables [A6, A8, A9, Ci7, Cn3].

Pour le même type d'antenne, nous avons proposé une approche alternative basée sur une décomposition multidimensionnelle des données en sous-espaces orthogonaux *signal / bruit*, de type *Higher-Order Singular Value Decomposition (HOSVD)* [Ci18]. Grâce à l'orthogonalité multi-modale de la décomposition, l'algorithme proposé, appelé *Tensor-MUSIC*, présente des performances supérieures aux méthodes classiques fondées sur des décompositions matricielles des données.

Dans le but d'exploiter de manière efficace la diversité spatiale d'une antenne, j'ai introduit une nouvelle structure hiérarchique, de type multi-échelle, pour les antennes scalaires, avec un algorithme tensoriel pour l'estimation des DDAs [A12, Ci19]. Comparé à des méthodes similaires de la literature, cet algorithme s'est montré plus performant, particulièrement dans des scénarios difficiles avec un faible rapport signal à bruit et un faible nombre d'échantillons temporels. Un deuxième algorithme pour les antennes multi-échelles, fondé sur l'approximation parcimonieuse et qui exploite la structure Khatri-Rao du dictionnaire, a été également proposé [Ci22, Cn7]. Nous avons montré que cette approche permet de diminuer la charge de calcul, et dans certains cas difficiles, donne de meilleurs résultats que l'approche tensorielle. La problématique des antennes multi-échelle rejoigne les recherches très récentes concernant les antennes de très grande dimension. Un exemple dans ce sens est le « Vaste Réseau d'antennes Sub-Millimétriques de l'Atacama », déployé dans les Andes chiliennes par l'Observatoire européen austral. Avec une envergure de 15 km il permet d'atteindre une résolution angulaire de 35 millisecondes d'arc, près de 2 fois mieux que le télescope spatial Hubble.

Traitement d'antenne vectorielle non-colocalisée Un autre aspect de mes travaux sur l'estimation des DDAs est représenté par les antennes vectorielles non-colocalisées, des antennes électromagnétiques avec les dipôles et les boucles réparties dans l'espace. L'avantage principal de cette configuration est la diminution considérable des couplages entre les composantes de l'antenne. Cette configuration permet également d'augmenter l'ouverture physique de l'antenne, ce qui améliore les performances de la détection. L'utilisation des capteurs non-colocalisés nécessite cependant l'adaptation des algorithmes d'estimation de la DDA conçus pour les capteurs colocalisés. Nous avons proposé un algorithme pour estimer les DDAs des sources polarisées incidentes sur ce type d'antenne [Ci17, Cn6]. Il est fondé sur la minimisation d'une fonction coût non-linéaire. La solution exacte est obtenue de façon efficace, grâce à une transformation linéaire des paramètres d'intérêt, et à une recherche sur une grille discrète. Nous avons, par ailleurs, étudié et proposé des conditions d'identifiabilité pour ce modèle, ainsi que des configurations géométriques respectant ces conditions.

Traitement d'antenne acoustique Ce volet de mes recherches regroupe les résultats des travaux développés dans la cadre du contrat avec PSA Peugeot-Citröen (Vélizy-Villacoublay) et les résultats de la collaboration avec Northwestern Polytechnical University (Chine).

La collaboration avec PSA visait à améliorer l'environnement du chauffeur et de ses passagers dans l'habitacle de la voiture, en limitant notamment les nuisances sonores dues à l'écoulement de l'air à la surface du véhicule. Dans ce contexte, nous avons étudié la séparabilité des composantes d'origine aéroacoustique d'une part, et aérodynamique d'autre part, présentes dans un champ de pression pariétale (au contact d'une paroi). Les deux composantes étant enregistrées sur un réseau de capteurs, nous avons montré que les méthodes classiques de filtrage de type $FK-3D$ sont inefficaces dans les basses fréquences pour séparer les deux composantes. Nous avons proposé deux méthodes de filtrage matriciel multicomposante ($3D-SVD$ notamment) permettant d'obtenir une séparabilité de ces deux composantes après une transformation spatiale des données [A4, Cn2].

En 2015, en collaboration avec mon ancien doctorant Xijing Guo (voir Section 1.6) et ses collègues de Northwestern Polytechnical University, on a initié une collaboration sur les antennes acoustiques *superdirectives* et sur les méthodes de *formation de voies* (*beamforming*), pour ce type d'antennes. Une antenne superdirective peut être vue comme une réalisation du concept de « capteur acoustique d'ordre supérieur », *i. e.*, un capteur capable de mesurer les dérivées spatiales partielles d'ordre $N \geq 2$, du champ de pression acoustique. Par conséquent, le champ acoustique incident peut être décomposé en une série de composantes, appelées *modes*. Dans la pratique, ces modes sont estimés de façon soustractive, à partir des mesures effectuées sur une antenne de capteurs acoustiques (pression / vélocité) avec une configuration particulière. Cela permet de construire des antennes super-directives de très petites dimensions par rapport à la longueur d'onde. Dans [A13], nous avons proposé une méthode de *beamforming* dans le domaine modal pour une antenne rectangulaire uniforme

3×3 de capteurs vectoriels acoustiques bidimensionnels ; l'espacement entre les capteurs est beaucoup plus petit que les longueurs d'onde dans la bande de fréquences de travail. Les modes acoustiques ont été extraits à partir des mesures de vélocité des particules, et ont été ensuite combinés pour obtenir le *beamformer* souhaité. La méthode proposée est limitée à un ordre maximum $N = 5$. Les performances de l'antenne ont été calculées analytiquement et validées lors d'une campagne d'acquisition dans le lac Qiandao (Chine).

Un problème bien connu des antennes super-directives est leur faible *gain de bruit blanc* (WNG[2]), qui mesure la capacité de l'antenne à rejeter le bruit spatialement non corrélé, dû à certaines causes non-acoustiques (le bruit propre du capteur). Par conséquent, le gain de l'antenne peut être faible, malgré un indice de directivité (ID) élevé. Pour maximiser le gain de l'antenne, un compromis entre l'ID et le WNG est nécessaire. Nous avons fourni dans [A14] une limite supérieure pour l'ID afin d'obtenir le meilleur gain d'antenne possible, pour une famille d'antennes miniaturisées pour les applications sous-marines. La formation superdirective de voies est sensible aux distorsions des caractéristiques du capteur, *i. e.,* le gain, la phase et la position. Nous avons proposé dans [A15] une méthode de régularisation probabiliste pour rendre la formation superdirective de voies robuste aux distorsions des caractéristiques des capteurs. L'objectif est d'augmenter la directivité tout en assurant une robustesse avec une probabilité élevée. Le problème de régularisation probabiliste est résolu par un algorithme SOCP (*Second-Order Cone Programming*), où le paramètre de régularisation est choisi par une analyse statistique des perturbations du système, basée sur des simulations de Monte Carlo. Les expériences ont été réalisées sur une antenne rectangulaire uniforme miniaturisée de dimension 3×3, sans calibration. Les résultats montrent que la méthode proposée est robuste aux distorsions des paramètres des capteurs et permet d'obtenir un niveau de directivité plus élevé par rapport à d'autres approches robustes de formation de voies adaptative.

1.5.2.2 Analyse de mélanges multidimensionnels sous contraintes

Mes travaux de recherche menés dans ce cadre se déclinent en plusieurs sous-thèmes :

Modèles et décompositions tensorielles pour les signaux en spectroscopie et imagerie Les contributions présentées dans cette partie émanent des collaborations avec les chercheurs des laboratoires LIEC (Laboratoire Interdisciplinaire des Environnements Continentaux), LCPME (Laboratoire de Chimie Physique et Microbiologie pour les Matériaux et l'Environnement), Laboratoire Matériaux Optiques, Photonique et Systèmes (LMOPS) et de l'équipe STICMO (Cibles moléculaires dans une démarche translationnelle) du CRAN.

Un premier volet de ces travaux s'articule autour de la thèse de Fabrice Caland ([D.2]) réalisée en collaboration avec le LIEC. Ces recherches avaient comme objectif d'identifier, à l'aide d'outils mathématiques d'algèbre multilinéaire, les réponses de bio-senseurs bactériens

[2] *White Noise Gain*, en anglais

fluorescents dans des conditions environnementales variées, qu'il s'agisse d'un stress engendré par la présence à forte dose d'un métal ou d'une carence nutritive engendrée par son absence. Cette identification est fondée sur l'analyse quantitative à l'échelle d'une population bactérienne de signaux multidimensionnels. Elle repose en particulier sur (i) l'acquisition de données spectrales (fluorescence) multivariées sur des suspensions de biosenseurs multicolores interagissant avec des métaux, et sur (ii) le développement d'algorithmes de décompositions tensorielles. Les méthodes proposées et développées ont permis d'identifier sans *a priori*, la réponse fonctionnelle de bio-senseurs sous différentes conditions environnementales. Elles tirent parti de la variabilité des réponses systémiques et permettent de déterminer les sources élémentaires, caractérisant le système, ainsi que leur comportement en fonction des paramètres extérieurs. Ces approches sont inspirées des méthodes de type CPD, et CPD avec dépendances linéaires dans les modes (Paralind/Confac). L'avantage de ce type d'approche, par rapport aux approches classiques, est l'identification unique des réponses des bio-senseurs sous faibles contraintes. Ces méthodes ont permis d'améliorer l'identification des spectres de fluorescence et des réponses des bio-senseurs, en optimisant les mesures physiques par l'utilisation de spectres synchrones ou en apportant une diversité suffisante aux plans d'expériences. L'usage des spectres synchrones s'est avéré déterminant, à la fois pour améliorer la séparation des sources de fluorescence, mais aussi pour augmenter le rapport signal sur bruit des bio-senseurs les plus faibles. Cette méthode d'analyse spectrale originale a permis d'élargir fortement la gamme chromatique des bio-senseurs fluorescents multicolores utilisables simultanément. Par conséquent, une nouvelle méthode d'estimation de la concentration de polluants métalliques présents dans un échantillon, à partir de la réponse spectrale d'un mélange de bio-senseurs non-spécifiques, a pu être développée. Les résultats de ces recherches ont été validés par des publications internationales et nationales ([A11], [Ci15], [Ci14], [Ci12], [Ci9], [Cn5]).

Un deuxième volet concerne l'analyse des signaux à diversité multiple en spectroscopie Raman et cytométrie en flux. L'analyse de la polarisation en spectroscopie Raman constitue un outil puissant en physique chimique pour identifier la symétrie des modes de vibration des molécules, macromolécules et cristaux, en fonction de leur structure cristallographique et de leur orientation spatiale locale. Nous avons proposé ([A5], [O2], [Ci10]) deux nouvelles approches pour l'analyse des signaux en spectroscopie Raman polarisée pour les données à diversité de rotation et spatiale. Nous avons montré que l'utilisation conjointe de la polarisation parallèle et croisée donne des résultats plus précis de séparation des modes vibrationnels des molécules, et améliore les propriétés d'unicité de la solution pour les algorithmes de séparation de sources non-négatives. Les approches proposées ont été validées sur deux ensembles de données Raman polarisées réelles. Une étude de faisabilité pour un nouveau système de microscopie Raman a été également réalisée dans [Ci13]. En utilisant la lumière polarisée pour l'analyse Raman de l'échantillon étudié, ainsi que les diversités spatiale, de rotation et spectrale, nous avons démontré la possibilité de séparer les contributions des différents micro-cristaux présents sur l'échantillon et à estimer leurs orientations spatiales. L'approche est basée sur une CPD d'ordre trois des données acquises et a été validée sur des simulations

numériques réalistes.

Nous avons proposé également une approche tensorielle pour l'analyse des données de cytométrie ([A10], [Ci20]). La cytométrie en flux est une technique d'investigation largement utilisée en biologie et en médecine pour la caractérisation et la quantification de la densité, de la fluorescence et des propriétés morphologiques des cellules. Une approche souvent utilisée dans les études biomédicales consiste à étudier la réponse d'une population cellulaire par rapport à des paramètres environnementaux/ chimiques/ biologiques. Nous avons développé une nouvelle approche pour le traitement des séquences de données de cytométrie en flux. Elle consiste à décomposer une séquence de densités de probabilités multidimensionnelles en utilisant la décomposition en tenseurs blocs. Pour illustrer l'efficacité de l'approche, une étude du potentiel de la membrane mitochondriale de la lignée cellulaire T47D, en fonction de la concentration d'un agent de découplage, a été réalisée. Nous avons montré que la capacité de tri des cellules est considérablement améliorée par rapport aux méthodes classiques. La méthode proposée a permis d'observer directement la dépendance du potentiel de la membrane mitochondriale de la phase du cycle cellulaire.

Décomposition des tableaux multidimensionnels de données binaires Les données binaires sont de plus en plus présentes dans divers domaines. Elles peuvent être soit générées par des processus/phénomènes à deux états (et naturellement encodées sur les valeurs binaires $0, 1$), *e.g.* capteurs de proximité, boutons poussoirs, interrupteurs électriques, réponses oui/non dans les enquêtes, *etc.*, soit obtenues artificiellement, par la quantification sur 1 bit, des données à valeurs réelles. C'est la raison pour laquelle, la factorisation/ décomposition des tableaux à valeurs binaires a récement suscité beaucoup d'intérêt, et a déjà été utilisée avec succès dans de nombreux domaines tels que les télécommunications, l'identification des phénotypes des gènes, la classification des documents, les systèmes de recommandation, *etc.* Je me suis penché sur ce problème dans le cadre de la thèse de Mamadou Diop ([D.3]) réalisée en collaboration avec le CEA-Saclay. Nous avons proposé ([A19], [Ci23], [Cn8]) une nouvelle approche pour la *factorisation booléenne en matrices binaires* fondée sur un modèle de mélange *post non-linéaire*. Contrairement aux méthodes de factorisation de matrices binaires existantes, fondées sur le produit matriciel classique, le modèle proposé est équivalent au modèle booléen de factorisation matricielle lorsque les entrées des facteurs sont exactement binaires. Ainsi, comparée aux approches existantes, cette approche permet d'obtenir des résultats plus interprétables dans le cas de sources binaires corrélées, et avec des rangs d'approximation matricielle plus faibles. Nous avons démontré également un ensemble de conditions d'unicité pour ce type de décomposition. Deux algorithmes s'appuyant sur des mises à jour multiplicatives des facteurs, ont été proposés. Les performances de cette approche ont été illustrées, entre autre, sur les données Netflix prize [Ci28]. Nous avons proposé une généralisation de cette approche au cas de tableaux multidimensionnels (tenseurs) binaires. Un algorithme efficace pour la décomposition des tenseurs à valeurs binaires, s'appuyant sur une stratégie de

type AO-ADMM (*Alternating Optimization -ADMM*[3]), a été développé dans ([Ci27]).

Factorisations non-négatives, sous contraintes, en imagerie hyperspectrale Les travaux décrits dans cette partie concernent la séparation de sources en imagerie hyperspectrale, avec des contraintes liées au processus d'acquisition et à la nature des signaux enregistrés.

Un premier problème est issu du projet ANR OPTIFIN (OPTImisation des FINitions) qui avait pour objectif le développement de méthodes d'analyse permettant de prédire et de classifier la qualité des rendus de pièces de bois en fonction d'une analyse de surface en temps réel. Nous avons abordé dans le cadre de la thèse de Ludivine Nus ([D.4]), le problème de démélange en-ligne d'images hyperspectrales acquises par un imageur *pushbroom*[4], pour la caractérisation en temps réel du matériau bois. Nous avons proposé un modèle de mélange en-ligne s'appuyant sur la factorisation en matrices non-négatives. A partir de ce modèle, trois algorithmes pour le démélange séquentiel en-ligne, fondés respectivement sur les règles de mise à jour multiplicatives, le gradient optimal de Nesterov et l'optimisation ADMM ont été développés ([A17], [Ci25], [Ci24], [Cn11]). Ces algorithmes sont spécialement conçus pour réaliser le démélange en temps réel, au rythme d'acquisition de l'imageur *pushbroom*. Afin de régulariser le problème d'estimation (généralement mal posé), deux sortes de contraintes sur les endmembers ont été utilisées : une contrainte de dispersion minimale ainsi qu'une contrainte de volume minimal. Une méthode pour l'estimation automatique du paramètre de régularisation a été également proposée ([Ci26]), en reformulant le problème de démélange hyperspectral en-ligne comme un problème d'optimisation bi-objectif. Nous avons aussi proposé dans [Ci31] une approche permettant de gérer la variation du nombre de sources, *i. e.* le rang de la décomposition, au cours du traitement. Les algorithmes en-ligne préalablement développés ont été modifiés, en introduisant une étape d'apprentissage d'une bibliothèque hyperspectrale, ainsi que des pénalités de parcimonie, permettant de sélectionner uniquement les sources actives. Ces algorithmes ont été appliqués sur des données réelles, à la détection et à la classification des singularités du matériau bois.

Un deuxième problème est représenté par le démélange d'images hyperpsectrales polarisées. Nous avons développé ([A16], [Ci30], [Cn10]) dans le cadre du stage postdoctoral de Julien Flamant ([pD.2]), un nouvel outil particulièrement bien adapté à cette tâche : la factorisation non-négative de matrices à valeurs quaternioniques (QNMF[5]). Cet outil étend au cas des signaux polarisés le concept de factorisation en matrices non-négatives (NMF), très utilisé en démélange hyperspectral. L'information sur la polarisation est représentée par les paramètres de Stokes, un ensemble de 4 paramètres énergétiques largement utilisés dans

[3] *Alternating Direction Method of Multipliers*

[4] Technique permettant d'acquérir les images hyperspectrales tranche par tranche, en deplasant la caméra au-dessus de l'échantillon analysé.

[5] *Quaternion Non-negative Matrix Factorization*, en anglais.

l'imagerie polarimétrique. La QNMF repose sur deux ingrédients clés : (i) la représentation algébrique des paramètres de Stokes grâce aux quaternions et (ii) l'exploitation des contraintes physiques sur les paramètres de Stokes. Ces contraintes généralisent la non-négativité au cas des signaux polarisés, codant le fait que la matrice de covariance associée à chaque source est semi-définie positive. Des conditions d'unicité pour la QNMF sont également données, et on montre qu'elles englobent les conditions suffisantes d'unicité connues pour la NMF. De surcroît, en exploitant l'information de polarisation, la QNMF assouplit les conditions d'unicité de la NMF qui exigent que les sources présentent une certaine configuration de zéros. Nous introduisons un algorithme simple mais efficace, pour résoudre le problème QNMF en pratique. Des mises à jour sous forme explicite sont obtenues, en utilisant des résultats récents sur le calcul quaternionique. Nous avons montré la pertinence de l'approche sur des données synthétiques réalistes.

Identifiabilité des modèles tensoriels de mélange avec dépendances linéaires dans les modes L'identifiabilité d'un modèle de mélange (ou l'unicité de la décomposition associée) est un problème fondamental dans les méthodes d'analyse de données, car il permet de garantir l'interprétabilité des résultats obtenus. Cela a été une préoccupation constante durant mes recherches, et transversale à toutes les problématiques que j'ai abordées. L'un des avantages principaux des modèles tensoriels est leur identifiabilité sous de faibles contraintes, liées principalement au rang des matrices des facteurs (modes). Lorsque les vecteurs dans les modes présentent de dépendances linéaires, l'unicité de la décomposition n'est plus garantie, d'où le besoin de nouveaux résultats d'identifiabilité, plus spécifiques.

Nous avons démontré, dans le contexte du traitement d'antenne, une condition suffisante pour l'unicité de la CPD d'ordre 3, dans le cas particulier où l'une des trois matrices de facteurs est de rang plein. Cette condition est plus relaxée que les conditions similaires existantes dans la literature. Sur cette base, on a proposé une borne supérieure sur le nombre maximal de sources polarisées détectables à l'aide d'une antenne vectorielle à six composantes. Ces résultats ont fait l'objet de plusieurs publications [A6], [Ci7], [Cn4].

Dans [A7] et [Cn4], nous avons introduit trois conditions suffisantes, garantissant l'unicité d'une des matrices des facteurs pour la CPD d'ordre trois. Sur la base de ces conditions, un résultat d'unicité partielle a été proposé, qui autorise des facteurs colinéaires dans un seul mode. Nous avons montré que, si on assure l'unicité de l'un des modes, alors le modèle CPD initial peut être décomposé de manière unique en somme des tenseurs de rang inférieur, pour lesquels, l'identifiabilité peut être évaluée de façon indépendante. Cette condition est plus simple et plus facile à vérifier que d'autres conditions similaires existant dans la littérature. Nous avons également examiné l'unicité du modèle CPD d'ordre 4, dans le cas où les seules dépendances linéaires possibles entre les colonnes des matrices des facteurs sont les colinéarités [Ci12]. Pour cette configuration spéciale, nous démontrons une condition nécessaire et suffisante garantissant le rang de colonne plein du produit de Khatri-Rao de deux matrices. A partir de ce résultat, nous avons proposé une condition suffisante pour l'unicité des CPD

d'ordre 4 avec colinéarités dans maximum trois modes. Ce résultat est illustré par l'analyse des données de fluorescence présentant quatre diversités. Dans [Ci21], nous avons étudié l'identifiabilité du modèle Paralind d'ordre 3 avec des matrices d'interaction parcimonieuses (S-Paralind). Nous avons fourni des résultats théoriques sur la manière d'obtenir les matrices d'interaction les plus parcimonieuses possible, dans certaines configurations particulières. Ces résultats pourraient être utilisés dans la conception et l'analyse d'algorithmes décomposition basés sur la norme ℓ_0.

Avec le développement des capteurs et des technologies de mesure, les données collectées sont intrinsèquement multidimensionnelles dans un grand nombre d'applications. Cela peut être interprété comme une croissance de la dimensionnalité/ordre du tenseur associé. Dans ce contexte, je me suis intéressé récemment à l'identifiabilité d'une nouvelle classe de modèles tensoriels, le *train de tenseurs (TT)*, et plus particulièrement aux modèles TT-CPD. Il s'agit des représentations hiérarchiques des modèles CPD d'ordre élevé. Nous avons proposé dans [Ci29], [Cn9] des conditions suffisantes d'identifiabilité pour ces modèles, dans le cas où les matrices des facteurs présentent des dépendances linéaires. Nous avons comparé ces conditions avec les conditions existantes dans la littérature pour les modèles CPD d'ordre élevé, et avons montré leurs avantages et inconvénients.

Une partie des travaux présentés dans cette section sera détaillée dans la Partie II du manuscrit. Cependant, pour des raisons de lisibilité, l'organisation des thématiques scientifiques dans la Partie II sera différente par rapport à la présente section.

1.6 Activités d'encadrement

1.6.1 Thèses de doctorat

[D.4] **Ludivine Nus** (2016-2019) : *Méthodes rapides de traitement d'images hyperspectrales. Application à la caractérisation en temps réel du matériau bois.*

- Financement : contrat ANR OPTIFIN. Thèse débutée en octobre 2016 et soutenue en décembre 2019.
- Encadrement : David Brie (50%), **Sebastian Miron** (50%).
- Publications : [A17, Ci31, Ci26, Ci25, Ci24, Cn11].

[D.3] **Mamadou Diop** (2016-2018) : *Décomposition booléenne des tableaux multi-dimensionnels de données binaires : une approche par modèle de mélange post non-linéaire.*

- Financement : contrat CEA-Tech / Région Lorraine. Thèse débutée en janvier 2016 et soutenue en décembre 2018.
- Encadrement : David Brie (50%), **Sebastian Miron** (50%).

- Publications : [A19, Ci28, Ci27, Cn8, Cn5].

[D.2] **Fabrice Caland** (2009-2013) : *Décomposition tensorielle de signaux luminescents émis par des biosenseurs pour l'identification de Systèmes Minéral-Bactérie.*

- Financement : CNRS/Région Lorraine (ANR HAESPRI). Thèse débutée en octobre 2009 et soutenue en septembre 2013.
- Encadrement : David Brie (34%), **Sebastian Miron** (33%), Christian Mustin (33%, LIMOS).
- Publications : [Ci15, Ci14, Ci12, Ci9, Cn5].

[D.1] **Xijing Guo** (2007-2012) : *Analyse multilinéaire en traitement d'antenne vectorielle.*

- Financement : bourse de l' Ecole doctorale franco-chinoise.Thèse en co-tutelle avec Xi'an Jiaotong University, débutée en janvier 2007 et soutenue en mars 2012 en Chine.
- Encadrement : David Brie (30%), **Sebastian Miron** (50%), Shihua Zhu (20%, Xi'an Jiaotong University).
- Publications : [A8, A7, A6, O1, Ci11, Ci8, Ci7, Ci6, Cn4, Cn3].

1.6.2 Post-doctorats

[pD.3] **Yassine Zniyed** (2019-2020) : *Apprentissage des réseaux de neurones par les méthodes tensorielles.*

- Financement : contrat ANR LeaFleT.
- Encadrement : Konstantin Usevich (34%), **Sebastian Miron** (33%), David Brie (33%).
- Publications : [Ci32].

[pD.2] **Julien Flamant** (2018-2019) : *Apport du formalisme quaternionique au démélange d'images hyperspectrales polarisées.*

- Financement : CRAN.
- Encadrement : David Brie (50%), **Sebastian Miron** (50%).
- Publications : [A16, Ci30, Cn10].

[pD.1] **Marian Dragomir** (mars-mai 2010) : *Design of adaptive algorithms for default detection on electric power transmission lines.*

- Financement : contrat BRAIN (Université « Gh. Asachi », Iasi, Roumanie).
- Encadrement : **Sebastian Miron** (50%), Mircea Guşă (50%, Université « Gh. Asachi », Iaşi, Roumanie).

1.6.3 Ingénieur d'études (CDD)

[IE.1] **Eddy Robin** (2018-2019) : *Réalisation d'un démonstrateur pour la séparation des sources binaires.*

- Financement : contrat CEA-Tech / Région Lorraine.
- Encadrement : **Sebastian Miron** (50%), David Brie (50%).
- Publications : [A19].

1.6.4 Stages de master 2

[S.6] **Maxime Desplanches.** *Fusion de données (géolocalisation, scanner 3D, etc.) pour la cartographie spatio-temporelle des polluants dans l'air des lieux de travail.* Master Méthodes et Applications pour le Traitement du Signal et de l'Image ; CRAN, Université de Lorraine - INRS, Nancy, 2021.

[S.5] **Tarek Nsiri.** *Compression des couches convolutives dans les réseaux profonds de neurones, implémentation sous PyTorch.* Master Ingénierie des Systèmes Numériques ; CRAN, Université de Lorraine, Nancy, 2020.

[S.4] **Philippe Flores.** *Analyse des données issues d'instruments de mesure en temps réel pour la cartographie spatio-temporelle des polluants dans l'air des lieux de travail.* Master Méthodes et Applications pour le Traitement du Signal et de l'Image ; CRAN, Université de Lorraine - INRS, Nancy, 2020.

[S.3] **Marc-Abel Bisch.** *Estimation des directions d'arrivées avec un réseau d'antennes utilisant des méthodes parcimonieuses.* Master Ingénierie de Systèmes Complexes ; CRAN, Université de Lorraine, 2014. Publications : [Ci22, Cn7].

[S.2] **Yazid Merah.** *Traitement d'antenne vectorielle non-colocalisée.* Master Ingénierie de Systèmes Complexes ; CRAN, Université de Lorraine, 2012. Publications : [Ci17].

[S.1] **Fabrice Caland.** *Décomposition tensorielle de données de fluoréscence pour l'analyse de système bactérien.* Master Ingénierie de Systèmes Complexes ; CRAN, Université de Lorraine, 2009. Publications : [Ci9].

1.6.5 Projet de master 1

[P.6] **Vladislav Garabă.** *Tensor methods for compressing neural network.* Master 1 Computers and Information Technology, Université « V. Alecsandri », Bacău, Roumanie ; CRAN, Université de Lorraine, Erasmus 2020.

[P.5] **Andrada Leca.** *Segmentation de textures sur des images d'échantillons de bois.* Master 1 Computers and Information Technology, Université « V. Alecsandri », Bacău, Roumanie ; CRAN, Université de Lorraine, Erasmus 2019.

[P.4] **Octavian Butură.** *Classification algorithms for hyperspectral wood imaging system.* Master 1 Computers and Information Technology, Université « V. Alecsandri », Bacău, Roumanie ; CRAN, Université de Lorraine, Erasmus 2018.

[P.3] **Mihail Țârdea.** *Implémentation d'une approche « greedy » pour l'estimation de la direction d'arrivée d'une source électromagnétique avec une antenne vectorielle non-colocalisée.* Master 1 Computers and Information Technology, Université « V. Alecsandri », Bacău, Roumanie ; CRAN, Université de Lorraine, Erasmus 2016.

[P.2] **Ionuț Ștefănescu.** *Une approche par clustering pour les décompositions de type Block-PARAFAC d'images hyperspectrales.* Master 1 Computers and Information Technology, Université « V. Alecsandri », Bacău, Roumanie ; CRAN, Université de Lorraine, Erasmus 2015.

[P.1] **Ionuț Dumitrescu.** *Block-PARAFAC non-negative decomposition of hyperspectral images.* Master 1 Computers and Information Technology, Université « V. Alecsandri », Bacău, Roumanie ; CRAN, Université de Lorraine, Erasmus 2014.

1.7 Contrats et projets

• Projets internationaux :

1.7.1 A new paradigm for geophysical imaging

- Financement : PHC Procore (France - Hong Kong)
- Durée : 2 ans (2011 - 2013)
- Responsable pour la partie française : **Sebastian Miron**
- Autre participants : Kainam Thomas Wong (Hong Kong Polytechnic University), David Brie
- Budget (part CRAN) : 8k€

Résumé : La collaboration mise en place dans le cadre de ce projet a permis de développer des méthodes et résultats originaux relatifs à la conception d'antennes et au développement des algorithmes de traitement associés. Plus spécifiquement, nous avons proposé une nouvelle configuration d'antenne présentant des invariances spatiales d'échelle. Cette configuration géométrique est particulièrement bien adaptée au développement d'antennes acoustiques utilisées en géophysique et permet de représenter les données reçues sous la formes d'un tenseur admettant une décomposition multilinéaire CPD. Ses avantages par rapport aux antennes classiques sont multiples. Elle permet d'enregistrer des signaux source utilisant différentes

résolutions spatiales. Un algorithme efficace pour le traitement des signaux issus de ce type d'antenne, qui exploite sa géométrie particulière, a été également proposé. Ainsi, il est possible d'estimer les directions d'arrivée des ondes enregistrées, à partir d'un seul échantillon temporel, ce qui n'est pas le cas avec les approches classiques. Nous avons également montré que ce type d'antenne est beaucoup moins sensible aux problèmes liés aux trajets multiples (réflexions multiples) des ondes.

• Projets nationaux évalués :

1.7.2 Learning neural networks with flexible nonlinearities by tensor methods

- Financement : ANR JCJC LeaFleT
- Durée : 3 ans (2019 - 2022)
- Porteur : Konstantin Usevich (CRAN)
- Autre chercheurs : David Brie, **Sebastian Miron**, Pierre Comon, Simon Barthelmé (Gipsa-Lab), Christophe Cerisara (LORIA), Mariya Ishteva (Vrije Universiteit Brussel), Caroline Chaux (Institut de Mathématiques de Marseille)
- Budget : 207 k€

Résumé : Les progrès récents dans l'apprentissage automatique sont liés aux réseaux de neurones profonds, qui sont basés sur une structure multicouche du perceptron, avec un nombre extrêmement important de couches. Un inconvénient majeur des réseaux de neurones est le manque d'interprétabilité des résultats, aspect essentiel pour prendre des décisions crédibles fondées sur l'apprentissage profond. Dans ce projet, nous proposons d'utiliser des réseaux de neurones avec des fonctions d'activation non linéaires flexibles (par opposition aux fonctions fixes, comme les sigmoïdes ou les unités ReLU). Ces modèles ont été utilisés avec succès dans l'identification de systèmes dynamiques non-linéaires structurés par bloc. L'ingrédient clé de notre approche est l'utilisation du lien entre ces architectures et les décompositions tensorielles. Cela permettrait non seulement d'utiliser de puissants outils de calcul pour l'apprentissage, mais aussi d'utiliser les propriétés d'identifiabilité des décompositions tensorielles, qui pourraient ouvrir la voie à des modèles interprétables de réseaux de neurones.

1.7.3 Développement d'un procédé de classification et d'analyse permettant de prédire le rendu des finitions sur bois

- Financement : ANR OPTIFIN

- Partenaires : CRITT Bois, CRAN, LCPME, BERRY WOOD, OGF Industrie
- Durée : 4 ans (2015 - 2019)
- Porteur : Eric Masson (CRITT Bois)
- Autre chercheurs : David Brie, **Sebastian Miron**, Vincent Bombardier (CRAN), Cédric Carteret, André Erwan (LCPME)
- Budget (part CRAN) : 270 k€

Résumé : Le bois est un matériau dont le rendu naturel est apprécié dans de nombreuses applications (parquets, ameublement, agencement intérieur ou extérieur, *etc.*). Cependant son hétérogénéité et sa variabilité de surface rendent souvent problématique la maitrise et l'harmonie des rendus, générant parfois des non-qualités. Ce rendu et cette harmonie sont majoritairement regardés au travers d'une finition appliquée sur le bois (vernis, peinture, *etc.*). La qualité de la finition et le rendu final sont impactés par le couple bois-finition. La texture physique, l'acidité, la présence de molécules pouvant migrer ou réagir avec la finition sont des facteurs potentiellement impactants. Le projet OPTIFIN a pour objectif le développement d'une méthode d'analyse permettant de prédire et de classifier la qualité des rendus en fonction d'une analyse de surface (voire dans une faible épaisseur de surface) en temps réel, du bois. Cette analyse de surface nécessite le couplage de plusieurs techniques d'acquisition utilisables en industrie (caméra simple spectre, caméra hyperspectrale, polarisation, télémétrie, autre spectrométrie, *etc.*) ainsi que le couplage de plusieurs méthodes de traitement du signal (méthodes de séparation de signaux classiques ou récursives, classification, *etc.*).

1.7.4 Méthodes de factorisation en matrices non-négatives pour l'apprentissage sur des données binaires

- Financement : Feder, région Lorraine
- Partenaires : CEA LIST
- Durée : 3 ans (2016 - 2018)
- Porteur : David Brie (CRAN)
- Autre chercheurs : **Sebastian Miron**, Anthony Larue, Antoine Souloumiac (CEA LIST),
- Budget (part CRAN) : 150 k€ (dont 1 bourse de thèse, 6 mois d'ingénieur d'études)

Résumé : Le cadre général du projet est l'apprentissage de dictionnaire sur des bases de données binaires par des techniques de factorisation en matrices non-négatives (NMF). C'est un problème générique qui a un grand nombre d'applications potentielles. A titre d'exemple, on peut mentionner la surveillance énergétique de bâtiments où les séquences de commutation (interrupteurs, chauffage, eau, *etc.*) sont stockées dans une base de données. Il s'agit de retrouver dans cette base de données des profils d'utilisation et, le cas échéant, d'identifier des comportements « anormaux ». Un second exemple est relatif à la surveillance des systèmes industriels de grande dimension. Dans ce type de systèmes, les différents sous-systèmes sont surveillés de façon distribuée ; seules les alarmes binaires sont centralisées dans une base de données. Il s'agit cette fois d'identifier des séquences de pannes.

1.7.5 Analyse hyperspectrale des interactions bactérie-minéral

- Financement : ANR HÆSPRI, CNRS, Région Lorraine
- Laboratoires partenaires : LIMOS, LEM, LCPME, CRAN
- Durée : 4 ans (2009 - 2013)
- Porteur : Christian Mustin (LIMOS)
- Autre chercheurs : David Brie, **Sebastian Miron**, Charles Soussen (CRAN), Patrick Billard (LIMOS), Manuel Dossot, Bernard Humbert (LCPME)
- Budget (part CRAN) : 170 k€

Résumé : Un défi scientifique fondamental est de comprendre les interactions entre les microbes et les minéraux. Les relations entre les activités microbiennes, les transformations continues des minéraux et le cycle des éléments sont essentielles pour de nombreuses facettes de l'activité humaine. L'objectif principal de ce projet est de révéler et d'analyser la variabilité cachée des interactions entre bactéries et minéraux, afin de délimiter les processus d'altération des minéraux par voie biologique et d'identifier leurs principes généraux. Dans ce contexte, et compte tenu de la compétence synergique des quatre partenaires, nous proposons une méthodologie originale et transversale pour étudier la variabilité écologique de l'altération bactérienne des silicates, en combinant des approches moléculaires, physico-chimiques, microbiologiques, avec des méthodes d'analyse de données et des signaux. Le projet réunit quatre équipes de Nancy : deux expertes en sciences de la terre, une en chimie et une en traitement du signal.

1.7.6 Méthodes tensorielles pour l'analyse du trafic des réseaux de transport multimodaux

- Financement : PEPS INS2I META-TRAM

- Laboratoires partenaires : Gipsa-Lab, I3S, CRAN
- Durée : 2 ans (2013 - 2014)
- Porteur : Alain Kibangou (Gipsa-Lab)
- Autre chercheurs : Pierre Comon (Gipsa-Lab), **Sebastian Miron**, Gérard Favier (I3S)
- Budget : 16 k€

Résumé : Afin de mieux gérer la mobilité dans nos agglomérations (éviter ou mieux contrôler les épisodes de congestion, prédire avec précision l'évolution du trafic, analyser avec finesse les plans de déplacements urbains et périurbains via des réseaux multimodaux), il est nécessaire de développer des outils d'analyse adéquats qui intègrent la multimodalité et l'hétérogénéité des réseaux à partir de mesures intrinsèquement multidimensionnelles. Ce projet a pour but d'explorer l'apport que peuvent avoir les décompositions tensorielles dans le contexte des réseaux de transport. Trois axes sont étudiés : la modélisation tensorielle en vue de l'estimation des matrices origine-destination, la clusterisation dynamique des flux et la synthèse d'algorithmes distribués adaptés au volume important des données, à la diversité des capteurs, à leur dispersion spatiale ainsi qu'à celle des opérateurs.

1.7.7 Imagerie de l'activité optique Raman pour l'analyse de molécules chirales

- Financement : PEPS INSIS ImRaChir
- Laboratoires partenaires : CRAN, LCPME
- Durée : 18 mois (2010 - 2011)
- Porteur : **Sebastian Miron**
- Autre chercheurs : David Brie (CRAN), Manuel Dossot, Cédric Carteret, Fabienne Quilles (LCPME)
- Budget : 20 k€

Résumé : L'utilisation d'une lumière polarisée circulairement droite ou gauche en spectroscopie Raman permet d'interagir de manière sélective avec les molécules chirales. Les données ainsi obtenues se composent de deux images hyperspectrales de microscopie Raman prises selon deux polarisations circulaires opposées. Ce projet a pour objectif l'analyse par microscopie Raman des molécules chirales et repose sur la mise en évidence de l'effet ROA (Raman Optical Activity) grâce à l'utilisation de lumière laser polarisée circulairement. D'un point de

vue instrumental, le défi majeur est de réussir à transférer cette approche classiquement utilisée en mode macroscopique (spectroscopie Raman) vers la microscopie confocale. Au niveau méthodologique, il s'agit d'étudier la possibilité de caractériser les molécules chirales par des techniques d'analyse de mélanges non-négatifs et multivariés. Le projet a également comme objet de formaliser la collaboration entre le CRAN et le LCPME sur l'analyse de mélanges en spectroscopie et imagerie Raman polarisées.

1.7.8 Biopsie optique des tumeurs cutanées : analyse Raman polarisée par micro- sonde d'imagerie fibrée et chimiométrie des signaux

- Financement : PEPS ST2I MIRACLE
- Laboratoires partenaires : MEDyC, CRAN, CReSTIC
- Durée : 2 ans (2008 - 2009)
- Porteur : Olivier Piot (MEDyC)
- Autre chercheurs : **Sebastian Miron**, Valeriu Vrabie (CReSTIC)
- Budget : 20 k€

Résumé : Les techniques de spectroscopie vibrationnelle telles que la diffusion Raman, basées sur l'interaction non destructive photons/matière, sont particulièrement bien adaptées à l'étude des échantillons biologiques pour le diagnostic et le pronostic de certaines pathologies. Les applications les plus prometteuses portent sur la détection des tissus cancéreux dans les stades précoces du développement tumoral et sur la caractérisation des cellules cancéreuses. Notre objectif est de développer une micro-sonde d'imagerie Raman dédiée au diagnostic *in vivo* et en temps réel des tumeurs cutanées de type carcinome BCC (*Basal Cell Carcinoma*) et SCC (*Squamous Cell Carcinoma*). Cette instrumentation innovante permettrait donc de réaliser des « biopsies optiques » directement sur des patients, sur la base d'informations moléculaires, contenues dans les spectres Raman et spécifiques de l'état physiopathologique du tissu ; et ceci sans aucun marquage ni préparation préalable du tissu. Notre approche vise donc à faciliter le diagnostic notamment dans les stades précoces du développement tumoral et à optimiser la stratégie thérapeutique médico-chirurgicale pour chaque type de lésion (médecine prédictive).

• Projets industriels :

1.7.9 Fusion des données (géolocalisation, scanner 3D *etc.*) pour la cartographie spatio-temporelle des polluants dans l'air des lieux de travail

- Financement : INRS
- Durée : 9 mois (mars - novembre 2021)
- Responsable : **Sebastian Miron**
- Budget : 8 k€

Résumé : Ce projet est un travail exploratoire visant à étudier la possibilité de fusionner les données issues de la cartographie spatio-temporelle de la pollution avec des données issues de la numérisation des environnements de travail (par le biais d'un scanner 3D) et de localisation des personnes/instruments dans ces environnements (système UWB), disponibles à l'INRS. A terme, l'objectif est de proposer un systèmes « clés en main » pour caractériser et comprendre les risques liées à la pollution dans les lieux de travail ainsi que pour estimer l'exposition à la pollution des travailleurs durant un cycle industriel.

1.7.10 Analyse des données issues d'instruments de mesure en temps réel pour la cartographie spatio-temporelle des polluants dans l'air des lieux de travail

- Financement : INRS
- Durée : 9 mois (mars - novembre 2020)
- Responsable : **Sebastian Miron**
- Budget : 10 k€

Résumé : L'analyse spatio-temporelle appliquée aux environnements professionnels consiste à étudier les phénomènes inscrits dans le temps et dans l'espace, notamment ceux concernant les pollutions atmosphériques et la tâche des travailleurs. En particuliers, elle décrit l'hétérogénéité dans l'espace et dans le temps des polluants chimiques et biologiques présents. Les informations collectées par les réseaux de capteurs concernent les niveaux de concentrations, les distributions en taille des particules portant ces polluants et leur variation dans le temps et l'espace. Cette cartographie spatio-temporelle appliquée aux environnements professionnels génère une masse de donnée importante. Ce projet correspond à une demande exprimée par l'INRS pour l'analyse et l'interprétation de ces données. Les objectifs sont : (1) définir

et proposer des modèles et des outils pour l'analyse des données et des signaux multidimensionnels, adaptés à l'étude des variations spatio-temporelles des concentrations en polluants chimiques dans l'air des lieux de travail ; (2) appliquer les méthodes et les approches définies au premier point à l'analyse des données réelles collectées dans des entreprises par le biais d'instruments de mesure ou de réseaux de capteurs spécifiques.

1.7.11 Echantillonnage compressé pour le traitement tensoriel d'antenne multi-variée

- Financement : entreprise High Tech Communication Services (HTCS)
- Durée : 3 ans (2013 - 2015)
- Co-responsables : **Sebastian Miron**, David Brie
- Budget : 95 k€

Résumé : Les avancées technologiques dans le domaine des capteurs et de la puissance de calcul ont rendu possible récemment la mise au point des réseaux d'antennes de très grande dimension (des champs d'antennes). A titre d'exemple, un télescope radio composé de 13 000 antennes vient d'être mis en fonction au New Mexico par l'université de New Mexico et la NASA. Ce type de réseau d'antenne permet d'acquérir une immense quantité de données, ce qui pose un réel problème quant à leur traitement. Le projet étudie l'implémentation des algorithmes de traitement d'antenne basés sur les techniques d'échantillonnage compressé dans un cadre tensoriel (multi-varié). Le premier objectif est de quantifier le gain dû à l'utilisation de l'échantillonnage compressé dans les algorithmes multilinéaires de traitement d'antenne. Le deuxième objectif est de proposer des configurations d'antenne permettant d'optimiser ce gain, en vue de la réalisation d'un démonstrateur.

1.8 Implication dans la vie collective

1.8.1 Collaborations internationales

- Kainam Thomas Wong (Hong Kong Polytechnic University, actuellement Beihang University, Chine) : Traitement d'antenne multidimensionnelle.
- Xijing Guo (Northwestern Polytechnical University, Chine) : Traitement d'antennes acoustiques.
- Alwin Stegeman (University of Groningen, Pays-Bas) : Unicité des décompositions tensorielles.
- André de Almeida, D. C. Araújo (Federal University of Ceará, Brésil) : Méthodes tensorielles en traitement d'antenne.

1.8.2 Collaborations nationales

- Xavier Luciani (SeaTech, Toulon) : Tenseurs à valeurs quaternioniques.
- Rémy Boyer (CRIStAL, Lille) : Décompositions en trains de tenseurs.
- Christian Mustin et Patrick Billard (LIEC, Nancy) dans la cadre de la thèse de Fabrice Caland ([D.2]).
- Manuel Dossot et Cédric Carteret (LCPME, Nancy) : Analyse de signaux de spectroscopie proche infrarouge et Raman.
- Saïd Moussaoui (LS2N, Nantes) : Séparation de sources en imagerie hyperspectrale.

1.8.3 Rayonnement et animation scientifique

- Membre du comité éditorial des revues indexées JCR *Physical Communication* (depuis 2012) et *IET Signal Processing* (depuis 2017).
- Membre extérieur du jury de thèse de Yassine Zniyed (thèse de l'Université Paris-Saclay, soutenue le 15 octobre 2019).
- Co-organisateur, avec Rémy Boyer (CRIStAL, Lille) et Hénrique Goulart (I3S, Nice), de la journée thématique GDR ISIS : « Nouvelles méthodes tensorielles et applications » (17 juin 2019).
- Conférencier invité à l'école d'été de Peyresq en traitement du signal et des images (25 juin - 1er juillet 2017) : Méthodes multilinéaires en traitement d'antenne.
- Conférencier invité à la journée du labcom P3A (6 décembre 2016, Lyon) : Antennes multi-échelle.
- Membre extérieur du jury de thèse de Abbas Rammal (thèse de l'Université de Reims Champagne-Ardenne, soutenue le 25 janvier 2016).
- Expert pour l'évaluation des projets nationaux ANR (2015).
- Membre extérieur du comité de sélection pour les postes : 61 MCF 0192 (Université de Bordeaux, 2012) et 61 MCF 1700 (Université de Paris-Sud 11, 2011).
- Titulaire de la prime d'excellence scientifique (2016-2020), (évaluation du CNU : A).

1.8.4 Responsabilités scientifiques

- Membre associé du comité technique (TC) de *IEEE Sensor Array and Multichannel (SAM)* (depuis 2018).
- Membre du comité de programme technique pour les conférences : *2017 International Conference on Latent Variable Analysis and Signal Separation (LVA-ICA 2017)*, *2020 IEEE Statistical Signal Processing Workshop (SSP 2020, 2021)*, *2020 IEEE 11th Sensor Array and Multichannel Signal Processing Workshop (SAM 2020)*.
- Editeur invité du numéro spécial *Polarization in wireless communications* de la revue *Physical Communication* (2012).
- Co-organisateur, avec Kainam Thomas Wong (Hong Kong Polytechnic University), de la session spéciale *Polarized signal processing* du *2011 IEEE Statistical Signal Processing Workshop*.
- Relecteur pour les revues *IEEE Transaction on Signal Processing, IEEE Signal Processing Letters, IEEE Transactions on Geosciences and Remote Sensing, IEEE Transactions on Aerospace and Electronic Systems, IET Signal Processing, EURASIP Signal Processing, EURASIP Journal on Advances in Signal Processing, Journal of Sound and Vibration, Sensors, Physical Communication, Chemometrics and Intelligent Laboratory Systems, Traitement du Signal*.

Chapitre 2

Production scientifique

2.1 Articles dans des revues internationales avec comité de lecture

[A20] J. Flamant, **S. Miron**, and D. Brie, "A general framework for constrained convex quaternion optimization," *IEEE Transactions on Signal Processing [submitted]*, 2021. [Online]. Available : https://arxiv.org/abs/2102.02763

[A19] **S. Miron**, M. Diop, A. Larue, E. Robin, and D. Brie, "Boolean decomposition of binary matrices using a post-nonlinear mixture approach," *Signal Processing*, vol. 178, p. 107809, 2021. [Online]. Available : https://doi.org/10.1016/j.sigpro.2020.107809

[A18] **S. Miron**, Y. Zniyed, R. Boyer, A. de Almeida, G. Favier, D. Brie, and P. Comon, "Tensor methods for multisensor signal processing," *IET Signal Processing*, vol. 178, no. 10, pp. 693 – 709, 2020. [Online]. Available : https://doi.org/10.1049/iet-spr.2020.0373

[A17] L. Nus, **S. Miron**, and D. Brie, "An ADMM-based algorithm with minimum dispersion regularization for on-line blind unmixing of hyperspectral images," *Chemometrics and Intelligent Laboratory Systems*, vol. 204, p. 104090, 2020.

[A16] J. Flamant, **S. Miron**, and D. Brie, "Quaternion non-negative matrix factorization : definition, uniqueness and algorithm," *IEEE Transactions on Signal Processing*, vol. 68, no. 1, pp. 1870–1883, 2020.

[A15] X. Guo, **S. Miron**, Y. Yang, and S. Yang, "Second-order cone programming with probabilistic regularization for robust adaptive beamforming," *The Journal of the Acoustical Society of America*, vol. 141, no. 3, pp. EL199–EL204, 2017.

[A14] X. Guo, **S. Miron**, Y. Yang, and S. Yang, "An upper bound for the directivity index of superdirective acoustic vector sensor arrays," *The Journal of the Acoustical Society of America*, vol. 140, no. 5, pp. EL410–EL415, 2016.

[A13] X. Guo, S. Yang, and **S. Miron**, "Low-frequency beamforming for a miniaturized aperture three-by-three uniform rectangular array of acoustic vector sensors," *The Journal of the Acoustical Society of America*, vol. 138, no. 6, pp. 3873–3883, 2015.

[A12] **S. Miron**, Y. Song, D. Brie, and K. T. Wong, "Multilinear direction finding for sensor-array with multiple scales of invariance," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 3, pp. 2057–2070, 2015.

[A11] D. Parrello, C. Mustin, D. Brie, **S. Miron**, and P. Billard, "Multicolor whole-cell bacterial sensing using a synchronous fluorescence spectroscopy-based approach," *PloS one*, vol. 10, no. 3, 2015.

[A10] D. Brie, R. Klotz, **S. Miron**, S. Moussaoui, C. Mustin, P. Bécuwe, and S. Grandemange, "Joint analysis of flow cytometry data and fluorescence spectra as a non-negative array factorization problem," *Chemometrics and Intelligent Laboratory Systems*, vol. 137, pp. 21–32, 2014.

[A9] R. Boyer and **S. Miron**, "Study of the asymptotic Cramér–Rao bound for the COLD uniform linear array," *Physical Communication*, vol. 5, no. 4, pp. 296–307, 2012.

[A8] X. Guo, **S. Miron**, and D. Brie, "The effect of polarization separation on the performance of CANDECOMP/PARAFAC-based vector sensor array processing," *Physical Communication*, vol. 5, no. 4, pp. 289–295, 2012.

[A7] X. Guo, **S. Miron**, D. Brie, and A. Stegeman, "Uni-mode and partial uniqueness conditions for CANDECOMP/PARAFAC of three-way arrays with linearly dependent loadings," *SIAM Journal on Matrix Analysis and Applications*, vol. 33, no. 1, pp. 111–129, 2012.

[A6] X. Guo, **S. Miron**, D. Brie, S. Zhu, and X. Liao, "A CANDECOMP/PARAFAC perspective on uniqueness of DOA estimation using a vector sensor array," *IEEE Transactions on Signal Processing*, vol. 59, no. 7, pp. 3475–3481, 2011.

[A5] **S. Miron**, M. Dossot, C. Carteret, S. Margueron, and D. Brie, "Joint processing of the parallel and crossed polarized Raman spectra and uniqueness in blind nonnegative source separation," *Chemometrics and Intelligent Laboratory Systems*, vol. 105, no. 1, pp. 7–18, 2011.

[A4] J. I. Mars, J. Guillon, **S. Miron**, V. Grulier, and C. Picard, "Séparation des composantes compressibles et incompressibles dans un champ de pression pariétale," *Traitement du Signal*, vol. 25, no. 4, pp. 346–356, 2008.

[A3] N. Le Bihan, **S. Miron**, and J. I. Mars, "MUSIC algorithm for vector-sensors array using biquaternions," *IEEE Transactions on Signal Processing*, vol. 55, no. 9, pp. 4523–4533, 2007.

[A2] **S. Miron**, N. Le Bihan, and J. I. Mars, "Quaternion-MUSIC for vector-sensor array processing," *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1218–1229, 2006.

[A1] **S. Miron**, N. Le Bihan, and J. I. Mars, "Vector-sensor MUSIC for polarized seismic sources localization," *EURASIP Journal on Applied Signal Processing*, vol. 2005, pp. 74–84, 2005.

2.2 Chapitres dans des ouvrages scientifiques

[O2] D. Brie, S. Moussaoui, **S. Miron**, C. Carteret, and M. Dossot, "Bayesian positive source separation for spectral mixture analysis," in *Resolving Spectral Mixtures with Applications from Ultrafast Time-Resolved Spectroscopy to Super-Resolution Imaging*, ser. Data Handling in Science and Technology. Elsevier, 2016, vol. 30, pp. 279–309, chapitre 8. [Online]. Available : https://hal.archives-ouvertes.fr/hal-01367514

[O1] **S. Miron**, X. Guo, and D. Brie, "Vector sensor array processing for polarized sources using a quadrilinear representation of the data covariance," in *Signal Processing*. In-Tech, 2010, pp. 19–36.

2.3 Conférences internationales avec actes et comité de lecture

[Ci32] Y. Znyied, K. Usevich, **S. Miron**, and D. Brie, "Learning nonlinearities in the decoupling problem with structured CPD," in *2021 19th IFAC Symposium on System Identification (SYSID)*. Padova, Italy : IFAC, July 2021.

[Ci31] L. Nus, **S. Miron**, B. Jaillais, S. Moussaoui, and D. Brie, "A semi-supervised rank tracking algorithm for on-line unmixing of hyperspectral images," in *2020 45th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Barcelona, Spain : IEEE, May 2020.

[Ci30] J. Flamant, **S. Miron**, and D. Brie, "Quaternion non-negative matrix factorization : a new tool for spectropolarimetric imaging," in *2019 8th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Guadeloupe, France, December 2019, pp. 336–340.

[Ci29] Y. Zniyed, **S. Miron**, R. Boyer, and D. Brie, "Uniqueness of tensor train decomposition with linear dependencies," in *2019 8th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Guadeloupe, France, December 2019, pp. 460–464.

[Ci28] M. Diop, **S. Miron**, A. Larue, and D. Brie, "Binary matrix factorization applied to Netflix dataset analysis," in *2019 5th IFAC Symposium on Telematics Applications (TA)*, Chengdu, China, September 2019.

[Ci27] M. Diop, **S. Miron**, A. Souloumiac, and D. Brie, "Boolean CP decomposition of binary tensors : uniqueness and algorithm," in *2019 44th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Brighton, UK : IEEE, May 2019.

[Ci26] L. Nus, **S. Miron**, and D. Brie, "Estimation of the regularization parameter of an on-line NMF with minimum volume constraint," in *2018 IEEE 10th Sensor Array and Multichannel Signal Processing Workshop (SAM)*. Sheffield, U.K. : IEEE, July 2018, pp. 425–429.

[Ci25] L. Nus, **S. Miron**, and D. Brie, "On-line blind unmixing for hyperspectral pushbroom imaging systems," in *2018 IEEE Statistical Signal Processing Workshop (SSP)*. Freiburg, Germany : IEEE, June 2018, pp. 418–422.

[Ci24] B. Jaillais, K. Meghar, L. Nus, **S. Miron**, D. Brie, and S. Moussaoui, "Unsupervised processing of hyperspectral images," in *2018 CHIMIOMETRIE XIX*, Paris, France, January 2018.

[Ci23] M. Diop, A. Larue, **S. Miron**, and D. Brie, "A post-nonlinear mixture model approach to binary matrix factorization," in *2017 25th European Signal Processing Conference (EUSIPCO)*. Kos Island, Greece : IEEE, August 2017, pp. 321–325.

[Ci22] M.-A. Bisch, **S. Miron**, D. Brie, and D. Lejeune, "A sparse approach for DOA estimation with a multiple spatial invariance sensor array," in *2016 IEEE Statistical Signal Processing Workshop (SSP)*. Palma de Mallorca, Spain : IEEE, June 2016.

[Ci21] **S. Miron** and D. Brie, "Some rank conditions for the identifiability of the sparse Paralind model," in *2015 International Conference on Latent Variable Analysis and Signal Separation (LVA-ICA)*. Liberec, Czech Republic : Springer, August 2015, pp. 41–48.

[Ci20] D. Brie, **S. Miron**, P. Bécuwe, and S. Grandemange, "A non-negative multilinear block tensor decomposition approach to flow cytometry data analysis," in *2014 IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*. Reims, France : IEEE, September 2014.

[Ci19] **S. Miron**, Y. Song, D. Brie, and K. T. Wong, "CANDECOMP/PARAFAC (CP) direction finding with multi-scale array," in *2013 5th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Saint Martin, France, December 2013, pp. 224–227.

[Ci18] M. Boizard, G. Ginolhac, F. Pascal, **S. Miron**, and P. Forster, "Numerical performance of a tensor MUSIC algorithm based on HOSVD for a mixture of polarized sources," in *2013 21st European Signal Processing Conference (EUSIPCO)*, Marrakech, Marocco, September 2013.

[Ci17] Y. Merah, **S. Miron**, and D. Brie, "A generalized acquisition scheme for vector cross-product direction finding with spatially spread vector-sensor components," in *2013 38th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Vancouver, Canada : IEEE, May 2013, pp. 3977–3980.

[Ci16] A. L. De Almeida, A. Y. Kibangou, **S. Miron**, and D. C. Araújo, "Joint data and connection topology recovery in collaborative wireless sensor networks," in *2013 38th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Vancouver, Canada : IEEE, May 2013, pp. 5303–5307.

[Ci15] F. Caland, **S. Miron**, D. Brie, and C. Mustin, "A blind sparse approach for estimating constraint matrices in Paralind data models," in *2012 20th European Signal Processing Conference (EUSIPCO)*. Bucharest, Romania : IEEE, August 2012, pp. 839–843.

[Ci14] F. Caland, **S. Miron**, D. Brie, and C. Mustin, "A CANDECOMP/PARAFAC approach to the estimation of environmental pollutant concentrations using biosensors," in *2011 IEEE Statistical Signal Processing Workshop (SSP)*. Nice, France : IEEE, June 2011, pp. 801–804.

[Ci13] **S. Miron**, D. Brie, M. Dossot, and C. Carteret, "Micro-crystal orientations estimation in polarized Raman microscopy using an acquisition scheme with multiple diversities," in *2011 IEEE Statistical Signal Processing Workshop (SSP)*. Nice, France : IEEE, June 2011, pp. 21–24.

[Ci12] D. Brie, **S. Miron**, F. Caland, and C. Mustin, "An uniqueness condition for the 4-way CANDECOMP/PARAFAC model with collinear loadings in three modes," in *2011 36th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Prague, Czech Republic : IEEE, May 2011, pp. 4108–4111.

[Ci11] X. Guo, S. Zhu, **S. Miron**, and D. Brie, "Approximate joint diagonalization by non-northogonal nonparametric Jacobi transformations," in *2010 35th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Dallas, U.S.A. : IEEE, March 2010, pp. 3774–3777.

[Ci10] **S. Miron**, D. Brie, M. Dossot, and B. Humbert, "On the modeling and processing of polarization in Raman spectroscopy," in *2009 First Workshop on Hyperspectral Image and Signal Processing : Evolution in Remote Sensing (WHISPERS)*. Grenoble, France : IEEE, August 2009.

[Ci9] C. Soussen, **S. Miron**, F. Caland, D. Brie, P. Billard, and C. Mustin, "Complexity reduction by convex cone detection for unmixing hyperspectral images of bacterial biosensors," in *2009 17th European Signal Processing Conference (EUSIPCO)*, Glasgow, Scotland, August 2009, pp. 1938–1942.

[Ci8] **S. Miron**, X. Guo, and D. Brie, "DOA estimation for polarized sources on a vector-sensor array by PARAFAC decomposition of the fourth-order covariance tensor," in *2008 16th European Signal Processing Conference (EUSIPCO)*, Lausanne, Switzerland, August 2008.

[Ci7] X. Guo, **S. Miron**, and D. Brie, "Identifiability of the PARAFAC model for polarized source mixture on a vector sensor array," in *2008 33rd IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Las Vegas, U.S.A. : IEEE, March 2008, pp. 2401–2404.

[Ci6] X. Guo, **S. Miron**, and D. Brie, "Three-way array analysis on polarized signals for direction-finding and blind source separation," in *2007 IAR*, Grenoble, France, November 2007.

[Ci5] **S. Miron**, N. Le Bihan, and J. I. Mars, "High resolution vector-sensor array processing based on biquaternions," in *2006 31st IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Toulouse, France : IEEE, May 2006.

[Ci4] **S. Miron**, N. Le Bihan, and J. I. Mars, "High resolution vector-sensor array processing using quaternions," in *2005 IEEE Statistical Signal Processing Workshop (SSP)*. Bordeaux, France : IEEE, July 2005, pp. 918–923.

[Ci3] **S. Miron**, N. Le Bihan, and J. I. Mars, "Polarized source characterization using vector-MUSIC," in *2004 12th European Signal Processing Conference (EUSIPCO)*. Vienna, Austria : IEEE, September 2004, pp. 421–424.

[Ci2] **S. Miron**, N. Le Bihan, and J. I. Mars, "Joint estimation of direction of arrival and polarization parameters for multicomponent sensor array," in *66th European Association of Geoscientists and Engineers Conference and Exhibition (EAGE)*, Paris, France, June 2004.

[Ci1] **S. Miron**, M. Guillon, N. Le Bihan, and J. I. Mars, "Multidimensional signal processing using quaternions," in *3rd Workshop on Physics in Signal and Image Processing (PSIP)*, Grenoble, France, January 2003, pp. 57–60.

2.4 Conférences nationales avec actes et comité de lecture

[Cn11] L. Nus, **S. Miron**, and D. Brie, "Démélange d'images hyperspectrales à l'aide de la NMF en-ligne avec contrainte de dispersion minimale," in *2019 XXVII^{ème} Colloque GRETSI Traitement du Signal & des Images (GRETSI)*, Lille, France, August 2019.

[Cn10] J. Flamant, **S. Miron**, and D. Brie, "Factorisation en matrices quaternioniques non-négatives : un nouvel outil pour l'imagerie spectro-polarimétrique," in *2019 XXVII^{ème} Colloque GRETSI Traitement du Signal & des Images (GRETSI)*, Lille, France, August 2019.

[Cn9] Y. Zniyed, **S. Miron**, R. Boyer, and D. Brie, "Uniqueness of Tensor Train Decomposition with Linear Dependencies," in *2019 XXVII^{ème} Colloque GRETSI Traitement du Signal & des Images (GRETSI)*, Lille, France, August 2019.

[Cn8] M. Diop, A. Larue, **S. Miron**, and D. Brie, "Factorisation en matrices binaires par modèle de mélange post non-linéaire," in *2017 XXVI^{ème} Colloque GRETSI Traitement du Signal & des Images (GRETSI)*, Juan-les-Pins, France, September 2017.

[Cn7] M.-A. Bisch, **S. Miron**, D. Brie, and D. Lejeune, "Estimation parcimonieuse des directions d'arrivées pour les antennes multi-échelle," in *2015 XXV^{ème} Colloque GRETSI Traitement du Signal & des Images, (GRETSI)*, Lyon, France, September 2015.

[Cn6] Y. Merah, **S. Miron**, and D. Brie, "Estimation de la direction d'arrivée des sources polarisées avec un capteur électromagnétique à six composantes non-localisées," in *2013 XXIV^{ème} Colloque GRETSI Traitement du Signal & des Images (GRETSI)*, Brest, France, September 2013.

[Cn5] F. Caland, **S. Miron**, D. Brie, and C. Mustin, "Estimation de la réponse de biosenseurs par l'analyse des signaux spectraux multivariés," in *2011 XXIII^{ème} Colloque GRETSI Traitement du Signal & des Images, (GRETSI)*, Bordeaux, France, September 2011.

[Cn4] X. Guo, **S. Miron**, D. Brie, and A. Stegeman, "Identifiabilité partielle de mélanges trilinéaires de sources linéairement dépendantes," in *2011 XXIII^{ème} Colloque GRETSI Traitement du Signal & des Images, (GRETSI)*, Bordeaux, France, September 2011.

[Cn3] X. Guo, **S. Miron**, and D. Brie, "Identifiabilité du mélange de sources polarisées : le cas où le nombre de capteurs est inférieur au nombre de sources," in *2009 XXII^{ème} Colloque GRETSI Traitement du Signal & des Images (GRETSI)*, Dijon, France, September 2009.

[Cn2] J. Guillon, **S. Miron**, V. Grulier, C. Picard, and J. I. Mars, "Séparation des composantes compressible et incompressible d'un champ de pression pariétal," in *2007 XXI^{ème}*

Colloque GRETSI Traitement du Signal & des Images (GRETSI), Troyes, France, September 2007.

[Cn1] **S. Miron**, N. Le Bihan, and J. I. Mars, "Etude des performances du modele quaternionique en traitement d'antenne vectorielle," in *2005 XX^{ème} Colloque GRETSI Traitement du Signal & des Images (GRETSI)*, Louvain-la-Neuve, Belgium, September 2005.

2.5 Mémoire de thèse

[Th] **S. Miron**, "Méthodes multilinéaires et hypercomplexes en traitement d'antenne multicomposante à haute résolution," Ph.D. dissertation, Institut National Polytechnique de Grenoble (INPG), Grenoble, France, October 2005.

Part II

A summary of my research activites

Foreword

In this part of the manuscript I present a summary of my research activities, with results extracted from the Ph.D. thesis that I co-supervised and from my other academic collaborations. I chose to focus hereafter on topics related to tensors and constraint multivariate source unmixing, as they represent the core of my research activities since my Ph.D. defense. For the thematic coherence of the presentation, I will not detail my work on spatially spread antennas and acoustic sensor arrays in this part of the manuscript.

The summary of my contributions on tensors and constraint multivariate source unmixing is organized in three chapter as follows:

- Chapter 3: Uniqueness of tensor decompositions with linear dependencies in the modes. This chapter regroups contributions from the Ph.D. of Xijing Guo (Section 1.6: D.1) and several other academic collaborations;

- Chapter 4: Tensor-based DOA estimation with sensor arrays. The results presented here are mainly extracted from the Ph.D. of Xijing Guo (Section 1.6: D.1), and the collaboration with the Hong Kong Polytechnic University.

- Chapter 5: Tensor models and methods in spectroscopy. This chapter presents research from the Ph.D. of Fabrice Caland (Section 1.6: D.2), and collaborations with colleagues from CRAN, LCPME and LIEC.

- Chapter 6: Constrained unmixing of multidimensional sources. These methods were developed in the Ph.D.'s of Mamdou Diop, Ludivine Nus (Section 1.6: D.3, D.4) and the post-doctoral work of Julien Flamant (Section 1.6 : pD.2).

Within each section of these chapters, I privileged the chronological order of the results, unless the logical flow of the presentation required otherwise.

Notations

x, N, λ	:	scalars
$\boldsymbol{x}, \boldsymbol{\lambda}$	:	vectors
$\boldsymbol{X}, \boldsymbol{\Lambda}$	:	matrices
$\boldsymbol{\mathcal{X}}$	:	tensor
$\|\boldsymbol{x}\|_p$	:	ℓ_p norm of a vector ($p > 0$); $\|\boldsymbol{x}\|_p = (\sum_i x_i^p)^{1/p}$
$\|\boldsymbol{x}\|_0$	:	ℓ_0 pseudo-norm of a vector ; $\|\boldsymbol{x}\|_0 = \sum_{i,x_i \neq 0} \|x_i\|^0$
$\|\boldsymbol{X}\|_{\mathsf{F}}$	:	Frobenius norm of a matrix ; $\|\boldsymbol{X}\|_{\mathsf{F}} = \sqrt{\sum_{i,j} \|x_{ij}\|^2}$
$\|\boldsymbol{X}\|_{p,q}$	:	mixed ℓ_p/ℓ_q norm of a matrix; $\|\boldsymbol{X}\|_{p,q} = \left(\sum_i \|\boldsymbol{x}(i,:)\|_q^p\right)^{1/p}$, where $\boldsymbol{x}(i,:)$ is the i^{th} row of $\boldsymbol{X}$
$\boldsymbol{X}^{\mathsf{T}}$	:	transpose of $\boldsymbol{X}$
$\boldsymbol{X}^*$	:	conjugate of $\boldsymbol{X}$
$\boldsymbol{X}^{\mathsf{H}}$	:	conjugate transpose of $\boldsymbol{X}$
$\boldsymbol{X}^\dagger$	:	Moore-Penrose pseudo-inverse of $\boldsymbol{X}$
$\boldsymbol{X} \otimes \boldsymbol{Y}$	:	Kronecker product of two matrices
$\boldsymbol{X} \odot \boldsymbol{Y}$	:	Khatri-Rao product of two matrices having the same number of columns
$\boldsymbol{X} \boxdot \boldsymbol{Y}$	:	Hadamard (element-wise) product of two matrices having the same size
$\mathcal{U} \circ \mathcal{V}$	:	tensor product of vector spaces $\mathcal{U}, \mathcal{V}$
$\mathcal{U} \oplus \mathcal{V}$	:	direct sum of vector spaces $\mathcal{U}, \mathcal{V}$
$\boldsymbol{x} \circ \boldsymbol{y}$	:	outer product of two vectors: $[\boldsymbol{x} \circ \boldsymbol{y}]_{ij} = x_i y_j$
$\boldsymbol{\mathcal{X}} \underset{q}{\bullet} \boldsymbol{A}$	:	n-mode product between a tensor $\boldsymbol{\mathcal{X}}$ ($N_1 \times \cdots \times N_Q$) and a matrix $\boldsymbol{A}$ ($J \times N_q$): $[\boldsymbol{\mathcal{X}} \underset{q}{\bullet} \boldsymbol{A}]_{n_1,\ldots,n_{q-1},j,n_{q+1},\ldots,n_Q} = \sum_{n_q=1}^{N_q} [\boldsymbol{\mathcal{X}}]_{n_1,\ldots,n_{q-1},n_q,n_{q+1},\ldots,n_Q} [\boldsymbol{A}]_{j,n_q}$
$\boldsymbol{\mathcal{X}} \overset{p}{\underset{q}{\bullet}} \boldsymbol{\mathcal{Y}}$	:	contraction between two tensors: $[\boldsymbol{\mathcal{X}} \overset{p}{\underset{q}{\bullet}} \boldsymbol{\mathcal{Y}}]_{n_1,\ldots,n_{q-1},n_{q+1},\ldots,n_Q,m_1,\ldots,m_{p-1},m_{p+1},\ldots,m_P} = \sum_{k=1}^{N_q} [\boldsymbol{\mathcal{X}}]_{n_1,\ldots,n_{q-1},k,n_{q+1},\ldots,n_Q} [\boldsymbol{\mathcal{Y}}]_{m_1,\ldots,m_{p-1},k,m_{p+1},\ldots,m_P}$

Chapter 3

Uniqueness of tensor decompositions with linear dependencies in the modes

3.1 Preliminaries: tensors, CPD and uniqueness

Tensors are multilinear maps between two sets of linear spaces [Comon14]. Once these linear spaces are fixed as well as their bases, a tensor is characterized by its array of coordinates. In this manuscript, we shall assume that this is the case, and shall mainly work on array coordinates. Hence, with a usual abuse of vocabulary, we shall assimilate tensors with their array representations. The number of dimensions (indices) of such an array is called the order of the tensor. Thus, a tensor of order Q is a Q-way array:

$$\boldsymbol{\mathcal{X}} \in \mathbb{K}^{N_1 \times \cdots \times N_Q}, \tag{3.1}$$

with $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, and $N_1, \ldots, N_Q$, the sizes of the Q dimensions (modes).

In 1927 Hitchcock [Hitchcock27a, Hitchcock27b] introduced a Canonical Polyadic Decomposition (CPD) of a multiway array or tensor. The same decomposition, for three-way arrays, was proposed later independently by Carroll and Chang [Carroll70] in psychometrics, who named it CANDECOMP (CANonical DECOMPosition), and by Harshman [Harshman70] in phonetics, who called it PARAFAC (PARAllel FACtor decomposition). The CPD of a Q-order tensor can be expressed as:

$$\boldsymbol{\mathcal{X}} = \sum_{r=1}^{R} \boldsymbol{\mathcal{X}}(r), \tag{3.2}$$

where $\boldsymbol{\mathcal{X}}(r) = \boldsymbol{p}_1(r) \circ \ldots \circ \boldsymbol{p}_Q(r)$ is a rank-1 tensor ("$\circ$" denotes the outer product between vectors). The minimum value of R for which (3.3) holds exactly is called the rank of $\boldsymbol{\mathcal{X}}$. By regrouping the loading vectors $\boldsymbol{p}_q(r)$ in matrices $\boldsymbol{P}_q = [\boldsymbol{p}_q(1) \ldots \boldsymbol{p}_q(R)]$, $(q = 1, \ldots, Q)$, the CPD of $\boldsymbol{\mathcal{X}}$ can be alternatively expressed as:

$$\boldsymbol{\mathcal{X}} = [\![\boldsymbol{P}_1, \ldots, \boldsymbol{P}_Q]\!]. \tag{3.3}$$

Another notation for the CPD model used in this manuscript is:

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{I}}_{Q,R} \underset{1}{\bullet} \boldsymbol{P}_1 \underset{2}{\bullet} \boldsymbol{P}_2 \ldots \underset{Q}{\bullet} \boldsymbol{P}_Q, \tag{3.4}$$

where "$\underset{n}{\bullet}$"symbolizes the n-mode product between a tensor and a matrix, and $\boldsymbol{\mathcal{I}}_{Q,R}$ denotes the Q-order identity tensor of size $R \times \cdots \times R$. The CPD of a tensor can also be represented in a flattened version, as a q-mode unfolding matrix $\text{unfold}_q \boldsymbol{\mathcal{X}}$ of size $N_q \times \frac{N_1 \ldots N_Q}{N_q}$ given by:

$$\text{unfold}_q \boldsymbol{\mathcal{X}} = \boldsymbol{P}_q \left(\boldsymbol{P}_Q \odot \cdots \odot \boldsymbol{P}_{q+1} \odot \boldsymbol{P}_{q-1} \odot \cdots \odot \boldsymbol{P}_1\right)^{\mathsf{T}}. \tag{3.5}$$

In this chapter we address the problem of the uniqueness of the CPD $\boldsymbol{\mathcal{X}} = [\![\boldsymbol{P}_1, \ldots, \boldsymbol{P}_Q]\!]$ (or equivalently, the identifiability of the associated CPD model), in the case where linear dependencies are present between the columns of the loading matrices $\boldsymbol{P}_q$. The uniqueness of the CPD model is defined up to some trivial scale and permutation indeterminacies, leading to the notion of *essential uniqueness*. We say that the CPD (3.3) is *essentially unique* if for any another set of matrices $\bar{\boldsymbol{P}}_1, \ldots, \bar{\boldsymbol{P}}_Q$ satisfying $\boldsymbol{\mathcal{X}} = [\![\bar{\boldsymbol{P}}_1, \ldots, \bar{\boldsymbol{P}}_Q]\!]$, there exists a permutation matrix $\boldsymbol{\Pi}$ and Q invertible diagonal scaling matrices $\boldsymbol{\Delta}_1, \ldots, \boldsymbol{\Delta}_Q$, satisfying $\boldsymbol{\Delta}_1 \times \cdots \times \boldsymbol{\Delta}_Q = \boldsymbol{I}_R$, (where $\boldsymbol{I}_R$ is the $R \times R$ identity matrix), such that

$$\bar{\boldsymbol{P}}_1 = \boldsymbol{P}_1 \boldsymbol{\Pi} \boldsymbol{\Delta}_1, \quad \ldots \quad , \bar{\boldsymbol{P}}_Q = \boldsymbol{P}_Q \boldsymbol{\Pi} \boldsymbol{\Delta}_Q.$$

In the rest of the manuscript, the term *uniqueness* will be used to designate the *essential uniqueness*.

One of the most popular conditions for the uniqueness of the CPD is the Kruskal's condition [Kruskal77] relying on the concept of "Kruskal-rank", or simply "krank". The krank of an $N \times R$ matrix $\boldsymbol{P}$, denoted by $\text{krank}\{\boldsymbol{P}\}$, is the maximum value $\ell \in \mathbb{N}$ such that every ℓ columns of $\boldsymbol{P}$ are linearly independent. By definition, the krank of a matrix is less than or equal to its rank. Kruskal proved [Kruskal77] that the condition

$$\text{krank}\{\boldsymbol{P}_1\} + \text{krank}\{\boldsymbol{P}_2\} + \text{krank}\{\boldsymbol{P}_3\} \geq 2R + 2 \tag{3.6}$$

is sufficient for uniqueness of the CPD decomposition of third order tensors ($Q = 3$). Furthermore, it becomes a necessary and sufficient condition in the cases $R = 2$ or 3 (see [tenBerge02]).

The uniqueness condition (3.6) has been generalized to Q-order CPDs in [Sidiropoulos00a]. It states that the loading matrices $\boldsymbol{P}_q$ ($q = 1, \ldots, Q$) in (3.3) can be uniquely estimated from

$\boldsymbol{\mathcal{X}}$ if

$$\sum_{q=1}^{Q} \text{krank}\{\boldsymbol{P}_q\} \geq 2R + (Q-1). \tag{3.7}$$

This condition is sufficient but not necessary for the uniqueness of the CPD. Linear dependence among the loading vectors may violate these uniqueness conditions. In this case, CPD may (but does not necessarily) encounter difficulties. There are cases where some loading vectors are uniquely determined, while other subsets of loading vectors are subject to rotational indeterminacies. This phenomenon, known as *partial uniqueness*, was first reported in [Harshman70].

A particular case of CPD, that explicitly accounts for the linear dependencies between its loading matrices, was introduced in [Bro09] and [deAlmeida08a]. They propose to use prespecified matrices, known as constraint matrices, to describe the linear dependence patterns in the loading matrices. These models were called *PARAllel profiles with LINear Dependencies (PARALIND)* [Bro09] or *CONstraint FACtors (CONFAC)* [deAlmeida08a]. Instead of the $\boldsymbol{P}_q$ $(q = 1, \ldots, Q)$ matrices in the conventional CPD (3.3), the new loading matrices are given by $\widetilde{\boldsymbol{P}}_q \boldsymbol{\Psi}_q$, where $\widetilde{\boldsymbol{P}}_q$ are full-column rank matrices and $\boldsymbol{\Psi}_q$, are constraint matrices containing the patterns of linear dependencies.

In this chapter I present several uniqueness and partial uniqueness results that we developed for different tensor configurations.

3.2 Uniqueness of third-order CPD with a full-rank loading matrix

The previous section introduced the notations for the general case of a Q-order tensor. In the sequel, when dealing with third-order tensors, the following simplified CPD notation will be used:

$$\boldsymbol{\mathcal{X}} = [\![\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}]\!], \tag{3.8}$$

with $\boldsymbol{A} = [\boldsymbol{a}_1 \ \cdots \ \boldsymbol{a}_R]$ $(I \times R)$, $\boldsymbol{B} = [\boldsymbol{b}_1 \ \cdots \ \boldsymbol{b}_R]$ $(J \times R)$ and $\boldsymbol{C} = [\boldsymbol{c}_1 \ \cdots \ \boldsymbol{c}_R]$ $(K \times R)$. In the sequel, only the "noiseless" case is considered, but the presented results also hold for a "reasonable" amount of noise.

A situation often encountered in practical applications is the CPD with one full column rank matrix (say $\boldsymbol{C}$). In this case, Kruskal's condition reads:

$$\text{krank}\{\boldsymbol{A}\} + \text{krank}\{\boldsymbol{B}\} \geq R + 2. \tag{3.9}$$

However, when one loading matrix has full column rank, the condition (3.9) can be further relaxed. A *necessary and sufficient* condition for identifiability of the CPD model with one full-rank matrix was derived in [Jiang04]. This condition says that the CPD $\boldsymbol{\mathcal{X}} = [\![\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}]\!]$

is unique if *none* of the nontrivial[1] linear combinations of the columns of $\boldsymbol{A} \odot \boldsymbol{B}$ can be written as a tensor product of two vectors.

An easier-to-check, *sufficient* condition for uniqueness of the CP model with a full column rank loading matrix $\boldsymbol{C}$ was also provided in [Jiang04] (see also [Stegeman09a]). It depends on a prespecified $I^2J^2 \times R(R-1)/2$ matrix $\boldsymbol{U}$ whose elements are determined by the product of the second-order minors of $\boldsymbol{A}$ and $\boldsymbol{B}$. In [Jiang04] it was proven that if $\boldsymbol{U}$ has full column rank then identifiability of the model is guaranteed. The same condition was derived independently in [De Lathauwer06a].

The necessary and sufficient condition of [Jiang04] is the cornerstone of the main uniqueness result of this section (*i.e.*, Theorem 3.1 in the sequel) and leads us naturally to recall the relationship between the rank of $\boldsymbol{A} \odot \boldsymbol{B}$ and identifiability of the CPD model [Liu01], *i.e.*, CPD is identifiable only if $\boldsymbol{A} \odot \boldsymbol{B}$ has full column rank. We introduced [Guo11] a condition ensuring full column rank of $\boldsymbol{A} \odot \boldsymbol{B}$.

Lemma 3.1 ([Guo11])**.** *Let $\boldsymbol{A}(I \times R)$ and $\boldsymbol{B}(J \times R)$ be two matrices consisting of nonzero columns. If*

$$\max\{\text{rank}\{\boldsymbol{A}\} + \text{krank}\{\boldsymbol{B}\}, \text{rank}\{\boldsymbol{B}\} + \text{krank}\{\boldsymbol{A}\}\} \geq R + 1 \tag{3.10}$$

holds, then $\boldsymbol{A} \odot \boldsymbol{B}$ is full column rank.

In [Sidiropoulos00b], a similar lemma was presented but demanding $\text{krank}\{\boldsymbol{A}\}+\text{krank}\{\boldsymbol{B}\} \geq R+1$. Comparatively, it is a milder condition posed here in Lemma 3.1. This lemma shows that a full column rank Khatri-Rao matrix can be generated with two rank-deficient matrices, which is also the basis of signal decorrelation techniques such as spatial averaging [Shan85] or the polarization smoothing algorithm (PSA) [Rahamim04].

Generally, the necessary and sufficient uniqueness condition of [Jiang04] is not easy to verify in practice. Towards this end, we proved a sufficient condition for identifiability of the CPD model with a full column rank loading matrix expressed in terms of rank and krank of the loading matrices.

Theorem 3.1 ([Guo09, Guo11])**.** *The CPD model (3.8) is identifiable if the loading matrix $\boldsymbol{C}$ is full column rank and the other two loading matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ satisfy the conditions*

$$i)\ \min\{\text{krank}\{\boldsymbol{A}\}, \text{krank}\{\boldsymbol{B}\}\} \geq 2, \tag{3.11}$$

and

$$ii)\ \max\{\text{rank}\{\boldsymbol{A}\} + \text{krank}\{\boldsymbol{B}\}, \text{rank}\{\boldsymbol{B}\} + \text{krank}\{\boldsymbol{A}\}\} \geq R + 2. \tag{3.12}$$

The same condition was also proven independently by Stegeman [Stegeman09a], who found it *sufficient* for full column rank of the prespecified matrix $\boldsymbol{U}$ and hence arrived at the assertion of uniqueness for the details). However, we provided an alternative proof, using

[1] The *nontrivial* linear combination is referred to as a combination involving *at least two* vectors.

Lemma 3.1. It is also worth mentioning that the condition (3.12) is equivalent to Kruskal's condition if *e.g.,* $\text{rank}\{\boldsymbol{A}\} = \text{krank}\{\boldsymbol{A}\}$.

3.3 Partial uniqueness conditions for third-order tensors

As mentioned in Section 3.1, when linear dependencies are present between the columns of the loading matrices, *partial uniqueness* phenomena may be observed. Partial uniqueness of third-order CPD model has received quite a lot of attention over the last two decades, and some significant results can be found in [tenBerge04, Bro09, deAlmeida08a, Stegeman09b]. In this section, we consider the special case where Kruskal's condition (3.6) is not met and the linear dependencies may take the form of collinear columns in *only one* loading matrix, say $\boldsymbol{A}$. It should be pointed out that having collinear loading vectors in $\boldsymbol{A}$ implies nonuniqueness (or at most partial uniqueness) of the other two modes $\boldsymbol{B}$ and $\boldsymbol{C}$ if linear dependency of the corresponding loading vectors of $\boldsymbol{B}$ and $\boldsymbol{C}$ does not exist [Stegeman09b, Lemma 4.6].

Regarding the uniqueness issues, the following two questions arise naturally. The first is under what conditions essential uniqueness of the first mode loading matrix $\boldsymbol{A}$ is ensured. The second is whether essential/partial uniqueness holds for the loadings in the other two modes. Section 3.3.1 presents three sufficient conditions to answer the first question, whereas an answer to the second question is provided in Section 3.3.2. Throughout the entire section, we assume that $\boldsymbol{B}$ and $\boldsymbol{C}$ each have no collinear columns. Meanwhile, no assumptions are made about the dependencies in the columns of $\boldsymbol{A}$.

3.3.1 Uni-mode uniqueness of the third-order CPD with linearly dependent loadings

Uni-mode uniqueness is a special case of *partial uniqueness* where only one loading matrix is essentially unique. In [Guo12] we proved a first condition for uniqueness of the first mode loadings, which is given by the following theorem:

Theorem 3.2 ([Guo12])**.** *Recall the CPD model of a three-way array $\boldsymbol{\mathcal{X}}$ given by (3.8). If $\boldsymbol{A}$ has no zero columns and the condition*

$$\text{rank}\{\boldsymbol{A}\} + \text{krank}\{\boldsymbol{B}\} + \text{krank}\{\boldsymbol{C}\} \geq 2R + 2 \tag{3.13}$$

holds, the first mode loading matrix $\boldsymbol{A}$ is unique up to permutation and scaling of the columns.

Observe that, if $\text{krank}\{\boldsymbol{A}\} = \text{rank}\{\boldsymbol{A}\}$, the condition (3.13) becomes identical to Kruskal's condition implying uniqueness of all the loading matrices $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$. In the case where $\text{krank}\{\boldsymbol{A}\} < \text{rank}\{\boldsymbol{A}\}$, however, the second and the third mode loading matrices $\boldsymbol{B}$ and $\boldsymbol{C}$

may not necessarily be unique. In particular, if

$$(\mathfrak{A}) \quad \operatorname{krank}\{\boldsymbol{B}\} < \operatorname{rank}(\boldsymbol{B}) \text{ and } \operatorname{krank}\{\boldsymbol{C}\} < \operatorname{rank}\{\boldsymbol{C}\}$$

holds as well, the condition (3.13) can be further weakened, as stated in the following by our second uniqueness condition.

Theorem 3.3 ([Guo12]). *Let us recall the CP decomposition problem (3.8). If $\boldsymbol{A}$ has no zero columns,* $(\mathfrak{A})$ *holds and*

$$\operatorname{rank}\{\boldsymbol{A}\} + \operatorname{krank}\{\boldsymbol{B}\} + \operatorname{krank}\{\boldsymbol{C}\} \geq 2R + 1, \tag{3.14}$$

then the first mode loading matrix $\boldsymbol{A}$ is unique up to permutation and scaling of the columns.

Both Th.3.2 and Th.3.3 are generalized by the third condition, as shown below.

Theorem 3.4 ([Guo12]). *In the CPD (3.8), if $\boldsymbol{A}$ has no zero columns and the following two conditions:*

$$\begin{aligned} &\operatorname{rank}\{\boldsymbol{A}\} + \min\{\operatorname{krank}\{\boldsymbol{B}\}, \operatorname{krank}\{\boldsymbol{C}\}\} \geq R + 2, \\ &\operatorname{rank}\{\boldsymbol{A}\} + \operatorname{krank}\{\boldsymbol{B}\} + \operatorname{krank}\{\boldsymbol{C}\} + \max\{\operatorname{rank}(\boldsymbol{B}) - \operatorname{krank}\{\boldsymbol{B}\}, \operatorname{rank}\{\boldsymbol{C}\} - \operatorname{krank}\{\boldsymbol{C}\}\} \\ &\geq 2R + 2, \end{aligned} \tag{3.15}$$

hold, then $\boldsymbol{A}$ is unique up to permutation and scaling of the columns.

Discussion If $\operatorname{rank}\{\boldsymbol{A}\} = R$ then the two conditions in Theorem 3.4 implie the essential uniqueness of all three matrices $\boldsymbol{A}, \boldsymbol{B}$ and $\boldsymbol{C}$, as shown in [Stegeman09a] and [Guo11].

Regarding the relationships between the three sufficient conditions, it is worth noting that

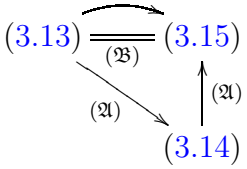
1. though the second condition (3.14) appears to be slightly weaker than the first one (3.13), it is subject to $(\mathfrak{A})$; hence, it can not completely substitute for (3.13);
2. the third condition (3.15) is necessary both for (3.13) and for (3.14) under $(\mathfrak{A})$.

It should be pointed out that condition (3.13) follows immediately from (3.15) if

$$(\mathfrak{B}) \quad \operatorname{krank}\{\boldsymbol{B}\} = \operatorname{rank}\{\boldsymbol{B}\} \text{ and } \operatorname{krank}\{\boldsymbol{C}\} = \operatorname{rank}\{\boldsymbol{C}\}$$

holds.

The following diagram illustrates the relationships between the aforementioned three conditions. In the diagram, we denote by "$\rightarrow$" as "being sufficient for", and by "$=$" as "being equivalent to". The notations $(\mathfrak{A})$ and $(\mathfrak{B})$ denote the respective conditions under which the sufficiency/equivalence holds.



3.3.2 A partial uniqueness condition for third-order tensors based on matrix partition

Based on Theorems 3.2-3.4, we proved a partial uniqueness condition for $\boldsymbol{B}$ and $\boldsymbol{C}$. Consider first a partition of matrix $\boldsymbol{A}$ in N submatrices:

$$\boldsymbol{A}\boldsymbol{\Pi}_A = [\boldsymbol{A}_1|\dots|\boldsymbol{A}_N], \quad \text{with} \quad \sum_{n=1}^{N} R_n = R \tag{3.16}$$

where R_n is the number of columns of the submatrix $\boldsymbol{A}_n$ and $\boldsymbol{\Pi}_A$ is a permutation matrix. The partition is such that

$$\text{span}\boldsymbol{A} = \text{span}\boldsymbol{A}_1 \oplus \cdots \oplus \text{span}\boldsymbol{A}_N \tag{3.17}$$

where "$\oplus$" denotes the *direct sum* of the subspaces. Hence, for each $1 \leq n \leq N$ it holds [Hoffman71]

$$\text{span}\boldsymbol{A}_n \cap \left(\cup_{\substack{1\leq i\leq N \\ i\neq n}} \text{span}\boldsymbol{A}_i\right) = \{\boldsymbol{0}\} \tag{3.18}$$

Consider also

$$\boldsymbol{B}\boldsymbol{\Pi}_A = [\boldsymbol{B}_1|\dots|\boldsymbol{B}_N] \quad \text{and} \quad \boldsymbol{C}\boldsymbol{\Pi}_A = [\boldsymbol{C}_1|\dots|\boldsymbol{C}_N] \tag{3.19}$$

the partitioned matrices for the two other modes corresponding to the partition of $\boldsymbol{A}$. Then our partial uniqueness condition on $\boldsymbol{B}$ and $\boldsymbol{C}$ can be stated as follows.

Theorem 3.5 ([Guo12])**.** *Consider a partition of $\boldsymbol{A}$ into N submatrices meeting the conditions above. If $\boldsymbol{A}$ is essentially unique, then $\boldsymbol{\mathcal{X}}$ can be uniquely decomposed into the sum of N R_n-component lower rank tensors $[\![\boldsymbol{A}_n, \boldsymbol{B}_n, \boldsymbol{C}_n]\!]$. Inside each of these lower rank tensors, the first mode loadings, i.e., the columns of $\boldsymbol{A}_n$, can be uniquely determined. The submatrices $\boldsymbol{B}_n$ and $\boldsymbol{C}_n$ can be uniquely determined if the R_n-component CPD of $[\![\boldsymbol{A}_n, \boldsymbol{B}_n, \boldsymbol{C}_n]\!]$ is unique.*

In other words, Theorem 3.5 implies that if one of the conditions given by Theorem (3.2)-(3.4) is satisfied then the identifiability problem of $[\![\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}]\!]$ can be divided into N independent identifiability subproblems of lower rank CP models $[\![\boldsymbol{A}_n, \boldsymbol{B}_n, \boldsymbol{C}_n]\!]$, with $n = 1, \dots, N$.

A direct consequence of this theorem is that if a submatrix $\boldsymbol{A}_n$ in the partition of $\boldsymbol{A}$ has only one column, then the associated loadings in the second and the third mode are uniquely identifiable. This is obvious since the associated lower rank tensor is a rank-1 three-way array which is proven to be essentially unique [Harshman72].

In [Bro09, tenBerge04] the authors explored the cases where $\boldsymbol{B}$ and $\boldsymbol{C}$ are full column rank. Obviously, our conditions are not restricted to this case and hold even if the two modes present linear dependencies but only one mode ($\boldsymbol{A}$) has proportional loadings. This will be illustrated next by some numerical examples. A discussion on the case with collinearity in more than one mode is provided in [Guo12] and in [Stegeman09b].

Numerical examples and comparisons to other uniqueness conditions for CPD such as the ones proposed in [Kruskal77, Stegeman09b], can be found in [Guo12].

3.4 Identifiability of the sparse PARALIND model

Let us consider the PARALIND version of the third-model CPD model (3.8), given by:

$$\boldsymbol{\mathcal{X}} = [\![\widetilde{\boldsymbol{A}}\boldsymbol{\Psi}, \widetilde{\boldsymbol{B}}\boldsymbol{\Phi}, \widetilde{\boldsymbol{C}}\boldsymbol{\Omega}]\!], \tag{3.20}$$

where matrices $\boldsymbol{\Psi}(R_1 \times R)$, $\boldsymbol{\Phi}(R_2 \times R)$, $\boldsymbol{\Omega}(R_3 \times R)$, contain the linear dependency patterns between the columns of $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, respectively and $\widetilde{\boldsymbol{A}}(I \times R_1)$, $\widetilde{\boldsymbol{B}}(J \times R_2)$ and $\widetilde{\boldsymbol{C}}(K \times R_3)$ are full column rank matrices. According to [Bro09], identifiability of the PARALIND model is essentially the same as that of the CPD model. If the interaction matrices are fixed and known, identifiability conditions specific to PARALIND can be found in [Stegeman09b]. If these interaction matrices are not known, the identifiability problem can be much more complicated. In particular, it may happen that only some components of the three matrices, or only one matrix (among the three) are identifiable, resulting in partial uniqueness or uni-mode uniqueness results, as we saw in the previous section.

Assume next that the uniqueness of matrix $\boldsymbol{A}$ is fulfilled and that we aim at estimating the constraint matrix $\boldsymbol{\Psi}$ together with the full column rank matrix $\widetilde{\boldsymbol{A}}$. The identifiability of $\boldsymbol{\Psi}$ and $\widetilde{\boldsymbol{A}}$ comes down to the uniqueness of the bilinear decomposition $\boldsymbol{A} = \widetilde{\boldsymbol{A}}\boldsymbol{\Psi}$. Without any further constraints, such a decomposition is not unique since an alternative decomposition can be obtained as $\boldsymbol{A} = \widetilde{\boldsymbol{A}}\boldsymbol{\Psi} = (\widetilde{\boldsymbol{A}}\boldsymbol{T}^{-1})(\boldsymbol{T}\boldsymbol{\Psi}) = \widetilde{\boldsymbol{A}}'\boldsymbol{\Psi}'$, for any non-singular matrix $\boldsymbol{T}$. By imposing sparsity on the constraint matrix $\boldsymbol{\Psi}$ (which should have a minimum number of non-zero entries), we try to explain the rank deficiency of matrix $\boldsymbol{A}$ by considering the the simplest dependency pattern between its columns.

We address next the following two questions:

- when is the matrix $\widetilde{\boldsymbol{A}}$, yielding the sparsest $\boldsymbol{\Psi}$, a submatrix of $\boldsymbol{A}$?
- when is the decomposition $\boldsymbol{A} = \widetilde{\boldsymbol{A}}\boldsymbol{\Psi}$ unique ?

3.4.1 Choosing a basis of the column space of $\boldsymbol{A}$ that yields the sparsest $\boldsymbol{\Psi}$

We provide next some results on what is the "best" matrix $\widetilde{\boldsymbol{A}}$ for having the sparsest $\boldsymbol{\Psi}$ matrix. Let $\boldsymbol{A}$ be a matrix of dimension $(M \times N), M \geq N$ and let $\text{rank}\{\boldsymbol{A}\} \leq N$. We aim at finding a factorization of the matrix $\boldsymbol{A} = \widetilde{\boldsymbol{A}}\boldsymbol{\Psi}$, where $\widetilde{\boldsymbol{A}}$ is a $(M \times \text{rank}\{\boldsymbol{A}\})$ (tall) matrix and $\boldsymbol{\Psi}$ is $(\text{rank}\{\boldsymbol{A}\} \times N)$ (fat) matrix. The considered factorization problem is known to be subject to permutation and scale ambiguities. To remove scale ambiguities we impose to have the maximum value of each column of $\boldsymbol{\Psi}$ equal to 1.

Let us denote the set of admissible bases of the column space of $\boldsymbol{A}$ by $\mathcal{A} = \{\widetilde{\boldsymbol{A}}$ of size $(M \times \text{rank}\{\boldsymbol{A}\})$ / span$\widetilde{\boldsymbol{A}} = \text{span}\boldsymbol{A}\}$. The problem of finding the factorization of $\boldsymbol{A}$ having the sparsest $\boldsymbol{\Psi}$ can then be formulated as follows:

$$\min_{(\widetilde{\boldsymbol{A}},\boldsymbol{\Psi}),\ \widetilde{\boldsymbol{A}}\in\mathcal{A},\ \boldsymbol{A}=\widetilde{\boldsymbol{A}}\boldsymbol{\Psi}} \|\boldsymbol{\Psi}\|_0 \tag{3.21}$$

where $\|\boldsymbol{\Psi}\|_0$ stands for the ℓ_0 pseudo-norm of matrix $\boldsymbol{\Psi}$, that is the number of non-zero entries of $\boldsymbol{\Psi}$.

Proposition 3.1 ([Miron15a]). *Let $\boldsymbol{A} = \widetilde{\boldsymbol{A}}_1\boldsymbol{\Psi}_1 = \widetilde{\boldsymbol{A}}_2\boldsymbol{\Psi}_2$ where both $\widetilde{\boldsymbol{A}}_1$ and $\widetilde{\boldsymbol{A}}_2$ are full column-rank matrices in $\mathcal{A}$ such that $\widetilde{\boldsymbol{A}}_1$ is composed of $\text{rank}\{\boldsymbol{A}\} \leq N$ independent columns of $\boldsymbol{A}$ and $\widetilde{\boldsymbol{A}}_2$ is composed of $\text{rank}\{\boldsymbol{A}\} \leq N$ independent linear combination of the columns of $\boldsymbol{A}_1$ which are not proportional to the columns of $\boldsymbol{A}$. If $\text{rank}\{\boldsymbol{A}\}$ satisfies $\text{rank}\{\boldsymbol{A}\}^2 - (N+1)\,\text{rank}\{\boldsymbol{A}\} + 2N \geq 0$, then $\|\boldsymbol{\Psi}_1\|_0 \leq \|\boldsymbol{\Psi}_2\|_0$.*

Remark 3.1 ([Miron15a]). *It can be noticed that $\text{rank}\{\boldsymbol{A}\}^2 - (N+1)\,\text{rank}\{\boldsymbol{A}\} + 2N \geq 0$ is satisfied for all $\text{rank}\{\boldsymbol{A}\} \leq N \leq 6$. This is no longer true when $N > 6$.*

Remark 3.2 ([Miron15a]). *The result of proposition 3.1 is based on the worst case scenario since it corresponds to the least favorable case $\|\boldsymbol{\psi}_1(i)\|_0 \leq \text{rank}\{\boldsymbol{A}\}$. This condition can be relaxed by imposing a more favorable situation such as : $\|\boldsymbol{\psi}_1(i)\|_0 \leq \text{rank}\{\boldsymbol{A}\} - k$ which results in $\text{rank}\{\boldsymbol{A}\} + (N - \text{rank}\{\boldsymbol{A}\})(\text{rank}\{\boldsymbol{A}\} - k) \leq 2N$, that is $\text{rank}\{\boldsymbol{A}\}^2 - (N+k+1)\,\text{rank}\{\boldsymbol{A}\} + (2+k)N \geq 0$.*

Remark 3.3 ([Miron15a]). *In proposition 3.1, it is assumed that $\widetilde{\boldsymbol{A}}_2$ does not include any column of $\boldsymbol{A}$. Let us now examine what happens if $\widetilde{\boldsymbol{A}}_2$ does include a number of $l \leq \text{rank}\{\boldsymbol{A}\}$ columns of $\boldsymbol{A}$. In such a case, the sufficient condition for having $\|\boldsymbol{\Psi}_1\|_0 \leq \|\boldsymbol{\Psi}_2\|_0$ is $\text{rank}\{\boldsymbol{A}\}^2 - (N+k+1)\,\text{rank}\{\boldsymbol{A}\} + (2+k)N \geq l$.*

The results corresponding to Proposition 3.1 and Remark 3.2 are shown in Figure 3.1 for different values of N. The bottom curve corresponds to the case of Proposition 3.1 ($k = 0$). The other curves are obtained for increasing values of k. For all the cases corresponding to the values of the plotted curves greater that the threshold (set to 0), the solutions obtained by considering independent columns of $\boldsymbol{A}$ are sparser than those obtained by considering linear combinations of columns of $\boldsymbol{A}$ which are not proportional to the columns of $\boldsymbol{A}$. As mentioned in Remark 3.1, Proposition 3.1 is always true for $N \leq 6$. The case $N = 6$ is depicted on the left-hand side of Figure 3.1. The case of Remark 3.3, which corresponds to having l of columns of $\boldsymbol{A}$ in $\widetilde{\boldsymbol{A}}_2$, is simply obtained by shifting the threshold to a value equal to l.

3.4.2 Uniqueness of the sparsest decomposition

In this part we aim at studying the uniqueness of the sparsest decomposition $\boldsymbol{A} = \widetilde{\boldsymbol{A}}\boldsymbol{\Psi}$. For all the uniqueness results presented in this part we assume that the full column rank

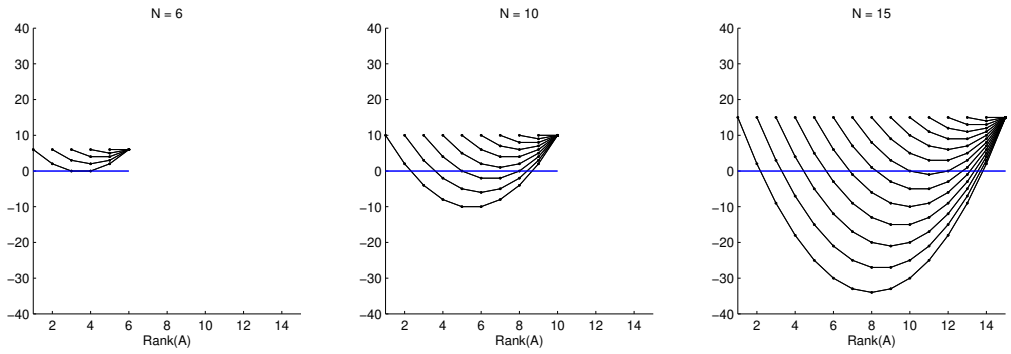


Figure 3.1: Illustration of the sparsity properties : the bottom curve corresponds to the case of proposition 3.1. The others correspond to the case of remark 2 for different values of k. As k is increasing, there are much more situations in which the condition $\|\boldsymbol{\Psi}_1\|_0 \leq \|\boldsymbol{\Psi}_2\|_0$ is fulfilled.

matrix $\widetilde{\boldsymbol{A}}$ yielding the sparsest solution is a submatrix of $\boldsymbol{A}$. The uniqueness properties of the case where $\widetilde{\boldsymbol{A}}$ is not a sub matrix of $\boldsymbol{A}$ are much more difficult to analyze. Thus, the matrix $\boldsymbol{A}$ can be written as: $\boldsymbol{A} = [\widetilde{\boldsymbol{A}} \quad \breve{\boldsymbol{A}}] = [\widetilde{\boldsymbol{a}}_1, \cdots, \widetilde{\boldsymbol{a}}_{\text{rank}\{\boldsymbol{A}\}}, \breve{\boldsymbol{a}}_{\text{rank}\{\boldsymbol{A}\}+1}, \cdots, \breve{\boldsymbol{a}}_N]$. Defining the matrix $\boldsymbol{A}^i = [\widetilde{\boldsymbol{A}} \quad \breve{\boldsymbol{a}}_i]$, with $i = \text{rank}\{\boldsymbol{A}\}+1, \ldots, N$, its Kruskal-rank satisfies $\text{krank}\{\boldsymbol{A}^i\} \leq \text{krank}\{\boldsymbol{A}\} \leq \text{rank}\{\boldsymbol{A}\}$ and each $\breve{\boldsymbol{a}}_i$ can be expressed as a linear combination of exactly $\text{krank}\{\boldsymbol{A}^i\}$ columns of $\widetilde{\boldsymbol{A}}$. In other words : $\breve{\boldsymbol{a}}_i = \widetilde{\boldsymbol{A}}\boldsymbol{\psi}(i)$ and $\|\boldsymbol{\psi}(i)\|_0 = \text{krank}\{\boldsymbol{A}^i\}$. We define the set $\widetilde{\mathcal{A}} = \{\widetilde{\boldsymbol{A}} \;/\; \widetilde{\boldsymbol{A}} \text{ is a submatrix of } \boldsymbol{A} \text{ of size } (M \times \text{rank}\{\boldsymbol{A}\}), \text{span}\widetilde{\boldsymbol{A}} = \text{span}\boldsymbol{A}\} \subset \mathcal{A}$.

Proposition 3.2 ([Miron15a]). *Let $\widetilde{\boldsymbol{A}}_1 \neq \widetilde{\boldsymbol{A}}_2$ two matrices of $\widetilde{\mathcal{A}}$ and $\boldsymbol{\Psi}_1$ and $\boldsymbol{\Psi}_2$ the two matrices satisfying $\boldsymbol{A} = \widetilde{\boldsymbol{A}}_1\boldsymbol{\Psi}_1 = \widetilde{\boldsymbol{A}}_2\boldsymbol{\Psi}_2$. Then $\|\boldsymbol{\Psi}_1\|_0 \leq \|\boldsymbol{\Psi}_2\|_0$ if and only if $\sum_{\text{rank}\{\boldsymbol{A}\}+1}^{N} \text{krank}\{\boldsymbol{A}_1^i\} \leq \sum_{\text{rank}\{\boldsymbol{A}\}+1}^{N} \text{krank}\{\boldsymbol{A}_2^i\}$.*

A straightforward extension of proposition 3.2 is given by the following proposition which gives the condition for having the sparsest and possibly unique decomposition of $\boldsymbol{A}$:

Proposition 3.3 ([Miron15a]). *If $\exists\, \widetilde{\boldsymbol{A}}_0 \in \widetilde{\mathcal{A}}$ such as $\forall \widetilde{\boldsymbol{A}} \in \widetilde{\mathcal{A}}, \widetilde{\boldsymbol{A}} \neq \widetilde{\boldsymbol{A}}_0$, $\sum_{\text{rank}\{\boldsymbol{A}\}+1}^{N} \text{krank}\{\boldsymbol{A}_0^i\} \leq \sum_{\text{rank}\{\boldsymbol{A}\}+1}^{N} \text{krank}\{\boldsymbol{A}^i\}$ then the decomposition $\boldsymbol{A} = \widetilde{\boldsymbol{A}}_0\boldsymbol{\Psi}_0$ is the sparsest decomposition. If the inequality is strict ($<$), the sparsest decomposition is unique.*

It should be noted that Proposition 3.3 does not provide any effective means to find the sparsest decomposition of $\boldsymbol{A}$. Indeed, finding it would require to find all the possible basis consisting in $\text{rank}\{\boldsymbol{A}\}$ columns of $\boldsymbol{A}$ which is actually an NP-hard problem.

The result of Proposition 3.3 can be specialized into the following cases:

- the linear dependencies between the columns of $\boldsymbol{A}$ are only colinearities. In that case $\widetilde{\mathcal{A}}$ only includes a single element since any selection of $\text{rank}\{\boldsymbol{A}\}$ independent columns of $\boldsymbol{A}$ will result in the same basis up to scale and permutation. Thus the decomposition is unique.

- $\boldsymbol{A}$ has a Kruskal-rank equal to its rank *i.e.* $\text{krank}\{\boldsymbol{A}\} = \text{rank}\{\boldsymbol{A}\}$. In that case, any selection of $\text{rank}\{\boldsymbol{A}\}$ columns of $\boldsymbol{A}$ is a basis and $\forall \widetilde{\boldsymbol{A}} \in \widetilde{\mathcal{A}}$, $\text{krank}\{\boldsymbol{A}^i\} = \text{krank}\{\boldsymbol{A}\}, \forall i = \text{rank}\{\boldsymbol{A}\} + 1, \cdots, N$. Thus $\forall \widetilde{\boldsymbol{A}} \in \widetilde{\mathcal{A}}$, $\|\boldsymbol{\Psi}\|_0 = (N - \text{rank}\{\boldsymbol{A}\} + 1)\,\text{rank}\{\boldsymbol{A}\}$.

In practice, having a unique sparsest matrix $\boldsymbol{\Psi}$ is not crucial. Indeed, from a PARALIND point of view, having a number of decompositions yielding the same degree of sparsity simply means that all these decompositions are equivalent.

3.5 Uniqueness of fourth-order CPD in the presence of collinear loadings

As the number of dimensions (the order) of a tensor increases, the uniqueness of its CPD becomes less problematic. We have seen in the previous sections that, for third-order tensors, the presence of collinear loadings in one mode is sufficient for the non-uniqueness of its CPD. However this is no longer the case for fourth-order tensors, where CPD uniqueness can be achieved even in the presence of collinearities. This section analyzes the CPD uniqueness in the case where the only possible linear dependencies between the columns of the loading matrices take the form of collinear loadings. We termed this particular dependency pattern the *collinear loadings only (CLO)* assumption. This has many practical applications in signal processing [Sidiropoulos01, deAlmeida08b], chemometrics [Bro09], etc.

We consider the CPD of a fourth-order tensor given by:

$$\boldsymbol{\mathcal{X}} = [\![\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D}]\!], \tag{3.22}$$

with $\boldsymbol{A} = [\boldsymbol{a}_1 \ \cdots \ \boldsymbol{a}_R]$ $(I \times R)$, $\boldsymbol{B} = [\boldsymbol{b}_1 \ \cdots \ \boldsymbol{b}_R]$ $(J \times R)$, $\boldsymbol{C} = [\boldsymbol{c}_1 \ \cdots \ \boldsymbol{c}_R]$ $(K \times R)$ and $\boldsymbol{D} = [\boldsymbol{d}_1 \ \cdots \ \boldsymbol{d}_R]$ $(L \times R)$. With these notations, an alternative way of writing the CPD model (3.22) is:

$$\boldsymbol{\mathcal{X}} = \sum_{r=1}^{R} \boldsymbol{a}_r \circ \boldsymbol{b}_r \circ \boldsymbol{c}_r \circ \boldsymbol{d}_r. \tag{3.23}$$

We denote the fact that two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are collinear (proportional) by $\boldsymbol{x} \propto \boldsymbol{y}$ and the contrary by $\boldsymbol{x} \not\propto \boldsymbol{y}$. Thus, the *CLO* assumption for one of the loading matrices, say $\boldsymbol{A}$, can be formulated as: *$\boldsymbol{A}$ is rank deficient if and only if one (or more columns) of the loading matrix is (or are) proportional to another column.* In that case:

$$\exists \quad r_1 \neq r_2 \text{ such as } \boldsymbol{a}_{r_1} \propto \boldsymbol{a}_{r_2}, \tag{3.24}$$

which results in $\text{krank}\{\boldsymbol{A}\} = 1$. As a consequence, under the CLO assumption, the krank of a loading matrix can only be equal to either R or 1 while its rank may vary between 1 and R.

3.5.1 The rank of the Khatri-Rao product of two matrices under the CLO assumption

In this section, a necessary and sufficient condition ensuring the full column rank of the Khatri-Rao product of two CLO matrices is provided. This rank is upper-bounded by $\text{rank}\{\boldsymbol{A}\odot\boldsymbol{B}\} \leq \text{rank}\{\boldsymbol{A}\}\,\text{rank}\{\boldsymbol{B}\}$. Thus, from now on we assume that $R \leq \text{rank}\{\boldsymbol{A}\}\,\text{rank}\{\boldsymbol{B}\}$. First we prove a necessary and sufficient condition under which two vector Kronecker products are collinear.

Proposition 3.4 ([Brie11]). *Let $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{v}$ be non zero vectors. Then, $(\boldsymbol{x}\otimes\boldsymbol{u}) \not\propto (\boldsymbol{y}\otimes\boldsymbol{v})$ iff $\boldsymbol{x} \not\propto \boldsymbol{y}$ or $\boldsymbol{u} \not\propto \boldsymbol{v}$.*

We provide next a lemma stating the condition under which the Khatri-Rao product of two matrices satisfying the CLO assumption is full column rank.

Lemma 3.2 ([Brie11]). *Consider the two matrices $\boldsymbol{A} = [\boldsymbol{a}_1 \cdots \boldsymbol{a}_R]$ and $\boldsymbol{B} = [\boldsymbol{b}_1 \cdots \boldsymbol{b}_R]$ satisfying the CLO assumption. $\boldsymbol{A}\odot\boldsymbol{B}$ is full column rank iff $\forall n \neq m$, $\boldsymbol{a}_n \not\propto \boldsymbol{a}_m$ or $\boldsymbol{b}_n \not\propto \boldsymbol{b}_m$.*

The proof of Lemma 3.2 is done by induction and exploits the result of Proposition 3.4.

3.5.2 A sufficient condition for the uniqueness of the 4-way CPD model with collinear loadings in at most three modes

We re-write the CPD model (3.22) by separating the set of columns of, say $\boldsymbol{A}$, in N sub-matrices $\boldsymbol{A}_i$, each sub-matrix containing only collinear columns and a sub-matrix $\breve{\boldsymbol{A}}$ containing the rest of the non-proportional columns. The matrix $\boldsymbol{A}$ can be partitioned (after re-ordering) as:

$$\boldsymbol{A} = [\boldsymbol{A}_1|\cdots|\boldsymbol{A}_N|\ \breve{\boldsymbol{A}}],$$

with $\boldsymbol{A_i} = [\boldsymbol{a}_i \cdots \boldsymbol{a}_i]$ of dimensions $(I \times R_i)$ and $\breve{\boldsymbol{A}} = [\boldsymbol{a}_{\widetilde{R}+1} \cdots \boldsymbol{a}_R]$ of dimensions $(I \times \breve{R})$. We denote $\widetilde{R} = \sum_i R_i$ and $\breve{R} = R - \widetilde{R}$. The matrices $\boldsymbol{B}$, $\boldsymbol{C}$ and $\boldsymbol{D}$ can then be written using the corresponding partition of $\boldsymbol{A}$ as: $\boldsymbol{B} = [\boldsymbol{B}_1|\cdots|\boldsymbol{B}_N|\ \breve{\boldsymbol{B}}]$, $\boldsymbol{C} = [\boldsymbol{C}_1|\cdots|\boldsymbol{C}_N|\ \breve{\boldsymbol{C}}]$ and $\boldsymbol{D} = [\boldsymbol{D}_1|\cdots|\boldsymbol{D}_N|\ \breve{\boldsymbol{D}}]$. With these notations, the 4-way CPD model can be re-written as:

$$[\![\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D}]\!] = \sum_{i=1}^{N} \boldsymbol{a}_i \circ [\![\boldsymbol{B}_i, \boldsymbol{C}_i, \boldsymbol{D}_i]\!] + [\![\breve{\boldsymbol{A}}, \breve{\boldsymbol{B}}, \breve{\boldsymbol{C}}, \breve{\boldsymbol{D}}]\!]. \tag{3.25}$$

The uniqueness result is summarized by the following theorem:

Theorem 3.6 ([Brie11]). *Consider a 4-way array $\boldsymbol{\mathcal{X}}$ given by (3.25) and admitting an order-R CPD decomposition. The CPD decomposition is essentially unique if but not necessarily:*

- *one loading matrix, say $\boldsymbol{D}$, is full column rank;*
- *the other loading matrices $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}$ satisfy the CLO assumption;*

- $\forall n \neq m,\ (\boldsymbol{a}_n \not\propto \boldsymbol{a}_m, \boldsymbol{b}_n \not\propto \boldsymbol{b}_m)$ or $(\boldsymbol{a}_n \not\propto \boldsymbol{a}_m, \boldsymbol{c}_n \not\propto \boldsymbol{c}_m)$ or $(\boldsymbol{b}_n \not\propto \boldsymbol{b}_m, \boldsymbol{c}_n \not\propto \boldsymbol{c}_m)$.

The proof is based on the expression (3.25) of the CPD and consists in the following two steps:

- proving the partial uniqueness of the corresponding 3-way CPD model obtained by unfolding along one dimension. By doing so, we ensure that the original 4-way problem is transformed into a number of independent 3-way CPD problems;
- proving the uniqueness of each subproblem.

An application of this uniqueness result on a real dataset generated by biosensors can be found in [Brie11].

3.6 Uniqueness of tensor train decompositions with linear dependencies

Q-way arrays with $Q \gg 3$, are nowadays frequently encountered in a large number of applications. Meanwhile, the storage memory and algorithmic complexity grow exponentially with the order of the tensor, a phenomenon known as the "the curse of dimensionality". For example, the rank-R CPD of a Q-order tensor with dimensions $N_1 = \cdots = N_Q = N$, has a computational cost of $\mathcal{O}(QRN^Q)$. A solution to break the "curse of dimensionality" is to break a high-order tensor in a sequence of lower-order tensors, which leads to Tensor Networks (TNs) models [Cichocki14]. Tensor Train Decomposition (TTD) [Oseledets11] is one the most compact and simple TNs. Indeed, TTD breaks a high Q-order tensor into a set of Q lower-order tensors, called TT-cores. These TT-cores have orders at most equal to 3.

Definition 3.1 (Tensor Train Decomposition)**.** *A Q-order tensor of size $N_1 \times \ldots \times N_Q$ that follows a tensor train decomposition [Oseledets11] of TT-ranks $\{R_1, \ldots, R_{Q-1}\}$ admits the following definition:*

$$\boldsymbol{\mathcal{X}} = \boldsymbol{G}_1 \overset{1}{\bullet}_{2} \boldsymbol{\mathcal{G}}_2 \overset{1}{\bullet}_{3} \boldsymbol{\mathcal{G}}_3 \overset{1}{\bullet}_{4} \cdots \overset{1}{\bullet}_{Q-1} \boldsymbol{\mathcal{G}}_{Q-1} \overset{1}{\bullet}_{Q} \boldsymbol{G}_Q, \tag{3.26}$$

where the TT-cores $\boldsymbol{G}_1, \boldsymbol{\mathcal{G}}_q$, and $\boldsymbol{G}_Q$ are, respectively, of dimensions $N_1 \times R_1$, $R_{q-1} \times N_q \times R_q$, and $R_{Q-1} \times N_Q$, for $2 \leq q \leq Q-1$, and we have $\text{rank}\{\boldsymbol{G}_1\} = R_1$, $\text{rank}\{\boldsymbol{G}_Q\} = R_{Q-1}$, $\text{rank}\{\text{unfold}_1 \boldsymbol{\mathcal{G}}_q\} = R_{q-1}$, *and* $\text{rank}\{\text{unfold}_3 \boldsymbol{\mathcal{G}}_q\} = R_q$.

3.6.1 Equivalence between TTD and CPD

In a recent work [Zniyed18], an equivalence between the CPD and the TTD was proposed. In fact, it has been shown that a Q-order CPD of rank-R is equivalent to a train of 3-order CPD(s) of rank-R, denoted TT-CPD. This result makes it easier to jointly reduce the

dimension and estimate the CPD factors using the TT-cores when the original tensor has a high order. Otherwise, when Q is high, the CPD factors estimation becomes a difficult task using ALS-based techniques [Bro99].

In [Zniyed19b], we focused on the case where linear dependencies are present between the columns on the factor matrices of the TT-CPD model. We provided partial and full uniqueness conditions for the new model, that we called TT-PARALIND, and compared these condition to the uniqueness conditions fo the classical Q-order CPD.

It is straightforward to see that the TTD of $\boldsymbol{\mathcal{X}}$ in (3.26) is not unique since

$$\boldsymbol{\mathcal{X}} = \boldsymbol{A}_1 \overset{1}{\bullet}_{2} \boldsymbol{\mathcal{A}}_2 \overset{1}{\bullet}_{3} \boldsymbol{\mathcal{A}}_3 \overset{1}{\bullet}_{4} \cdots \overset{1}{\bullet}_{Q-1} \boldsymbol{\mathcal{A}}_{Q-1} \overset{1}{\bullet}_{Q} \boldsymbol{A}_Q, \tag{3.27}$$

where

$$\boldsymbol{A}_1 = \boldsymbol{G}_1 \boldsymbol{U}_1^{-1}, \tag{3.28}$$

$$\boldsymbol{A}_Q = \boldsymbol{U}_{Q-1} \boldsymbol{G}_Q, \tag{3.29}$$

$$\boldsymbol{\mathcal{A}}_q = \boldsymbol{U}_{q-1} \overset{1}{\bullet}_{2} \boldsymbol{\mathcal{G}}_q \overset{1}{\bullet}_{3} \boldsymbol{U}_q^{-1}. \tag{3.30}$$

For $1 \leq q \leq Q-1$, $\boldsymbol{U}_q$ are square nonsingular matrices of dimension $R_q \times R_q$.

In practice, the TTD is performed thanks to the state-of-art TT-SVD algorithm [Oseledets11]. It is a sequential algorithm that recovers the TT-cores $\boldsymbol{\mathcal{G}}_q$ based on $(Q-1)$ SVDs applied to several matrix-based reshapings using the original tensor $\boldsymbol{\mathcal{X}}$. This algorithm allows to recover the true TT-cores up to a post and pre-multiplication by transformation (*change-of-basis*) matrices due to the extraction of dominant subspaces when using the SVD.

Consider Q-order tensor $\boldsymbol{\mathcal{X}}$ of size $N_1 \times \cdots \times N_Q$ that follows a rank-R CPD:

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{I}}_{Q,R} \overset{1}{\bullet}_{1} \boldsymbol{P}_1 \overset{1}{\bullet}_{2} \boldsymbol{P}_2 \cdots \overset{1}{\bullet}_{Q} \boldsymbol{P}_Q, \tag{3.31}$$

where the loading matrices $\boldsymbol{P}_q$ are of size $N_q \times R$. It was shown in [Zniyed18, Zniyed19a] that if the loading matrices $\boldsymbol{P}_q$ are full-column rank for $1 \leq q \leq Q$, then they can be recovered from the TT-cores by order-3 CPD decompositions.

3.6.2 Uniqueness of the TT-PARALIND model

We study next the case where linear dependencies are present between the columns of the loading matrices of (3.31). Thus, following the PARALIND model [Bro09], a loading matrix $\boldsymbol{P}_q$ can be expressed as:

$$\boldsymbol{P}_q = \widetilde{\boldsymbol{P}}_q \boldsymbol{\Phi}_q, \tag{3.32}$$

where $\widetilde{\boldsymbol{P}}_q$ is full column rank of size $N_q \times R_q$ ($R_q \leq R$) and $\boldsymbol{\Phi}_q$ is a rank deficient matrix of size $R_q \times R$ containing the dependency pattern between the columns of $\widetilde{\boldsymbol{P}}_q$.

Theorem 3.7 (TT-PARALIND [Zniyed19b]). *Decomposing tensor $\boldsymbol{\mathcal{X}}$ in (3.31) into a TT format, where $\boldsymbol{P}_1$ and $\boldsymbol{P}_Q$ are full column rank matrices, and $\boldsymbol{P}_q$ $(2 \leq q \leq Q-1)$ follow (3.32), recovers the estimated TT-cores such that*

$$\boldsymbol{G}_1 = \boldsymbol{P}_1 \boldsymbol{U}_1^{-1}, \tag{3.33}$$

$$\boldsymbol{\mathcal{G}}_q = \boldsymbol{\mathcal{I}}_{3,R} \underset{1}{\bullet} \boldsymbol{U}_{q-1} \underset{2}{\bullet} (\widetilde{\boldsymbol{P}}_q \boldsymbol{\Phi}_q) \underset{3}{\bullet} \boldsymbol{U}_q^{-\mathsf{T}}, \ 2 \leq q \leq Q-1 \tag{3.34}$$

$$\boldsymbol{G}_Q = \boldsymbol{U}_{Q-1} \boldsymbol{P}_Q^{\mathsf{T}}, \tag{3.35}$$

where, for $1 \leq q \leq Q-1$, $\boldsymbol{U}_q$ is a square $R \times R$ nonsingular matrix. The TT-cores $\boldsymbol{G}_1, \boldsymbol{\mathcal{G}}_q$, and $\boldsymbol{G}_Q$ are, respectively, of dimensions $N_1 \times R, R \times N_q \times R$, and $R \times N_Q$, given TT-ranks all equal to R.

We called this new tensor model given by Theorem 3.7 *TT-PARALIND* and proved in [Zniyed19b] several uniqueness conditions for its identifiability.

Theorem 3.8 (Partial uniqueness [Zniyed19b]). *The loading matrix $\boldsymbol{P}_q$ can be uniquely recovered from the estimated TT decomposition of $\boldsymbol{\mathcal{X}}$ if there exist q_1 and q_2 $(q_1 \neq q_2 \neq q)$, such that:*

$$\begin{cases} \operatorname{rank}\{\boldsymbol{P}_{q_1}\} = \operatorname{rank}\{\boldsymbol{P}_{q_2}\} = R, \\ \operatorname{rank}\{\boldsymbol{P}_q\} \geq 2. \end{cases} \tag{3.36}$$

Theorem 3.9 (Full uniqueness [Zniyed19b]). *The loading matrices $\boldsymbol{P}_1, \ldots, \boldsymbol{P}_Q$ can be uniquely recovered from the estimated TT-cores $\boldsymbol{G}_1, \boldsymbol{\mathcal{G}}_2, \ldots, \boldsymbol{\mathcal{G}}_{Q-1}, \boldsymbol{G}_Q$ if:*

$$\begin{cases} \operatorname{rank}\{\boldsymbol{P}_1\} = \operatorname{rank}\{\boldsymbol{P}_Q\} = R \\ \operatorname{rank}\{\boldsymbol{P}_q\} \geq 2, \quad 2 < q < Q-1 \\ \operatorname{krank}\{\boldsymbol{P}_2\}, \operatorname{krank}\{\boldsymbol{P}_{Q-1}\} \geq 2. \end{cases} \tag{3.37}$$

An interesting question is wether the uniqueness of the TT-PARALIND model requires stronger conditions compared to the equivalent Q-order PARALIND model. It is difficult to give a definite answer to this question but we can get some ideas on this matter by comparing the sufficient uniqueness conditions for the two models in some particular cases.

Discussion Our partial uniqueness conditions for third-order tensors [Guo12] have been generalized to the case of Q-order tensors in [Zhang13]. To the best of our knowledge, these are the most relaxed conditions existing in the literature, for the general case. For a fourth-order PARALIND, the partial uniqueness condition that guarantees that the loading matrix $\boldsymbol{P}_q$ can be uniquely estimated reads ([Zhang13]):

$$\operatorname{rank}\{\boldsymbol{P}_q\} + \sum_{i \neq q} \operatorname{krank}\{\boldsymbol{P}_i\} \geq 2R + 3. \tag{3.38}$$

Meanwhile, the if the partial uniqueness sufficient condition for TT-PARALIND, given by Theorem 3.8 is satisfied, then the condition (3.38) is also satisfied. Regarding the full-uniqueness, if the TT-PARALIND uniqueness conditions of Theorem 3.9 are satisfied, then the Kruskal's condition for the fourth order PARALIND model is satisfied as well. Thus, we can conjecture that the uniqueness conditions for TT-PARALIND are in general more restrictive than those for the classical PARALIND. This can be explained by the fact that TT-PARALIND model requires the estimation of more parameters, *i.e.,* the change-of-basis matrices $\boldsymbol{U}_q$, compared to PARALIND. However, more restrictive conditions is the price to pay for having a numerically stable and efficient estimation algorithm based on the SVD procedure.

3.7 Conclusions

Uniqueness of tensor decompositions in the presence of linear dependencies between the columns of the loading matrices has been a constant preoccupation in the linear algebra community for years. However, despite all these efforts we still do not dispose of easy-to-check, necessary and sufficient uniqueness conditions, except for some very special cases.

The first major contribution of my work (via Ph.D. supervising and academic collaborations) was to propose, based on rank considerations, full and partial uniqueness conditions for third order CPD. These conditions are more relaxed than the ones existing in the literature. For the case when these loading matrices with linear dependencies can be uniquely estimated, we derived several results on how to find the column space basis that yields the sparsest dependency pattern. The second contribution is a sufficient uniqueness condition for fourth-order CPD with collinear loadings in at most three modes. The third major contribution presented in this section is the derivation of the TT-PARALIND model and of two identifiability conditions for this model.

Chapter 4

Tensor-based DOA estimation with sensor arrays

This chapter is dedicated to an important part of my research: the application of the tensor models and decompositions to sensor-array processing for direction of arrival (DOA) estimation. This topic is in the continuity of my Ph.D. work on source detection and localization using multivariate arrays.

4.1 The DOA estimation problem

A fundamental problem in array processing is the estimation of the DOAs for multiple sources impinging on an array of sensors. Assume that K narrow-band far-field sources are impinging on an array of M identical sensors ($K < M$, in general). The direction of arrival of a source k in a Cartesian coordinate system $OXYZ$, associated with the array is given by the unit vector $\boldsymbol{p}_k = \begin{bmatrix}\sin\theta_k \cos\varphi_k & \sin\theta_k \sin\varphi_k & \cos\theta_k\end{bmatrix}^{\mathsf{T}}$, where θ and φ are the elevation and azimuth angles, respectively, as illustrated on Fig.4.1.

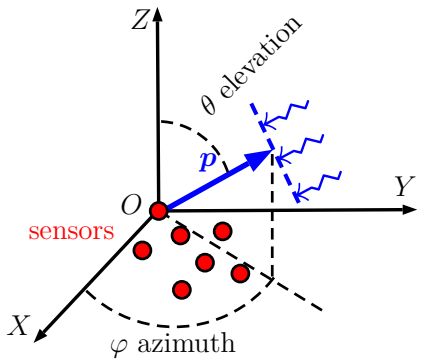


Figure 4.1: DOA estimation set-up

With these notations, a snapshot of the array at time instant t is given by:

$$\boldsymbol{y}(t) = \sum_{k=1}^{K} \boldsymbol{a}(\boldsymbol{k}_k) s_k(t) + \boldsymbol{b}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{b}(t), \tag{4.1}$$

where $\boldsymbol{y}(t)$ is the $(M \times 1)$ output vector of the array, $\boldsymbol{A} = [\boldsymbol{a}(\boldsymbol{p}_1), \ldots, \boldsymbol{a}(\boldsymbol{p}_K)]$ is the $(M \times K)$ steering vectors matrix of the K sources, $\boldsymbol{x}(t) = [s_1(t), \ldots, s_K(t)]^\mathsf{T}$ is the $(K \times 1)$ source signals vector at time instant t, and $\boldsymbol{b}(t)$ is a $(M \times 1)$ vector that accounts for the spatially and temporally white Gaussian noise on the array sensors. If N snapshots $t_1, t_2, \ldots, t_N$ are considered, the output of the array can be expressed in the matrix form as:

$$\boldsymbol{Y} = [\boldsymbol{y}(t_1), \ldots, \boldsymbol{y}(t_N)] = \quad \boldsymbol{A}(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_K)[\boldsymbol{x}(t_1), \ldots, \boldsymbol{x}(t_N)] + \boldsymbol{B} = \boldsymbol{A}(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_K)\boldsymbol{S}^\mathsf{T} + \boldsymbol{B} \tag{4.2}$$

with $\boldsymbol{S} = [\boldsymbol{s}_1, \ldots, \boldsymbol{s}_K]$, a $(N \times K)$ matrix gathering on its columns the N time samples for the K sources, $\boldsymbol{s}_k = [s_p(t_1), \ldots, s_p(t_N)]^\mathsf{T}$, with $k = 1, \ldots, K$ and $\boldsymbol{B} = [\boldsymbol{b}(t_1), \ldots, \boldsymbol{b}(t_N)]$, a $(M \times N)$ noise matrix. The DOA estimation problem consists in finding $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_K$ from the data $\boldsymbol{Y}$. By considering additional diversities, such as poarization, spatial structure, etc., the data matrix $\boldsymbol{Y}$ can be cast as a multi-way array and tensor-based techniques can be used to estimate the DOAs, as illustrated in the next sections of this chapter.

4.2 Tensors for polarized sensor array processing

The results presented in this section are extracted from the Ph.D. of Xijing Guo that I co-supervised with David Brie.

High-resolution techniques such as MUSIC [Bienvenu79, Schmidt81] or ESPRIT [Roy89], introduced in the late '70s and the '80s, gave an impetus to sensor-array signal processing research. A significant number of eigenstucture-based direction finding (DF) algorithms have been proposed since, for various types of sensors and array configurations. In [Sidiropoulos00b], Sidiropoulos *et al.* proposed for the first time a direction-of-arrival (DOA) estimation approach based on a CPD model of the data, and highlighted the link between CPD and ESPRIT.

We have shown in [Guo08, Guo11] that the data recorded on an electromagnetic (EM) vector-sensor array also follows a CPD model. We used the nice properties of the CPD to propose efficient DOA estimation algorithms and to derive bounds on the maximum number of sources that can be detected by such an array.

4.2.1 The trilinear model of the data recorded on an EM vector-sensor array

Consider a sensor array system of M displaced but otherwise identical electromagnetic vector-sensors deployed in the far-field of K narrowband sources. Figure 4.2 illustrates one of these vector-sensors. In accordance, a Cartesian coordinate system is established where the reference vector-sensor is positioned at the origin. It is assumed that the signals are completely polarized and the propagation medium is isotropic and homogeneous. Each of these vector-sensors forms a subarray of the entire sensor system, with the manifold given by [Nehorai94]:

$$\boldsymbol{b}(\varphi,\psi,\alpha,\beta) = \underbrace{\begin{bmatrix} -\sin\varphi & -\cos\varphi\sin\psi \\ \cos\varphi & -\sin\varphi\sin\psi \\ 0 & \cos\psi \\ -\cos\varphi\sin\psi & \sin\varphi \\ -\sin\varphi\sin\psi & -\cos\varphi \\ \cos\psi & 0 \end{bmatrix}}_{\boldsymbol{\Theta}(\varphi,\psi)} \underbrace{\begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} \cos\beta \\ j\sin\beta \end{bmatrix}}_{\boldsymbol{p}(\alpha,\beta)}, \tag{4.3}$$

for an incident signal with DOA (φ,ψ) and polarization (α,β). In (4.3), $\varphi \in [0,2\pi)$ denotes the azimuth angle, $\psi \in [-\pi/2,\pi/2]$ is the elevation angle, $\alpha \in (-\pi/2,\pi/2]$ is the orientation angle of the polarization ellipse and $\beta \in [-\pi/4,\pi/4]$ is the ellipticity angle. Throughout the manuscript, we assume that the DOA's and the polarizations are unknown but deterministic.

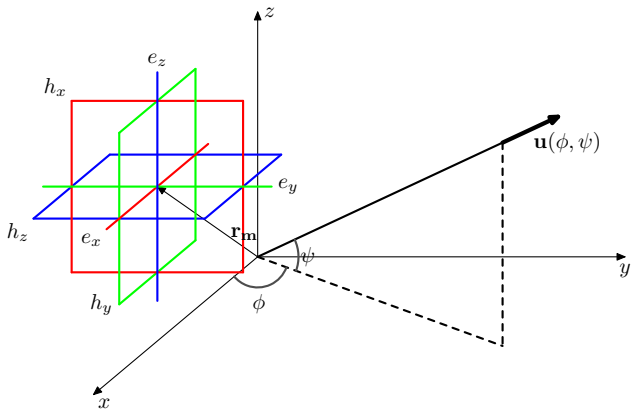


Figure 4.2: Geometry of a vector-sensor: $\{e_x, e_y, e_z\}$ denotes the dipole triad and $\{h_x, h_y, h_z\}$ the loop triad.

We denote by a set of three-dimensional vectors $\{\boldsymbol{r}_m\}_{m=1}^{M}$, the M observation points where the vector-sensors are positioned in the reference frame, and by

$$\boldsymbol{p}(\varphi,\psi) = [\cos\varphi\cos\psi \quad \sin\varphi\cos\psi \quad \sin\psi]^{\mathsf{T}},$$

the unit Poynting vector in the source direction. Here, ψ denotes the co-elevation angle. The expression for the phase shifts induced by the displacements of the mth ($m = 1, \ldots, M$) vector-sensor relative to the reference one is then given by $a_m(\varphi, \psi) = \exp\left\{j2\pi \boldsymbol{r}_m^{\mathsf{T}} \boldsymbol{u}(\varphi, \psi)/\lambda\right\}$, where λ is the wavelength. Furthermore, define $\boldsymbol{a}(\varphi, \psi) = [a_1(\varphi, \psi), \ldots, a_M(\varphi, \psi)]^{\mathsf{T}}$, the steering vector of a virtual scalar sensor array that has the same sensor configuration as the vector-sensor array [Tan96b]. Correspondingly, the steering vector of the entire vector-sensor array system is given by [Nehorai94]

$$\boldsymbol{d}(\varphi, \psi, \alpha, \beta) = \boldsymbol{a}(\varphi, \psi) \otimes \boldsymbol{b}(\varphi, \psi, \alpha, \beta),$$

where "$\otimes$" denotes the Kronecker product. Now suppose that the kth signal impinges on the array from a direction (φ_k, ψ_k), with the polarization parameterized by (α_k, β_k). Inductively, the output of the array at time t, *i.e.*, $\boldsymbol{y}(t)$, is obtained by summing up the contributions from all the K incident signals $s_k(t)$:

$$\boldsymbol{y}(t) = \sum_{k=1}^{K} s_k(t)\boldsymbol{d}(\varphi_k, \psi_k, \alpha_k, \beta_k), = \boldsymbol{D}\boldsymbol{s}(t) + \boldsymbol{n}(t) \tag{4.4}$$

where $\boldsymbol{D} = [\boldsymbol{d}(\varphi_1, \psi_1, \alpha_1, \beta_1), \ldots, \boldsymbol{d}(\varphi_K, \psi_K, \alpha_K, \beta_K)]$ is a $6M \times K$ matrix, $\boldsymbol{n}(t)$ is a vector accounting for the additive noise on the sensors and the model errors and $\boldsymbol{s}(t) = [s_1(t), \ldots, s_K(t)]^{\mathsf{T}}$ is a K-element vector. Let $\boldsymbol{A} = [\boldsymbol{a}(\varphi_1, \psi_1), \ldots, \boldsymbol{a}(\varphi_K, \psi_K)]$, and $\boldsymbol{B} = [\boldsymbol{b}(\varphi_1, \psi_1, \alpha_1, \beta_1), \ldots, \boldsymbol{b}(\varphi_K, \psi_K, \alpha_K, \beta_K)]$. Then, one observes $\boldsymbol{D} = \boldsymbol{A} \odot \boldsymbol{B}$ where "$\odot$" denotes the Khatri-Rao (column-wise Kronecker) product. In the sequel, we will occasionally omit the explicit dependence on $(\varphi_k, \psi_k, \alpha_k, \beta_k)$, for notational simplicity. If N snapshots at the discrete time instants $\{t_n\}_{n=1}^{N}$ are collected, the available data can be organized into a $6M \times N$ matrix $\boldsymbol{Y} = [\boldsymbol{y}(t_1), \ldots, \boldsymbol{y}(t_N)] = \boldsymbol{D}\boldsymbol{S}^{\mathsf{T}} + \boldsymbol{N}$, where $\boldsymbol{S} = [\boldsymbol{s}(t_1), \cdots, \boldsymbol{s}(t_N)]^{\mathsf{T}}$ and $\boldsymbol{N} = [\boldsymbol{n}(t_1), \cdots, \boldsymbol{n}(t_N)]$. Observing the multiple invariances of the data model, $\boldsymbol{Y}$ can also be expressed as [Guo08]:

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y}_1 \\ \vdots \\ \boldsymbol{Y}_M \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}\boldsymbol{D}_1(\boldsymbol{A}) \\ \vdots \\ \boldsymbol{B}\boldsymbol{D}_M(\boldsymbol{A}) \end{bmatrix} \boldsymbol{S}^{\mathsf{T}} + \begin{bmatrix} \boldsymbol{N}_1 \\ \vdots \\ \boldsymbol{N}_M \end{bmatrix} = (\boldsymbol{A} \odot \boldsymbol{B})\boldsymbol{S}^{\mathsf{T}} + \boldsymbol{N}, \tag{4.5}$$

where $\boldsymbol{D}_m(\boldsymbol{A}) = \mathrm{diag}\,(a_{m1}, \ldots, a_{mK})$ denotes the diagonal matrix with the mth row of $\boldsymbol{A} = [a_{mk}]_{M \times K}$ as its diagonal, and $\boldsymbol{Y}_m = \boldsymbol{B}\boldsymbol{D}_m(\boldsymbol{A})\boldsymbol{S}^{\mathsf{T}} + \boldsymbol{N}_m$. The factorization model (4.5) explicitly expresses an "unfolded" version of the three-way CPD model for the data recorded a vector-sensor array.

4.2.2 Source identifiability with a vector-sensor array

Based on the CPD uniqueness results presented in the Chapter 3, we derived in [Guo11] upper bounds on the number of identifiable source DOA's using a vector-sensor array. As we

are investigating identifiability of the model (4.4), only the noise free system is considered since the uniqueness problem is decoupled from the estimation problem by its nature [Wax89]. Regarding the DOA estimation problem explained in the previous section, the question is: what is the maximum number of the sources K, whose DOA's can be uniquely estimated from (4.5), without *a priori* knowledge on the exact DOA/polarization parameters of the signals? We restrict ourselves to the following two typical scenarios with respect to the level of correlation among the impinging signals. In the first scenario the signals are assumed to be uncorrelated or partially correlated, implying full column rank of the signal matrix $\boldsymbol{S}$. In the second scenario, fully correlated (coherent) signals are considered meaning that $\boldsymbol{S}$ is rank deficient. A natural assumption considered for both scenarios is *A1) the DOA's of the signals are distinct.*

The case of non-coherent signals We examine firstly the important case of *A2) uncorrelated or partially correlated signals.* Thus, the matrix $\boldsymbol{S}$ in (4.5) is full column rank. We proved in [Guo08] that for a linear equally spaced (LES) array with $M \geq K$ vector-sensors the assumptions *A1)* and *A2)* are sufficient for the DOA identifiability of all the K signals.

However, the assumption $M \geq K$ required in [Guo08] is not always necessary. Herein we consider the DOA identifiability problem for a sensor array of $M < K$, arbitrarily positioned, complete vector-sensors. To tackle this, one needs to know the linear dependence pattern of the steering vectors $\boldsymbol{a}(\varphi, \psi)$ of the corresponding scalar sensor array, and that of one vector sensor $\boldsymbol{b}(\varphi, \psi, \alpha, \beta)$.

For $\boldsymbol{a}(\varphi, \psi)$ we follow the common assumption that *A3) the manifold of the corresponding scalar sensor array* $\boldsymbol{a}(\varphi, \psi)$ *is free from rank-M ambiguity,* meaning that $\operatorname{krank}\{\boldsymbol{A}\} = M$. This assumption is justifiable in many cases, *e.g.*, for an LES array $\boldsymbol{a}(\varphi, \psi)$ is a Vandermonde vector. Under the given conditions, we can deduce from Theorem 3.1 an upper bound on the number K of identifiable sources using a vector-sensor array

$$K \leq M + \operatorname{rank}\{\boldsymbol{B}\} - 2. \tag{4.6}$$

For the same scenario, Kruskal's condition yields $K \leq M + \operatorname{krank}\{\boldsymbol{B}\} - 2$, an upper bound which is more restrictive than (4.6).

Regarding the manifold of one vector-sensor $\boldsymbol{b}(\varphi, \psi, \alpha, \beta)$, some interesting results can be found in [Tan96a] and [Ho98]. For instance, the manifold of a complete vector-sensor is free of rank-2 ambiguity [Tan96a], meaning $\operatorname{krank}\{\boldsymbol{B}\} \geq 3$. Moreover, higher ranks of $\boldsymbol{B}$ can be expected, depending on the polarization parameters of the signals, but in general it holds that $\operatorname{rank}\{\boldsymbol{B}\} \geq 4$, unless there exist four signals with identical ellipticity angles [Tan96a]. Particularly, when the ellipticity angles for four out of the K signals are identical and equal to $\pi/4$, *i.e.*, the four are circularly polarized with the same spin, $\operatorname{krank}\{\boldsymbol{B}\} = 3$ but $\operatorname{rank}\{\boldsymbol{B}\}$ is still likely to be 6.

The typical example mentioned above highlights the following two facts. The first one is that the number of signals that can be uniquely localized by a vector-sensor array is

underestimated if Kruskal's condition is used and there is a potential increase by applying (4.6). The second is that the increase, however, cannot be very significant for the vector-sensors since the maximum of rank$\{\boldsymbol{B}\}$ is the number of components of a vector-sensor, which is no greater than 6.

In order to get further insights on the link between Theorem 3.1 and DOA identifiability let us recall a result from [Tan96a], stating that for a vector-sensor array, the ambiguities become inevitable when $K \geq 3M$ signals are present. To *sufficiently* ensure uniqueness, our condition simply excludes all the possible values of K, for which ambiguities may occur, despite how unlikely it does happen. Therefore, a large gap can be observed between the derived bound (4.6), *sufficient* for uniqueness in localizing K signals without knowing exactly the true parameters of the signals, and the bound *necessary* for uniqueness, $K < 3M$. In between the two bounds there is a domain where a general uniqueness condition for our problem is still missing. From a practical point of view, a part of this domain can be explored on a case-by-case basis using the sufficient uniqueness condition of [Jiang04] based on a particular matrix $\boldsymbol{U}$ built from the 2×2 determinants of $\boldsymbol{A}$ and $\boldsymbol{B}$. Figure 4.3 illustrates a comparison between the two sufficient conditions for uniqueness. For the simulations, we used an $(M \times K)$ matrix $\boldsymbol{A}$ with rank$\{\boldsymbol{A}\}$ = krank$\{\boldsymbol{A}\} = M$ and a $(6 \times K)$ matrix $\boldsymbol{B}$. We set krank$\{\boldsymbol{B}\} = 3$ unless $\boldsymbol{B}$ becomes full column rank where krank$\{\boldsymbol{B}\}$ = rank$\{\boldsymbol{B}\} = K$. Meanwhile, observe that $\boldsymbol{B}$ can be a full column rank matrix only if $K \leq 6$. We plotted the condition of Theorem 3.1 and the condition based on the rank of $\boldsymbol{U}$ with respect to krank$\{\boldsymbol{A}\}$(= rank$\{\boldsymbol{A}\}$) and rank$\{\boldsymbol{B}\}$ for four different values of K. The black region in Figure 4.3 corresponds to both uniqueness conditions being satisfied, the dark grey indicates that only the second condition (based on the matrix $\boldsymbol{U}$) is satisfied while the clear shade of gray corresponds to values of (krank$\{\boldsymbol{A}\}$, rank$\{\boldsymbol{B}\}$) for which uniqueness cannot be assessed by either of these two conditions. The white area indicates cases that are impossible for a given K. One can see that the uniqueness test based on the matrix $\boldsymbol{U}$ becomes interesting in practical situations that are not covered by Theorem 3.1, especially when the number of components (*i.e.*, K) is high. Although the general value of this result cannot be assessed, simulations conducted with various realizations of $\boldsymbol{A}$ and $\boldsymbol{B}$ (*e.g.,* the entries of $\boldsymbol{A}$ and $\boldsymbol{B}$ randomly drawn from a continuous distribution), yielded the same result each time.

The case of coherent signals Usually, coherence among the signals occurs as a result of multipath propagation environment and results in proportional columns in the matrix $\boldsymbol{S}$ (Eq. (4.5)). This implies a maximum rank of $\boldsymbol{S}$ equal to $K - 1$ and krank$\{\boldsymbol{S}\} = 1$. In spite of this, we can assume, without loss of generality, that both assumptions *A1)* and *A3)* still hold. In the presence of correlated sources, the eigenstructure-based DOA estimation techniques, such as MUSIC [Bienvenu79, Schmidt81], encounter difficulties. To tackle this problem, "signal decorrelation" techniques, such as the polarization smoothing algorithm (PSA) [Rahamim04], can be used for vector-sensor arrays. The PSA can be regarded as forming a new signal matrix $\breve{\boldsymbol{S}} = \boldsymbol{B} \odot \boldsymbol{S}$ such that $\breve{\boldsymbol{S}}$ is full column rank. Following Lemma

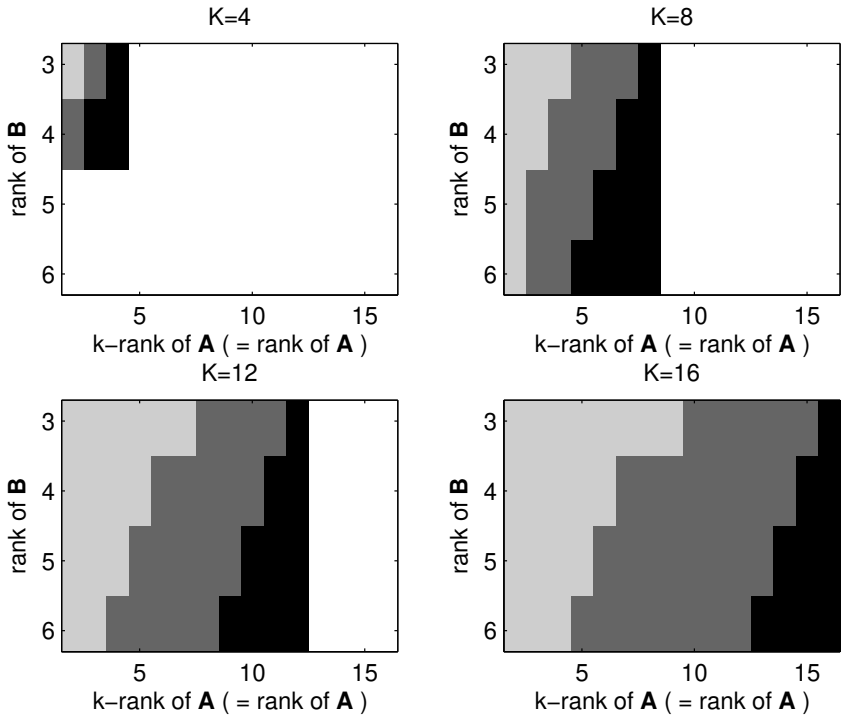


Figure 4.3: A comparison between the sufficient conditions for DOA identifiability given by Theorem 3.1 and the one based on matrix $\boldsymbol{U}$ [Jiang04]. The black region corresponds to both conditions being satisfied, the dark gray indicates that only the condition of [Jiang04] is verified and the clear shade of gray indicates that none of the two uniqueness conditions works.

3.1, $\check{\boldsymbol{S}}$ is ensured to be full column rank if

$$K \leq \text{krank}\{\boldsymbol{B}\} + \text{rank}\{\boldsymbol{S}\} - 1. \tag{4.7}$$

The newly obtained data structure can thus be expressed as $\check{\boldsymbol{Y}} = \boldsymbol{A}\check{\boldsymbol{S}}^{\mathsf{T}}$. To implement MUSIC-like algorithms on the new data $\check{\boldsymbol{Y}}$, a fundamental requirement is $M \geq K + 1$, *i.e.*, more vector-sensors than sources are needed [Rahamim04]. Combining this condition with (4.7), one obtains

$$K \leq \min\{M, \text{krank}\{\boldsymbol{B}\} + \text{rank}\{\boldsymbol{S}\}\} - 1 \tag{4.8}$$

We showed [Guo11] that, under the given assumptions, condition (4.8) guarantees that the K coherent signals can be uniquely localized. Particularly, in the special case where $\text{rank}\{\boldsymbol{S}\} = 1$, *i.e.*, signals are all coherent, (4.8) reduces to

$$K \leq \min\{\text{krank}\{\boldsymbol{B}\}, M - 1\}, \tag{4.9}$$

which recalls the result of [Rahamim04] in the case of an LES array (with *complete* vector-sensors): $K \leq \min(6, M - 1)$. Nevertheless, the result of [Rahamim04] is *not sufficient* to ensure uniqueness and it can be regarded as a derivation of (4.9) under the optimistic assumption that $\text{krank}\{\boldsymbol{B}\} = 6$. A counterexample to this assumption is that for a complete vector-sensor (6 components), one can have $\text{krank}\{\boldsymbol{B}\} = 3$ if four of the sources are circularly polarized with the same spin direction [Tan96a]. In this case, using the PSA only may not be sufficient to ensure system identifiability.

A study of the influence of polarization diversity on the performance of CPD-based DOA estimation can be found in [Guo08]. We showed that by optimally designing the source polarizations, the performance of the CPD-based method can approach that of non-blind separation methods.

4.2.3 Polarized array processing based on a quadrilinear representation of the data covariance

Th joint estimation of all the three parameters of the sources (DOA, polarization, and temporal sequence) is time-consuming, and it does not always have a practical interest, especially in array processing applications. We proposed in [Miron08, Miron10] an algorithm for DOA estimation of polarized sources, allowing the estimation of only two source parameters (DOA and polarization), and thus presenting a smaller computational complexity than its trilinear version [Guo08]. It is based on a quadrilinear (fourth-order tensor) representation of the polarized data covariance. The parameters are then obtained by the CPD of the covariance tensor of the polarized data, using a quadrilinear alternating least squares (QALS) approach. A significant advantage of the proposed algorithm lies in the fact that the methods based on statistical properties of the signals proved to outperform deterministic techniques

[Swindlehurst97], provided that the number of samples is sufficiently high.

The quadrilinear model of the data covariance Consider a uniform array of M identical sensors spaced by Δx along the x-axis, collecting narrowband signals emitted from K (known *a priori*) spatially distinct far-field sources. As shown at the beginning of this section, a snapshot of the array can be expressed by equation (4.4).

Let

$$\boldsymbol{S}_d = \begin{bmatrix} s_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & s_K \end{bmatrix}, \tag{4.10}$$

be a $K \times K$ diagonal containing the K source signals at some fixed instant. With these notations, a snapshot of the output of the array can be organized as a $M \times 6$ matrix

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{S}_d\boldsymbol{B}^{\mathsf{T}} + \boldsymbol{N}, \tag{4.11}$$

with $\boldsymbol{N}$ a $M \times 6$ matrix expressing the noise contribution on the antenna.

The following assumptions are made

(B1) the sources are zero-mean, stationary, mutually uncorrelated, ergodic processes,

(B2) the noise is i.i.d. centered, complex Gaussian process of variance σ^2, non-polarized and spatially white,

(B3) the sources have distinct DOAs.

We define the 4-way covariance of the received data as the $M \times 6 \times M \times 6$ tensor:

$$\boldsymbol{\mathcal{C}}_{YY} = \mathrm{E}\{\boldsymbol{Y} \circ \boldsymbol{Y}^*\}, \tag{4.12}$$

where E{.} denotes the mathematical expectation operation. We define also the source covariance as the $K \times K \times K \times K$ fourth-order tensor

$$\boldsymbol{\mathcal{C}}_{SS} = \mathrm{E}\{\boldsymbol{S}_d \circ \boldsymbol{S}_d^*\}. \tag{4.13}$$

From (4.11) and (4.13) and using assumptions (B1) and (B2) the covariance tensor of the received data takes the following form:

$$\boldsymbol{\mathcal{C}}_{YY} = \boldsymbol{\mathcal{C}}_{SS} \underset{1}{\bullet} \boldsymbol{A} \underset{2}{\bullet} \boldsymbol{B} \underset{3}{\bullet} \boldsymbol{A}^* \underset{4}{\bullet} \boldsymbol{B}^* + \boldsymbol{\mathcal{N}}, \tag{4.14}$$

where $\boldsymbol{\mathcal{N}}$ is a $M \times 6 \times M \times 6$ tensor containing the noise power on the sensors. Assumption (B1) implies that $\boldsymbol{\mathcal{C}}_{SS}$ is a hyperdiagonal tensor (the only non-null entries are those having all four indices identical), meaning that $\boldsymbol{\mathcal{C}}_{YY}$ presents a quadrilinear CPD structure [Harshman70]. The inverse problem for the direct model expressed by (4.14) is the estimation of matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ starting from the 4-way covariance tensor $\boldsymbol{\mathcal{C}}_{YY}$.

Identifiability of the quadrilinear model The polarized mixture model (4.14) is said to be identifiable if $\boldsymbol{A}$ and $\boldsymbol{B}$ can be uniquely determined (up to permutation and scaling indeterminacies) from $\boldsymbol{\mathcal{C}}_{YY}$. We proved the following result:

Proposition 4.1 ([Miron10]). *Under the assumptions (B1)-(B3), the DOA's of K uncorrelated sources can be uniquely determined using an M-element vector sensor array if $M \geq K-1$, regardless of the polarization states of the incident signals.*

This sufficient condition also sets an upper bound on the minimum number of sensors needed to ensure the identifiability of the polarized mixture model. However, the condition $M \geq K-1$ is not necessary when considering the polarization states, that is, a lower number of sensors can be used to identify the mixture model, provided that the polarizations of the sources are different. Also the symmetry properties of $\boldsymbol{\mathcal{C}}_{YY}$ are not considered and we believe that they can be used to obtain milder sufficient conditions for ensuring the identifiability.

Source parameters estimation Supposing that N snapshots of the array are recorded and using (B1) an estimate of the polarized data covariance (4.12) can be obtained as the temporal sample mean

$$\hat{\boldsymbol{\mathcal{C}}}_{YY} = \frac{1}{N}\sum_{n=1}^{N} \boldsymbol{Y}(n) \circ \boldsymbol{Y}^{*}(n). \tag{4.15}$$

For obvious matrix conditioning reasons, the number of snapshots should be greater or equal to the number of sensors, *i.e.* $N \geq K$.

The algorithm that we proposed includes three sequential steps, during which the DOA information is extracted and then refined to yield the final DOA's estimates.

Step 1: This first step of the algorithm is the estimation of the loading matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ from $\hat{\boldsymbol{\mathcal{C}}}_{YY}$. This estimation procedure can be accomplished via the *Quadrilinear Alternative Least Squares (QALS)* algorithm [Bro98a], as shown next.

Denote by $\hat{\boldsymbol{C}}_{pq} = \hat{\boldsymbol{\mathcal{C}}}_{YY}(:,p,:,q)$ the (p,q)th matrix slice $(M \times M)$ of the covariance tensor $\hat{\boldsymbol{\mathcal{C}}}_{YY}$. Also note $\mathrm{D}_p(\,\cdot\,)$ the operator that builds a diagonal matrix from the pth row of another and $\boldsymbol{\Delta} = \mathrm{diag}\,(E\{\|s_1\|^2\}, \ldots, E\{\|s_K\|^2\})$, the diagonal matrix containing the powers of the sources. The matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ can then be determined by minimizing the Least Squares (LS) criterion:

$$\mathcal{J}(\boldsymbol{A},\boldsymbol{B}) = \sum_{p,q} \left\| \hat{\boldsymbol{C}}_{pq} - \boldsymbol{A}\boldsymbol{\Delta}\mathrm{D}_p(\boldsymbol{B})\mathrm{D}_q(\boldsymbol{B}^{*})\boldsymbol{A}^{\mathsf{H}} \right\|_{\mathsf{F}}^{2}. \tag{4.16}$$

The optimization process in (4.16) can be implemented using QALS algorithm, briefly summarized in **Algorithm 1**.

Once the estimates $\hat{\boldsymbol{A}}$, $\hat{\boldsymbol{B}}$ are obtained, the following post-processing is needed for the refined DOA estimation.

Step 2: The second step of our approach extracts separately the DOA information contained by the columns of $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{B}}$. First the estimated matrix $\hat{\boldsymbol{B}}$ is exploited via the physical

Algorithm 1: QALS algorithm for four-way symmetric tensors

1: INPUT: the estimated data covariance $\hat{\boldsymbol{\mathcal{C}}}_{YY}$ and the number of the sources K
2: Initialize the loading matrices $\boldsymbol{A}$, $\boldsymbol{B}$ randomly, or using ESPRIT [Zoltowski00a] for a faster convergence
3: Set $\boldsymbol{C} = \boldsymbol{A}^*$ and $\boldsymbol{D} = \boldsymbol{B}^*$.
4: **repeat**
5: $\quad \boldsymbol{A} = \boldsymbol{Y}_{(1)}[(\boldsymbol{B} \odot \boldsymbol{C} \odot \boldsymbol{D})^\dagger]^\mathsf{T}$
6: $\quad \boldsymbol{B} = \boldsymbol{Y}_{(2)}[(\boldsymbol{C} \odot \boldsymbol{D} \odot \boldsymbol{A})^\dagger]^\mathsf{T}$
7: $\quad \boldsymbol{C} = \boldsymbol{Y}_{(3)}[(\boldsymbol{D} \odot \boldsymbol{A} \odot \boldsymbol{B})^\dagger]^\mathsf{T}$
8: $\quad \boldsymbol{D} = \boldsymbol{Y}_{(4)}[(\boldsymbol{A} \odot \boldsymbol{B} \odot \boldsymbol{C})^\dagger]^\mathsf{T}$,
where $(\ \cdot\ )^\dagger$ denotes Moore-Penrose pseudoinverse of a matrix
9: $\quad$ Update $\boldsymbol{C}$, $\boldsymbol{D}$ by $\boldsymbol{C} := (\boldsymbol{A}^* + \boldsymbol{C})/2$ and $\boldsymbol{D} := (\boldsymbol{B}^* + \boldsymbol{D})/2$
10: **until** convergence
11: OUTPUT: estimates of $\boldsymbol{A}$ and $\boldsymbol{B}$.

relationships between the electric and magnetic field given by the Poynting theorem : Recall the Poynting theorem, which reveals the mutual orthogonality nature among the three physical quantities related to the kth source: the electric field $\boldsymbol{e}_k$, the magnetic field $\boldsymbol{h}_k$, and the kth source's direction of propagation, *i.e.*, the normalized Poynting vector $\boldsymbol{p}_k$:

$$\boldsymbol{p}_k = \begin{bmatrix} \cos\varphi_k \cos\psi_k \\ \sin\varphi_k \cos\psi_k \\ \sin\psi_k \end{bmatrix} = \Re\left(\frac{\boldsymbol{e}_k \times \boldsymbol{h}_k^*}{\|\boldsymbol{e}_k\| \cdot \|\boldsymbol{h}_k\|} \right). \tag{4.17}$$

Equation (4.17) gives the cross-product DOA estimator, as suggested in [Nehorai94]. An estimate of the Poynting vector for the kth source $\hat{\boldsymbol{p}}_k$ is thus obtained, using the previously estimated $\hat{\boldsymbol{e}}_k$ and $\hat{\boldsymbol{b}}_k$.

Secondly matrix $\hat{\boldsymbol{A}}$ is used to extract the DOA information embedded in the Vandermonde structure of its columns $\hat{\boldsymbol{a}}_k$, via an optimisation procedure detailed in [Miron10].

Step 3: In the third step, the two DOA estimates, obtained at **Step 2**, are combined in order to get a refined estimation of the DOA parameters φ and ψ. This step can be formulated can be formulated as the following non-linear optimization problem

$$\min_{\psi,\varphi} \left\| \begin{bmatrix} \cos\varphi \cos\psi \\ \sin\varphi \cos\psi \\ \sin\psi \end{bmatrix} - \hat{\boldsymbol{p}} \right\| \quad \text{subject to } \cos\varphi \cos\psi = \varrho. \tag{4.18}$$

A closed form solution to (4.18) can be found by transforming it into an alternate problem of 3-D geometry, *i.e.* finding the point on the vertically posed circle $\cos\varphi\cos\psi = \varrho$ that minimizes its Euclidean distance to the point $\hat{\boldsymbol{p}}$, as shown in Figure 4.4.

To solve this problem, we make the orthogonal projection of $\hat{\boldsymbol{p}}$ onto the plane $x = \varrho$ in the 3-D space, then join the perpendicular foot with the center of the circle by a piece

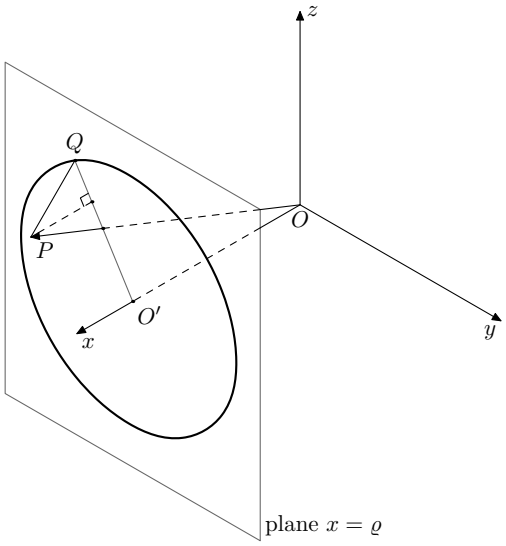


Figure 4.4: Illustration of the geometrical solution to the optimization problem (4.18). The vector $\vec{OP}$ represents the coarse estimate of Poynting vector $\hat{\boldsymbol{p}}$. It is projected orthogonally onto the $x = \varrho$ plane, forming a shadow cast $\overline{O'Q}$, where O' is the center of the circle $\odot O'$ on the plane given in the polar coordinates as $\cos\varphi\cos\psi = \varrho$. The refined estimate, obtained this way, lies on $\overline{O'Q}$. As it is also constrained on the circle, it can be sought as their intersection point Q.

of line segment. This line segment collides with the circumference of the circle, yielding an intersection point, that is the minimizer of the problem.

Let $\hat{\boldsymbol{p}} \triangleq [\hat{p}_1\, \hat{p}_2\, \hat{p}_3]^{\mathsf{T}}$ and define $\kappa \triangleq \hat{p}_3/\hat{p}_2$, then the intersection point is given by

$$\left[\begin{array}{ccc} \varrho & \pm\sqrt{\frac{1-\varrho^2}{1+\kappa^2}} & \pm|\kappa|\sqrt{\frac{1-\varrho^2}{1+\kappa^2}} \end{array}\right]^{\mathsf{T}} \tag{4.19}$$

where the signs are taken the same as their corresponding entries of vector $\hat{\boldsymbol{p}}$. Thus, the azimuth and elevation angles estimates are given by

$$\hat{\varphi} = \begin{cases} \arctan \frac{1}{|\varrho|}\sqrt{\frac{1-\varrho^2}{1+\kappa^2}}, & \text{if } \varrho \geq 0 \\ \pi - \arctan \frac{1}{|\varrho|}\sqrt{\frac{1-\varrho^2}{1+\kappa^2}}, & \text{if } \varrho < 0 \end{cases} \tag{4.20a}$$

$$\hat{\psi} = \arcsin\sqrt{\varrho^2 + \frac{1-\varrho^2}{1+\kappa^2}}, \tag{4.20b}$$

which completes the DOA estimation procedure. The polarization parameters can be obtained in a similar way from $\hat{\boldsymbol{B}}$.

It is noteworthy that this algorithm is not necessarily limited to uniform linear arrays. It can be applied to arrays of arbitrary configuration, with minimal modifications.

Discussion

- It was shown [Liu01] that higher dimensionality benefits in terms of CRB for a given

data set. In other words, the use of a quadrilinear ALS on the covariance tensor is better sounded that performing a trilinear ALS on the unfolded covariance tensor.

- Using estimator variance considerations, we showed [Miron10] that QALS estimates improves continuously with the sample size while this is not the case for trilinear ALS (TALS) [Guo08].

- Compared to MUSIC-like algorithms, which are also based on the estimation of the data covariance, the main advantage of QALS is the identifiability of the model. While MUSIC generally requires an exhaustive grid search for the estimation of the source parameters, the quadrilinear method yields directly the steering and the polarization vectors for each source.

Numerical simulation Figure 4.5 compares the accuracy of the QALS approach to three state of the art DOA estimation algorithms for polarized sources: polarized MUSIC, presented under different versions in [Ferrara83, Weiss93, Miron05, Gong09], TALS [Guo08], polarized ESPRIT [Zoltowski00a, Zoltowski00b] and to the Cramér-Rao bound for vector sensor arrays proposed in [Nehorai94]. We considered $K = 4$ uncorrelated sources, $M = 7$ sensors, $N = 100$ temporal snapshots and 500 runs per point. For simplicity, we assumed that all the elevation angles are zero, and some typical values are chosen for the azimuth angles, respectively: $\varphi_1 = 10°$, $\varphi_2 = 20°$, $\varphi_1 = 30°$, $\varphi_1 = 40°$. The polarizations parameters are $\alpha_2 = -45°$, $\beta_2 = -15°$ for the second source and for the others, the sources have equal orientation angles and ellipticity angles, 45° and 15° respectively. We can observe the advantage of the multilinear approaches in tackling DOA problem at low SNR can be observed. The quadrilinear approach seems to perform better than TALS as the SNR increases. The MUSIC algorithm is more sensitive to the noise than all the others, yet it reaches the CRB as the SNR is high enough. The estimate obtained by ESPRIT is mildly biased.

Extensive numerical simulations, with respect other parameters, can be found in [Miron10].

4.3 Tensors for multi-scale invariance sensor array processing

In the previous section we used the polarization diversity to build up tensor models for sensor arrays. Other types of diversities can also be used to this end. In [Miron13, Miron15b], we proposed a sensor array configuration that exploits the spatial scale invariance. This approach generalizes the results given in [Sidiropoulos00b] to an array that presents an arbitrary number of spatial invariances. This work is the result of the collaboration with colleagues from the Hong Kong Polytechnic University.

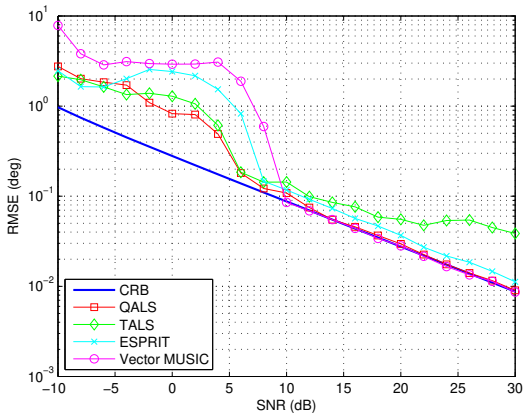


Figure 4.5: RMSE of azimuth angle estimation versus SNR for the second source in the presence of four uncorrelated sources

4.3.1 The geometric configuration of a multiple scale-invariant sensor-array

Consider a subarray composed of L_1 isotropic identical sensors indexed by $l_1 = 1, \ldots, L_1$. Consider then, L_2 identical replicas of this subarray, spatially translated to arbitrary, possibly unknown locations[1]. The L_2 distinct copies of the original subarray, indexed by $l_2 = 1, \ldots, L_2$, can now be seen as *subarrays* that together constitute a larger (*higher-level*) array. This proposed array structure can be further generalized by considering an additional level, composed of L_3 translated replicas of the previous sensor subarrays, indexed by $l_3 = 1, \ldots, L_3$. Let us generalize this scheme to a total of N such hierarchical levels, with the "highest" level consisting of L_N subarrays indexed by $l_N = 1, \ldots, L_N$. It is worth noting that two different subarrays at a given level n need not be disjoint, *i.e.* they may have in common subarrays/sensors of the previous level $(n-1)$. However, if all subarrays at each level are disjoint, then the entire array will contain a total number of $L = L_1 L_2 \ldots L_N$ identical sensors. Fig. 4.6 illustrates a three-level array of co-planar sensors.

Consider also a Cartesian coordinate system $OXYZ$ to describe the considered array. An impinging source is characterized in this coordinate system by its direction-cosines, u, v, w :

$$\boldsymbol{p} = \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} \sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ \cos\theta \end{bmatrix}, \tag{4.21}$$

where $\theta \in [0, \pi]$ denotes the source's incident elevation angle measured from the positive Z-axis, and $\varphi \in [0, 2\pi]$ symbolizes the azimuth angle measured from the positive X-axis.

[1]Information on the exact positions of the sensors is not required for the validity of the data model. However, to estimate the sources' DOAs from the measured data, the knowledge of the array manifold is necessary.

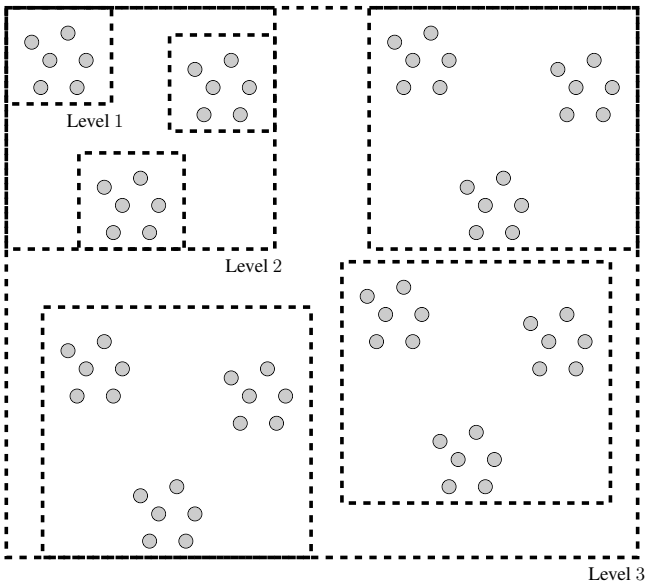


Figure 4.6: A multi-scale planar array with three hierarchical levels.

4.3.2 The CPD data model

Consider first a narrow-band plane wave impinging on the array described in section 4.3.1. Let us represent by $a_{l_1 l_2 \ldots l_N}$ its phase factor at the sensor indexed by $l_1, l_2, \ldots, l_N$ at the N various levels of the array. Define $\boldsymbol{k} = [u\ v\ w]^{\mathsf{T}}$ and $\boldsymbol{d}_{l_n}^{(n)} = [x_{l_n}^{(n)}\ y_{l_n}^{(n)}\ z_{l_n}^{(n)}]^{\mathsf{T}}$, with $n = 1, \ldots, n$. With the notation introduced above, the spatial phase factor is given by:

$$a_{l_1 l_2 \ldots l_N}(\boldsymbol{k}) = \exp\Big\{j\frac{2\pi}{\lambda}\ \sum_{n=1}^{N} \boldsymbol{k}^{\mathsf{T}} \boldsymbol{d}_{l_n}^{(n)}\Big\} = \prod_{n=1}^{N} \exp\Big\{j\frac{2\pi}{\lambda}\ \boldsymbol{k}^{\mathsf{T}} \boldsymbol{d}_{l_n}^{(n)}\Big\}. \tag{4.22}$$

Thus, the array manifold for the entire sensor-array is

$$\boldsymbol{a}(\boldsymbol{k}) = \boldsymbol{a}_1(\boldsymbol{k}) \otimes \ldots \otimes \boldsymbol{a}_N(\boldsymbol{k}), \tag{4.23}$$

with

$$\boldsymbol{a}_n(\boldsymbol{k}) = \begin{bmatrix} e^{j(2\pi/\lambda)\boldsymbol{k}^{\mathsf{T}}\boldsymbol{d}_1^{(n)}} \\ \vdots \\ e^{j(2\pi/\lambda)\boldsymbol{k}^{\mathsf{T}}\boldsymbol{d}_{Ln}^{(n)}} \end{bmatrix} \tag{4.24}$$

an $L_n \times 1$ vector, $\forall n = 1, \ldots, N$ and "$\otimes$" represents the Kronecker product of two matrices.

Next, consider P narrow-band plane-waves, having traveled through a nonconductive homogeneous isotropic medium, impinging upon the array from directions $\boldsymbol{k}_p = [u_p\ v_p\ w_p]^{\mathsf{T}}$, with $p = 1, \ldots, P$. Denote by $s_p(t)$ the time signal emitted by the pth narrow-band source[2].

[2]The incident signals are narrow-band in that their bandwidths are very small compared with the inverse of the wavefronts' transit time across the array.

Then, the output at time t of the entire sensor-array can be expressed as an $L \times 1$ vector,

$$\boldsymbol{z}(t) = \sum_{p=1}^{P} \Big(\boldsymbol{a}_1(\boldsymbol{k}_p) \otimes \ldots \otimes \boldsymbol{a}_N(\boldsymbol{k}_p) \Big) s_p(t) + \boldsymbol{n}(t), \tag{4.25}$$

where $\boldsymbol{n}(t)$ is a complex-valued zero-mean additive white noise.

Let us assume that we have at our disposal K snapshots at time instants, $t_1, t_2, \ldots, t_K$. Define the following $L_n \times P$ matrices:

$$\boldsymbol{A}_1 = \Big[\boldsymbol{a}_1(\boldsymbol{k}_1), \ldots, \boldsymbol{a}_1(\boldsymbol{k}_P) \Big] \tag{4.26}$$

$$\vdots$$

$$\boldsymbol{A}_N = \Big[\boldsymbol{a}_N(\boldsymbol{k}_1), \ldots, \boldsymbol{a}_N(\boldsymbol{k}_P) \Big] , \tag{4.27}$$

and the $K \times P$ matrix:

$$\boldsymbol{S} = \begin{bmatrix} s_1(t_1) & s_2(t_1) & \ldots & s_P(t_1) \\ s_1(t_2) & s_2(t_2) & \ldots & s_P(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ s_1(t_K) & s_2(t_K) & \ldots & s_P(t_K) \end{bmatrix} = \Big[\boldsymbol{s}_1, \ \boldsymbol{s}_2, \ldots, \boldsymbol{s}_P \Big] \ . \tag{4.28}$$

The collection of K snapshots of the array can then be organized into an $L \times K$ data matrix as

$$\boldsymbol{Z} = [\boldsymbol{z}(t_1), \ldots, \boldsymbol{z}(t_K)] = \Big(\boldsymbol{A}_1 \odot \cdots \odot \boldsymbol{A}_N \Big) \boldsymbol{S}^{\mathsf{T}} + \boldsymbol{N}, \tag{4.29}$$

where "$\odot$" denotes the Khatri-Rao product of two matrices, and $\boldsymbol{N}$ is a $(L \times K)$ complex-valued matrix, modeling the noise on the entire array for all K temporal snapshots. Equation (4.29) reveals a $(N+1)$-dimensional CPD structure.

In the case where only one time-sample is available, *i.e.* matrix $\boldsymbol{S}$ is a $1 \times P$ vector, the data model given by (4.29) becomes

$$\boldsymbol{z} = \Big(\boldsymbol{A}_1 \odot \cdots \odot \boldsymbol{A}_N \Big) \boldsymbol{s} + \boldsymbol{n}, \tag{4.30}$$

with $\boldsymbol{z} = \boldsymbol{z}(t_1)$, $\boldsymbol{s} = \boldsymbol{s}(t_1) = \Big(\boldsymbol{S}(1,:) \Big)^{\mathsf{T}}$ and $\boldsymbol{n} = \boldsymbol{N}(:,1)$. In the definitions above, we used the Matlab notations for columns and rows selection operators. Equation (4.30) is a vectorized representation of a N-dimensional CP data model. It is worth noting that if only one snapshot of the array is available, the $N+1$ CP model degenerates into an N-dimensional one.

A discussion on the identifiability of these CPD models, for different scenarios can be found in [Miron15b].

4.3.3 DOA estimation

We proposed a DOA estimation procedure that can be split into two stages. The first stage estimates the N steering vectors $\boldsymbol{a}_n(\boldsymbol{k}_p)$ $(n = 1, \ldots, N)$ for each of the P sources $(p = 1, \ldots, P)$, by exploiting the CPD structure (4.29) of the collected data. In this first stage, an Alternating Least Squares (ALS) procedure can be used to fit the CPD model. This ALS procedure recursively estimates one of the $N+1$ matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_N, \boldsymbol{S}$, by fixing the other N of them.

The second stage estimates the source's direction-cosines $\boldsymbol{k}_p$, $p = 1, \ldots, P$ from the steering vectors obtained at the previous stage. To this end, we proposed to formulate the DOA estimation as an optimization problem, and to adopt a new sequential procedure that exploits all the available information from the source's steering vectors encompassing all scale levels.

Define the following cost functions:

$$\mathcal{J}_n(\boldsymbol{k}_p) = \|\hat{\boldsymbol{a}}_n^p - \boldsymbol{a}_n(\boldsymbol{k}_p)\|^2, \text{ with } n = 1, \ldots, N, \tag{4.31}$$

where $\hat{\boldsymbol{a}}_n^p$ denotes the estimated steering vector at the nth level for the pth source. Estimating the DOAs for the pth source comes down to minimizing the following criterion:

$$\mathcal{I}_N(\boldsymbol{k}_p) = \sum_{n=1}^{N} \mathcal{J}_n(\boldsymbol{k}_p). \tag{4.32}$$

This function is non-convex and highly non-linear with respect to the direction-cosines; hence, a direct local optimization procedure would fail in most cases. We propose to adopt a new sequential strategy to minimize $\mathcal{I}_N(\boldsymbol{k}_p)$, progressing from one level to the next higher level, using an iterative refinement of the direction-cosine estimates within each level. The method is based on the fact that, when noise-free, the N cost-functions in (4.31) have the same global minimum.

The proposed algorithm can be summarized as follows:

The Two-Stage Estimation Algorithm

First Stage: Estimate $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_N$ by CP decomposition of the data $\boldsymbol{\mathcal{Z}}$ or $\boldsymbol{z}$ (see eq. (4.29) or (4.30)).

Second Stage:

- For $p = 1, \ldots, P$ and

  for $n = 1, \ldots, N$, compute

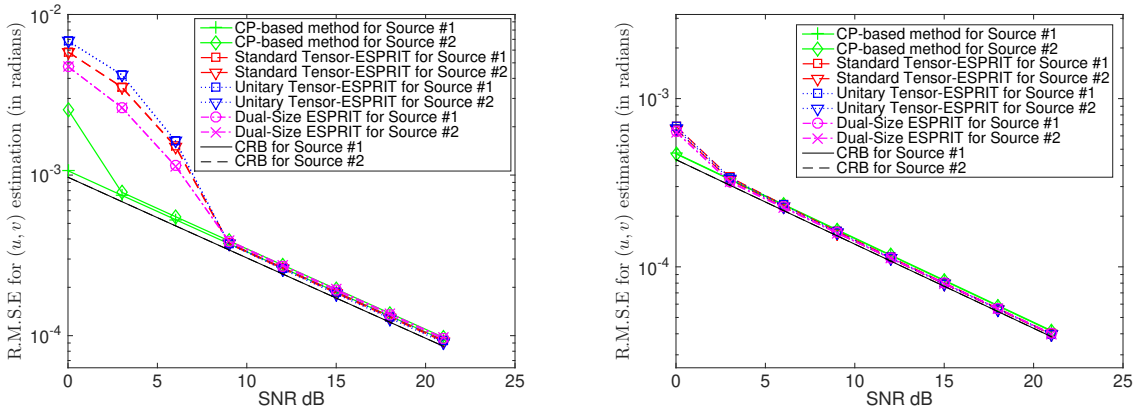
$$\boldsymbol{k}_{p,n}^* = \underset{\boldsymbol{k}_p}{\operatorname{argmin}}\, \mathcal{I}_n(\boldsymbol{k}_p) = \underset{\boldsymbol{k}_p}{\operatorname{argmin}} \sum_{i=1}^{n} \mathcal{J}_i(\boldsymbol{k}_p). \tag{4.33}$$

- *Output:* The estimated parameters for the P sources: $\hat{\boldsymbol{k}}_p = (\hat{u}_p, \hat{v}_p, \hat{w}_p) = \boldsymbol{k}^*_{p,N}$ with $p = 1, \ldots, P$.

A discussion on the convergence of the proposed algorithm can be found in [Miron15b].

4.3.4 Numerical simulations

We compare next, our approach with the two ESPRIT-based methods developed in [Wong98] and [Zoltowski00a] and to the R-D Standard Tensor-ESPRIT and the R-D Unitary Tensor-ESPRIT developed in [Haardt08]. We evaluate the performance of the four considered algorithms for different SNRs, in a statistically independent sources scenario. Figures 4.7(a) and 4.7(b) plot the "composite root-mean-square-error" (CRMSE) of the sources' Cartesian direction-cosine estimates versus SNR, for 10 and 50 snapshots, respectively. The array con-



(a) $K = 10$ (b) $K = 50$

Figure 4.7: Statistically independent sources: CRMSE versus the signal-to-noise power ratio (SNR).

figuration used for the simulations is given on Figure 4.8. At low SNRs and for a small number of snapshots, our method performs better than the ESPRIT-based algorithms. As expected, when the SNR and/or the number of snapshots increases (Figure 4.7(b)), all the methods present similar performance.

Further results, for various scenarios, and a detailed discussion are available in [Miron15b].

4.4 Conclusions

DOA estimation models and methods has been a constant research concern during my entire career. My main contributions, presented in this chapter, focus on the use of additional diversities to improve source localization performance.

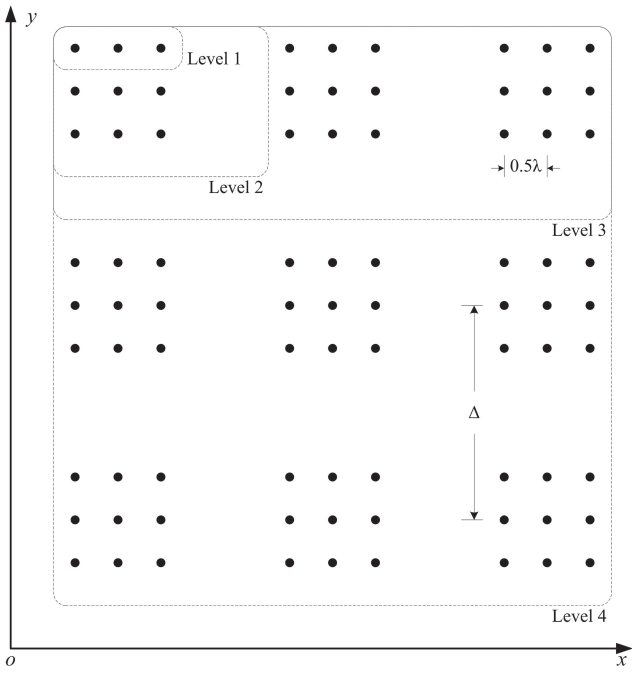


Figure 4.8: The four level array configuration used for the simulations.

We showed, for the first time in the array processing literature, that the data recorded on an electromagnetic vector-sensor array follows a CPD model. Based this result and on the CPD uniqueness results presented in Chapter 3, we derived bounds on the the maximum number of sources that can be detected by such an array. Moreover, we introduced a fourth-order tensor representation of the data covariance for a vector-sensor array and showed that this representation improves the accuracy of the DOA estimation algorithms.

By exploiting a particular type of spatial diversity, we proposed a new kind of sensor array for DOA estimation presenting multiple scales of invariance; a source localization algorithm for this special array was also developed and we showed that it outperforms state-of-the-art methods, in difficult scenarios.

Chapter 5

Tensor models and methods in spectroscopy

This chapter gathers the main results of my work on multi-way arrays decompositions and models for spectroscopy data processing. Two types of fluorescence are concerned, namely the cell fluorescence and the polarized Raman spectroscopy for cristal characterization.

5.1 Tensors for biosensors and cells fluorescence signal analysis

This section is dedicated to the use of tensor methods for cell fluorescence analysis. The biosensors signal analysis was developed in the ANR HÆSPRI (presented in Section 1.7); the results are extracted from the Ph.D. work of Fabrice Caland that I co-supervised with David Brie. The results on the flow cytometry are the fruit of the collaboration with biologist colleagues from CRAN.

5.1.1 Analysis of biosensors' responses to pollutants by tensor decomposition

A biosensor is a biological system of detection combined with a system of transduction [Tecon08, S. Ramanathan97]. As compared to other classical sensors, biosensors offer the advantage of being possibly produced at very low cost, being more sensitive to the presence of low concentration of chemical elements and providing a measure of the bioavailability on the studied element (the proportion of the element which actually interacts with the biological system).

The design of a bacterial biosensor consists in adding a reporter gene to the bacterium DNA so that it synthesizes fluorescent proteins when exposed to a stressing element [Leveau02]. Figure 5.1 shows a schematic diagram of a bacterial biosensor and its components interacting

to provide information on toxicity level. Biosensors have been developed to react to a number of environmental pollutants.

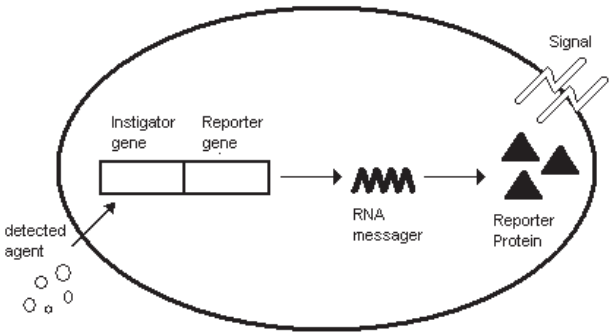


Figure 5.1: Schematic diagram of bacterial biosensor

One of the greatest limitations in the use of biosensors is the ambiguous nature of their response. Highly specific promoters responding to a single chemical element are rare. For example, the existence of unambiguous responses for the sequestration or expulsion of heavy metals is highly controversial. A multivariate approach is therefore highly useful to study this problem. The modularity of the biosensors makes it possible to multiply the sources of fluorescence by limiting the genetic modifications carried out by the markings of the studied promoter genes. The different sources of fluorescence will mix with each other and with the natural fluorescence of the medium to give the overall fluorescence. The use of source separation methods allows to extract the signals of interest and to reduce interference from other sources of fluorescence.

The approach that we proposed in the HÆSPRI project consists in fully exploiting the diversity of biosensor responses from which the relevant information can be recovered through signal processing methods. Supposing that the biosensor response to a number of $N-1$ factors have to be evaluated, the corresponding spectra can be arranged as a N-way array, as explained in the next section. We proposed in [Caland11, Caland12, Parrello15] to analyse the sensor responses to different pollutants by means of CPD.

5.1.1.1 The CPD model of biosensors' fluorescence data

A biosensor is producing a fluorescence signal $s(\lambda)$ (with λ, the wavelength) when exposed to a given pollutant. For the sake of simplicity let us consider the case of two pollutants only. The method can be extended straightforwardly to a larger number of pollutants. We denote by $a(x)$ and $b(y)$ the responses of a biosensor to two different pollutants, where x and y are representing the pollutant concentrations. Suppose now that a solution contains these two pollutants with respective concentrations x_0 and y_0. Following [Tecon06], the fluorescence is known to be proportional to the amount of bacteria denoted $c(z_0)$, where the variable z_0 is introduced for notational homogeneity and whose meaning will be explained in the sequel. Thus,

the fluorescence signal emitted by the biosensor is proportional to $(a(x_0) + b(y_0))c(z_0)s(\lambda)$. Supposing that there are F biosensor types marked with different fluorophores and responding differently to these two pollutants, the measured fluorescence signal can be expressed as:

$$\boldsymbol{\mathcal{D}}(x_0, y_0, z_0, \lambda) = \sum_{f=1}^{F} (a_f(x_0) + b_f(y_0))c_f(z_0)s_f(\lambda). \tag{5.1}$$

It should be noted that, within a certain range, the model is additive with respect to the response of the biosensors to the two pollutants. Building a multidimensional array by performing a metered addition of the two pollutants will not produce an array having a multilinear CPD structure. To tackle this problem we propose to perform successive metered additions of a single pollutant combined with a controlled concentration variation of the different types of biosensors. By performing a metered addition of the first pollutant, we obtain a three-way data array model:

$$\boldsymbol{\mathcal{D}}_1(x, z, \lambda) = \sum_{f=1}^{F} \alpha_f(x)c_f(z)s_f(\lambda), \tag{5.2}$$

where $\alpha_f(x) = a_f(x_0 + x) + b_f(y_0)$. The quantity x is representing the concentration variation of the first pollutant, starting from the initial concentration x_0. The variable z is representing the number of different biosensor concentrations considered in the experiment, z_0 indexing the first concentration. Similarly, for a metered addition of the second pollutant, we get:

$$\boldsymbol{\mathcal{D}}_2(y, z, \lambda) = \sum_{f=1}^{F} \beta_f(y)c_f(z)s_f(\lambda), \tag{5.3}$$

with $\beta_f(y) = a_f(x_0) + b_f(y_0 + y)$. Clearly, equations (5.2) and (5.3) express a CPD of datasets $\boldsymbol{\mathcal{D}}_1$ and $\boldsymbol{\mathcal{D}}_2$ from which the functions $\alpha_f(x)$ and $\beta_f(y)$ can be recovered.

The reason for considering three-way data rather than two-way data is mainly to uniqueness property of the CPD. Indeed, as mentioned in [Moussaoui06], an ICA method applied to these non negative data is likely to yield negative values. Similarly, the uniqueness of the NMF results cannot be ensured. On the contrary, introducing a third dimension together with the non negativity of the data results, under mild conditions, in unique non diverging CPD [Lim09]. The CPD of the three-way array can then be written as:

$$\boldsymbol{\mathcal{D}} = \sum_{f=1}^{F} \boldsymbol{\alpha}_f \circ \boldsymbol{c}_f \circ \boldsymbol{s}_f = [\![\boldsymbol{A}, \boldsymbol{C}, \boldsymbol{S}]\!]. \tag{5.4}$$

In (5.4), $\boldsymbol{A}$, $\boldsymbol{C}$ and $\boldsymbol{S}$ are the matrices formed by the column gathering of vectors $\boldsymbol{\alpha}_f$, $\boldsymbol{c}_f$, $\boldsymbol{s}_f$, $f = 1, \ldots, F$, respectively.

5.1.1.2 Estimation of iron responses of a mixture of antagonistic biosensors

Among the metallic elements of environmental importance, iron and cadmium receive special attention. Iron is present in many soil minerals and plays a fundamental role in the growth and activity of organisms. In situations of deficiency, microorganisms have developed various strategies for acquiring iron. Unlike iron, cadmium is a non-essential element that is toxic to living beings. In solution, cadmium is mainly found as a divalent cation Cd^{2+}. In this form, the metal is very toxic, even at low concentrations.

We present in this section an experimental result for the identification of biosensors' response in the presence of iron. Results on the estimation of the biosensors' response to cadmium can be found in [Caland13]. For the presented experiments we used the PAO1-2*Delta* strain *Pseudomonas aeruginosa* as host. The reporter systems are two genes involved in iron homeostasis. The first promoter *pvdA*, regulates siderophore (pyoverdine) synthesis, and is induced by iron deficiency (see Figure 5.2). The second promoter, induced by iron excess in the bacterial cytoplasm, *bfrB* regulates the synthesis of a ferritin bacterium, which is responsible for immobilizing the excessive iron. This intracellular iron regulatory system operates over a wide range of concentrations and continuously stabilizes the status of iron(III) within the cytoplasm.

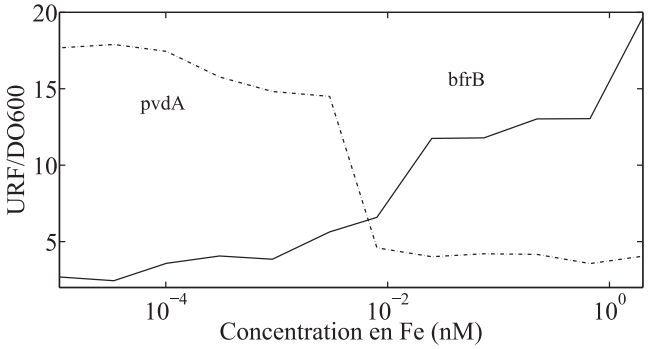


Figure 5.2: Response of the genes *bfrB* and *pvdA* to different concentrations of iron obtained by diluting $FeCl_3$. Fluorescence is expressed in units of fluorescence relative to the bacterial density estimated from an absorbance measurement at 600 nm. The fluorescence is measured at 510 nm after excitation at 490 nm.

The biosensors used in the following experiments are listed in the Table 5.1. The first column of the table contains the names of the biosensors and the following columns the construction *promoter*::*reporter* associated with the biosensor.

This series of experiments aims to identify, in a mixture, the antagonistic responses of two iron-responsive biosensors. The two reporter systems are constructed by combining the two promoters (*pvdA* and *bfrB*) with two fluorescent reporters (EYFP and GFP). This approach provides a more complete model of gene expression than a standard curve, while minimizing the number of experiments. The experimental design was constructed to generate a three-way data set respecting the CPD model. The measured fluorescence is a linear combination

Biosensor	Promoter	Fluorescent Probe	Genetic Construction
$\mathcal{B}_{1G}$	pvdA	GFP	*pvdA*::*gfp*
$\mathcal{B}_{1Y}$	pvdA	EYFP	*pvdA*::*eyfp*
$\mathcal{B}_{2G}$	bfrB	GFP	*bfrB*::*gfp*
$\mathcal{B}_{2Y}$	bfrB	EYFP	*bfrB*::*eyfp*

Table 5.1: Summary of biosensors and nomenclature used in the following experiments.

of F fluorescence sources whose intensities vary with respect to three parameters. In this experiment, the three parameters are: the iron concentration, biosensor mixture ratio and the wavelength. The data tensor $\boldsymbol{\mathcal{X}}$ is expressed by:

$$\boldsymbol{\mathcal{X}} = \sum_{f=1}^{F} \boldsymbol{a}_f \circ \boldsymbol{b}_f \circ \boldsymbol{s}_f, \tag{5.5}$$

where $\boldsymbol{a}_f$ represents source f as a function of iron concentration, $\boldsymbol{b}_f$, as a function of biosensor mixture ratio and $\boldsymbol{s}_f$, as a function of the wavelength.

The variations in the first parameter (iron concentration gradient) are obtained by a cascade dilution of an iron solution in an iron-deficient culture medium (DCAA). The range consists of twelve concentration levels from 2 mM to 11 nM obtained by successive dilutions (1/3 ratio) by a robot. The dilution plan is presented in the Table 5.2.

	1	2	3	4	5	6	7	8	9	10	11	12
A	2 mM 95:5	667 µM 95:5	222 µM 95:5	74 µM 95:5	25 µM 95:5	8 µM 95:5	3 µM 95:5	0.9 µM 95:5	0.3 µM 95:5	0.1 µM 95:5	34 nM 95:5	11 nM 95:5
B	2 mM 80:20	667 µM 80:20										11 nM 80:20
C	2 mM 60:40		222 µM 60:40								34 nM 60:40	
D	2 mM 50:50			74 µM 50:50						0.1 µM 50:50		
E	2 mM 40:60				25 µM 40:60				0.3 µM 40:60			
F	2 mM 20:80					8 µM 20:80		0.9 µM 20:80				
G	2 mM 5:95						3 µM 5:95					
H	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank

Table 5.2: Experiment design performed on a 96-well microplate. On the rows: variation of iron concentration over a range from 2mM to 11nM iron by cascade dilution of a solution of iron(III) chloride. On the columns: variation in the ratio α_{2G} or α_{2Y}, ranging from 95% to 5%, obtained by mixing the two biosensors.

The variations of the second parameter are obtained by mixing two biosensors in variable proportions. The used strain is the bacterium *Pseudomonas aeruginosa* whose iron-reactive

promoter genes *pvdA* and *bfrB* have been marked with two reporter systems *eyfp* and *gfp*. The first biosensor was designed for dose-dependent fluorescence production (GFP) via the fusion of a reporter gene to the *bfrB* promoter. The expression of the reporter gene is then under the control of the *bfrB* promoter whose induction is proportional to the amount of iron. The second biosensor also produces an iron-dependent dose-dependent fluorescence through the synthesis of a fluorescent protein EYFP whose maximum emission is shifted by 18nm towards the red with respect to GFP. The EYFP reporter gene is under the control of the *pvdA* promoter whose induction is proportional to iron deficiency. The two biosensors therefore have antagonistic iron concentration-dependent and spectrally overlapping responses.

The growth of the two biosensors ($\mathcal{B}_{1Y}$ and $\mathcal{B}_{2G}$) are carried out in a rich medium at 37°C under agitation (240 rpm). The biosensor suspensions are then rinsed in iron-deficient medium (DCAA) before being distributed in a 96-well microplate.

In each well, the same amount of biosensor is added, varying the mixing ratio in a range of 0.05 to 0.95. The biosensor suspensions are then rinsed in the iron-deficient medium (DCAA) before being distributed in a 96-well microplate:

$$\alpha_{2G} = \#\mathcal{B}_{2G}/(\#\mathcal{B}_{2G} + \#\mathcal{B}_{1Y}),$$

with $\#\mathcal{B}_{1G}$, $\#\mathcal{B}_{2G}$, $\#\mathcal{B}_{1Y}$ et $\#\mathcal{B}_{2Y}$ corresponding to the quntities of biosensors $\mathcal{B}_{1G}$, $\mathcal{B}_{2G}$, $\mathcal{B}_{1Y}$ and $\mathcal{B}_{2Y}$, respectively. An identical experimental design is used on the second biosensor mixture but with swapped markings. Green fluorescent protein (GFP) synthesis is under the control of the *pvdA* promoter, while the expression of the *bfrB* promoter is indicated by the synthesis of yellow protein (EYFP). The mixing ratio is expressed as follows:

$$\alpha_{2Y} = \#\mathcal{B}_{2Y}/(\#\mathcal{B}_{2Y} + \#\mathcal{B}_{1G}).$$

The tables 5.3 represent the two mixtures of biosensors performed. Each table represents the biosensors used and the associated theoretical behaviours of the promoters and markers. The responses of the different biosensors as a function of iron concentration ($\boldsymbol{a}_i$) are non-proportional ($\boldsymbol{a}_1 \neq \boldsymbol{a}_2$). Similarly, biosensor mixtures and markers were chosen to avoid collinearities ($\boldsymbol{b}_1 \neq \boldsymbol{b}_2$ and $\boldsymbol{s}_1 \neq \boldsymbol{s}_2$).

First mixture			
	a	b	s
$\mathcal{B}_{1G}$	$\boldsymbol{a}_1$	$\boldsymbol{b}_1$	$\boldsymbol{s}_1$
$\mathcal{B}_{2Y}$	$\boldsymbol{a}_2$	$\boldsymbol{b}_2$	$\boldsymbol{s}_2$

Second mixture			
	a	b	s
$\mathcal{B}_{1Y}$	$\boldsymbol{a}_1$	$\boldsymbol{b}_1$	$\boldsymbol{s}_2$
$\mathcal{B}_{2G}$	$\boldsymbol{a}_2$	$\boldsymbol{b}_2$	$\boldsymbol{s}_1$

Table 5.3: Representation of the expected behaviors ($\boldsymbol{a}_i$, $\boldsymbol{b}_i$) for the pairs of biosensors used in the two mixtures performed. In the second mixture, the promoter marking was swapped.

After incubation (24 hours at 37°C), different types of fluorescence spectra were measured in each of the microplate wells:

- fluorescence emission spectra measured between 506 and 560 nanometers with a step size of 2 nanometers (27 points) for an excitation wavelength of 485 nanometers,
- synchronous fluorescence spectra measured between 470 and 530 nanometers with a 20-point step for a 40 $\Delta\lambda = 20$ nanometers.

Thus, four experiments, in addition to a single-wavelength measurement, were carried out combining spectral measurements and biosensor constructions. The CPD of the emission and synchronous spectra for each of the two biosensor mixtures ($\mathcal{B}_{1Y}$, $\mathcal{B}_{2G}$) and ($\mathcal{B}_{1G}$, $\mathcal{B}_{2Y}$) were studied. We present hereafter some results for the emission spectra for the biosensor mixture ($\mathcal{B}_{1Y}$, $\mathcal{B}_{2G}$). More results can be found in [Caland13].

CPD analysis of emission spectra of the mixture of biosensors ($\mathcal{B}_{1Y}$, $\mathcal{B}_{2G}$) On each well, the emission spectra are measured between 506 and 560 nanometers with a step of 2 nanometers for a 495 nanometers excitation. For each experiment, the measurements are grouped in a three-dimensional table with a dimension of $27 \times 12 \times 7$. The first mode describes the wavelength of the fluorescence spectrum, the second describes the iron concentration and the third describes the biosensors ratio. The spectral measurements obtained are presented on Figure 5.3 as a function of iron concentration and biosensor mixture ratio.

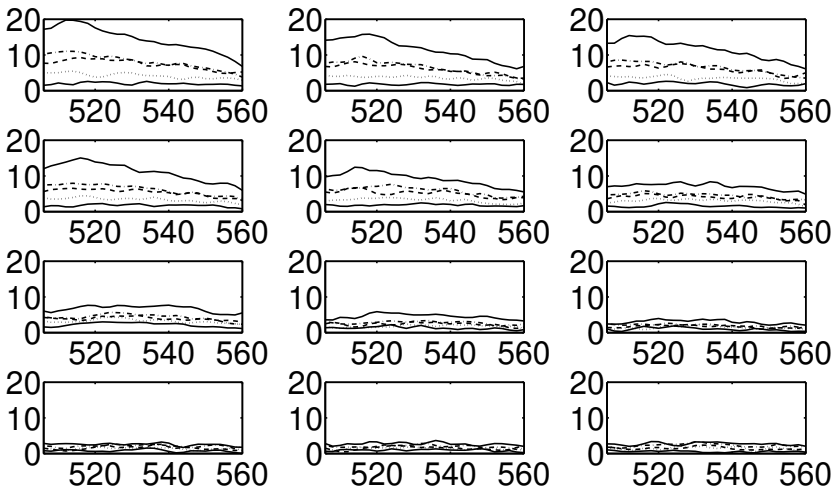


Figure 5.3: Raw fluorescence spectra mesured in the range 506 - 560 nm, with a step of 2 nm ($\lambda_{exc} = 405$ nm). Each subplot gathers the spectra obtained for one iron concentration (one well column). Each spectrum corresponds to the fluorescence mesured in one well for a given biosensor mixture ratio $\mathcal{B}_{2G}$ and $\mathcal{B}_{1Y}$.

Figure 5.4 shows the result of the CP decomposition for three sources, obtained with the first mixture of biosensors $\mathcal{B}_{2G}$ and $\mathcal{B}_{1Y}$. One can see the evolution of the intensity of the different spectral sources (Figure 5.4(a)) as a function of the iron concentration (Figure 5.4(b)) and the α_{2G} ratio of the mixture (Figure 5.4(c)). The Figure 5.4(d) represents

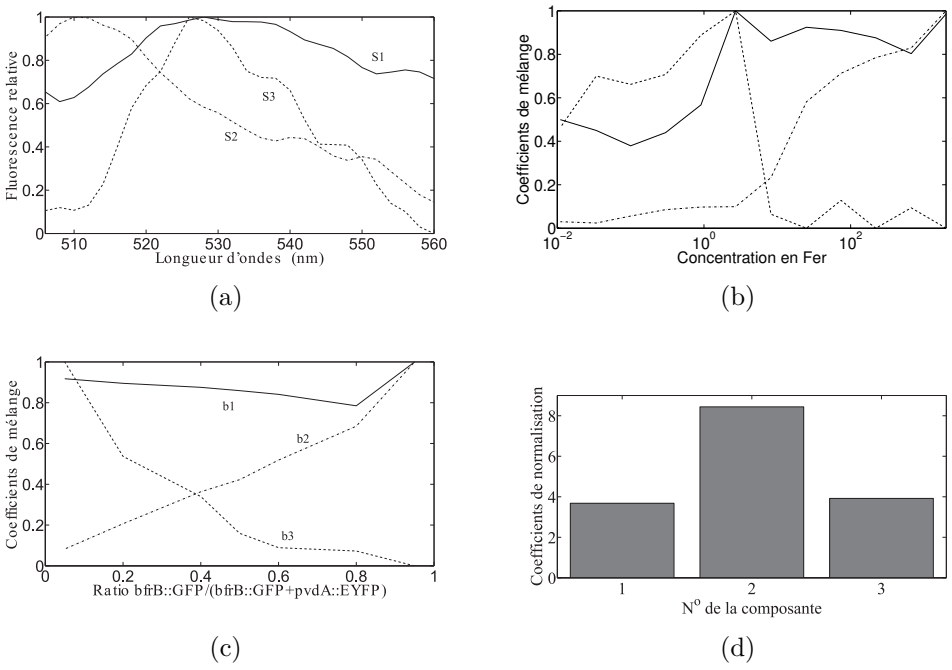


Figure 5.4: CPD result of the fluorescence spectra for the biosensors mixture ($\mathcal{B}_{2G}$, $\mathcal{B}_{1Y}$). Sources #1 (autofluorescence), #2 ($\mathcal{B}_{2G}$), #3 ($\mathcal{B}_{1Y}$). (a) Emission spectra for the identified sources. (b) Sources' responses with respect to the iron concentration. (c) Evolution of the estimated responses with respect to biosensors mixinng ratio α_{2G}. (d) Normalization coefficients obtained by setting to 1 the maximum value of responses in each of the three modes.

the normalization coefficients of the sources. The decomposition shows the existence of a #1 source, without any strong maximum, which covers almost the whole wavelength range. This source grows strongly for concentrations greater than 1 µM and is insensitive to the α_{2G} ratio of biosensors. The two other estimated spectral sources ($\boldsymbol{s}_2$, $\boldsymbol{s}_3$) have maxima at 512 nm and 527 nm, respectively. The estimation of mixing coefficients shows that these sources behave inversely as a function of iron concentration and biosensor ratio. The responses $\boldsymbol{a}_2$ and $\boldsymbol{b}_2$ are increasing and monotonous while the response $\boldsymbol{a}_3$ is increasing for concentrations 0.1 µM to 1 µM and then decreasing for higher concentrations, and the response $\boldsymbol{b}_3$ is decreasing. The normalization coefficients (Figure 5.4(d)) show a source of fluorescence #2 more intense than the other two sources that have similar levels.

Based on the estimated responses and spectra, the sources can be identified. It makes sens to attribute source #1 to an intrinsic autofluorescent signal from the bacterial cells. Its decrease for low iron concentrations is explained by a disturbed bacterial development: little iron, little growth. Moreover, since this source is constitutive of both biosensors, it does not evolve according to the ratio of biosensors.

The estimated spectral source $\boldsymbol{s}_2$ is consistent with the known emission spectrum for the fluorescent protein GFP, whose optimum is 510 nm. Estimation of mixing coefficients by the CPD method also shows an increase in the GFP signal with increasing iron concentration, which is the expected behaviour for the $bfrB$ promoter. The quasi-linear increase of this signal with the biosensor ratio follows the increase of the biosensor $\mathcal{B}_{2G}$ in the experimental mixture. Similarly, the estimated spectrum $\boldsymbol{s}_3$ corresponds to the expected spectrum of the fluorescent protein EYFP. Its production depends on the induction of the *pvdA* gene which decreases with increasing iron concentration. This phenomenon is highlighted by the response curve estimated by CPD (curve $\boldsymbol{a}_3$). We can also see an increase in the intensity of fluorescence for concentrations below 1 µM. This is explained by an increase in the quantity of bacteria. The decrease in mixing coefficients ($\boldsymbol{b}_3$) as a function of the ratio α_{2G} of is consistent with the association of the source #3 to the EYFP concentration.

The experiment results for the second biosensor mixture ($\mathcal{B}_{2Y}$, $\mathcal{B}_{1G}$), can be found in [Caland13]. The conclusion of these experiments is that the choice of the protein marking the gene under study can influence the quality of the estimation result. Marking a weakly expressed gene with a low-fluorescence protein reduces the quality of source separations and leads to a poor estimation of the responses. It is then important to use knowledge of the system in order to develop biosensors and protocols is in a way that improves the estimation of responses by source separation methods.

5.1.2 Non-negative block tensor decomposition for flow cytometry data

Flow cytometry is an investigation technique widely used in biology and medicine for the characterization and quantification of the morphological, density and fluorescence properties of cells. The recent technological progresses of cytometers allow the design of complex

multiparameter experiments yielding a large amount of data. Classical flow cytometry data analysis methods are no longer adapted to these data and hence there is a need for new algorithms that can efficiently retrieve the relevant information from this large amount of data.

In [Brie14a, Brie14b] we introduced a cell population sorting method based on a non-negative block tensor decomposition [De Lathauwer08a, De Lathauwer08b] of the data histograms. The key point is that it is a fully multidimensional approach in which cell sorting is done according to the response of the different sub-populations to a parameter. The main advantages of our method are: it is almost fully unsupervised (the only input parameter is the number of sub-populations sought in the data), it is non-parametric (there is no underlying parametric probability density function), it is not much affected by the overlapping of the sub-population distributions and it does not require any decision step since the problem is addressed as a source separation problem which provides both the amount of each sub-population and its probability density function. We also proved that there is a relationship between flow cytometry and bulk spectroscopy data, and proposed a joint processing approach of the two data modalities, that improves the separation accuracy and provides a more complete description of the analyzed cell populations. This joint data processing approach is not presented in this manuscript but the interested reader is referred to [Brie14a] for the details.

5.1.2.1 The probability density function of N-dimensional flow cytometry data model

Consider N-dimensional flow cytometry data. Each of the analyzed cells yields a length N vector measuring the amplitudes at N different wavelength values of the emitted fluorescence light. The set of measurements collected on a population of M different cells can be gathered in a $N \times M$ matrix $\boldsymbol{X} = [\boldsymbol{x}_1 \cdots \boldsymbol{x}_M]$, where $\boldsymbol{x}_m = [x_m(1), \cdots, x_m(N)]^\mathsf{T}$ and $m = 1, \ldots, M$. As illustrated in figure 5.5, this data matrix $\boldsymbol{X}$ can be characterized by its N-variate *pdf* denoted by $p(\boldsymbol{x}) = P(\boldsymbol{x}_m = \boldsymbol{x})$.

The bilinear model of a sequence of N-D pdfs Let us consider a cell population composed of K different sub-populations. A sub-population is defined as a set of cells exhibiting identical/similar behaviors with respect to the variation of a physical parameter. We represent the N-D data points in the analyzed sample by the *pdf* $p(\boldsymbol{x})$ of the measurement vector $\boldsymbol{x}$. This *pdf* is expressed as a mixture of K density functions f_k, corresponding to the K sub-populations:

$$P(\boldsymbol{x}_m = \boldsymbol{x}) = p(\boldsymbol{x}) = \sum_{k=1}^{K} \alpha_k f_k(\boldsymbol{x}) \tag{5.6}$$

with $\sum_{k=1}^{K} \alpha_k = 1$. Assume that we study the response of this cell population to the variation of some physical parameter, denoted s hereafter. For each value of s, a flow cytometry dataset can be recorded, resulting in a sequence of flow cytometry data matrices. Hence, each data

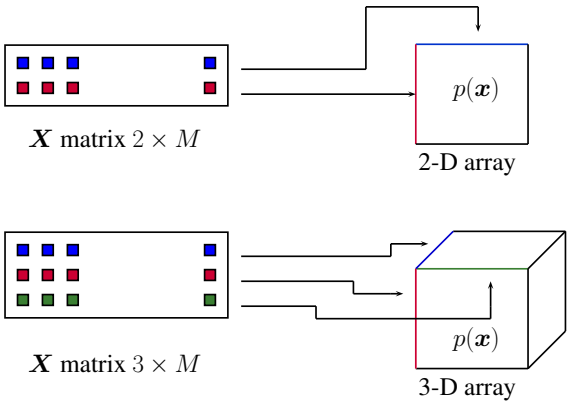


Figure 5.5: Transforming the $(N \times M)$ matrix $\boldsymbol{X}$ into N-D histograms; illustrations for $N = 2$ and 3.

matrix obtained for a given physical condition yields a *pdf* denoted by $p(\boldsymbol{x}, s)$, that can be modeled as:

$$p(\boldsymbol{x}, s) = \sum_{k=1}^{K} \alpha_k(s) f_k(\boldsymbol{x}). \tag{5.7}$$

The sequence of N-D histograms, obtained for different values of s, can thus be gathered into a $(N+1)$-D array (tensor) denoted by $\boldsymbol{\mathcal{P}}$. By unfolding this tensor along the dimensions corresponding to the N different wavelengths, we obtain a matrix $\boldsymbol{P}$ that admits the following bilinear factorization:

$$\boldsymbol{P} = \boldsymbol{A}\boldsymbol{F}^{\mathsf{T}}. \tag{5.8}$$

In (5.8), $\boldsymbol{P}$ has a number of rows equal to the number of values of the parameter s; the columns of $\boldsymbol{A}$, symbolized by $\boldsymbol{a}_k$ $(k = 1, \ldots, K)$, contain the mixing coefficients of the K sub-populations for the different values of s; the columns of $\boldsymbol{F}$ are the "unfolded" N-D density functions $\boldsymbol{f}_k$ of the K sub-populations.

The Block-CPD model of a sequence of N-D pdfs Assuming the independence of each coordinate of $\boldsymbol{x} = [x_1, \cdots, x_N]^{\mathsf{T}}$, the multivariate *pdf* $f_k(\boldsymbol{x})$ can be factorized as a product of N univariate *pdf*s: $f_k(\boldsymbol{x}) = f_k^1(x_1) \cdot f_k^2(x_2) \cdots f_k^N(x_N)$. Thus the data array can be written as a CPD model of order $N + 1$:

$$\boldsymbol{\mathcal{P}} = \sum_{k=1}^{K} \boldsymbol{a}_k \circ \boldsymbol{f}_{k,1} \circ \cdots \circ \boldsymbol{f}_{k,N}. \tag{5.9}$$

Equation (5.9) clearly expresses a $N + 1$ CPD model of rank K which can be alternatively written as:

$$\boldsymbol{\mathcal{P}} = [\![\boldsymbol{A}, \boldsymbol{F}_1, \cdots, \boldsymbol{F}_N]\!], \tag{5.10}$$

where $\boldsymbol{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_K]$ and $\boldsymbol{F}_n = [\boldsymbol{f}_{1,n}, \ldots, \boldsymbol{f}_{K,n}]$, with $n = 1, \ldots, N$. The link between models (5.10) and (5.8) is obtained by unfolding $\boldsymbol{\mathcal{P}}$ into a matrix, according to: $\boldsymbol{P} = \boldsymbol{A}(\boldsymbol{F}_1 \odot \cdots \odot \boldsymbol{F}_N)^{\mathsf{T}}$. Thus, we have $\boldsymbol{F} = \boldsymbol{F}_1 \odot \cdots \odot \boldsymbol{F}_N$. Admittedly, assuming the independence of each coordinate of $\boldsymbol{x}$ does not allow to represent general probability density function. Restricting our attention to the case of bi-dimensional density functions ($N = 2$), we propose to adopt for the data the rank-$(L_k, L_k, 1)$ Block Component Model, equally known as Block-CPD (BCPD) model, introduced by De Lathauwer in [De Lathauwer08a, De Lathauwer08b]. By doing so, it is possible to consider the more general case of non-separable *pdf*s. In fact, this is nothing but performing a low rank approximation of the (discretized) *pdf*s. Hence, the $(N+1)$-D data array can be written as:

$$\boldsymbol{\mathcal{P}} = \sum_{k=1}^{K} \boldsymbol{a}_k \circ \boldsymbol{E}_k, \tag{5.11}$$

where the rank L_k matrices $\boldsymbol{E}_k$ can be decomposed as:

$$\boldsymbol{E}_k = \sum_{l=1}^{L_k} \boldsymbol{f}_{k,1}^{l} \circ \boldsymbol{f}_{k,2}^{l} = \sum_{l=1}^{L_k} \boldsymbol{f}_{k,1}^{l} \boldsymbol{f}_{k,2}^{l\,\mathsf{T}}. \tag{5.12}$$

A graphical illustration of this model is given on figure 5.6.

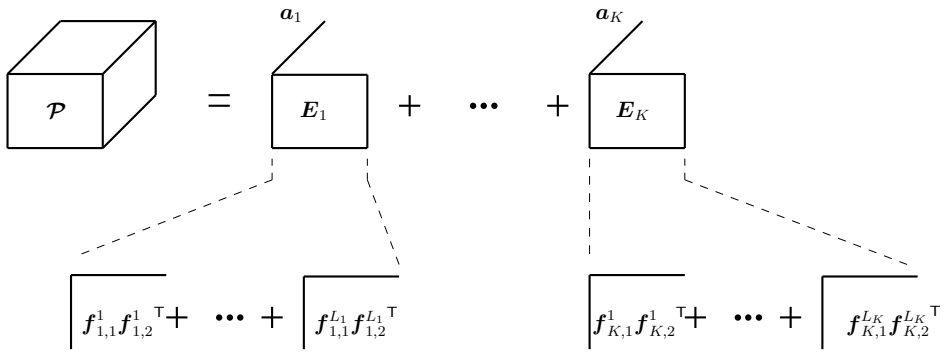


Figure 5.6: Graphical illustration of the rank-$(L_k, L_k, 1)$ BCPD model of $\boldsymbol{\mathcal{P}}$.

The BCPD model can be seen as a CPD model in which some of the loading vectors (columns) of the matrix $\boldsymbol{A}$ are collinear. The model in (5.11) can be easily generalized to higher dimensional ($N > 2$) data. The higher-dimensional case does not yield more complicated data processing situations, since it is known that higher-order CPD models require less restrictive identifiability conditions (see *e.g.,* [Brie11]).

Model identifiability The problem at hand can be embedded in the general framework of non-negative approximation of non-negative tensors using multilinear decompositions. This is still an open problem but a milestone was reached with the work of Lim and Comon [Lim09] where it is proved that non-negativity ensures the well posedness of the non-negative tensor approximation. As mentioned in [Lim09], this can actually be associated to the sparse

naive Bayes probabilistic model for *pdf* [Lowd05], since the underlying probabilistic model is a mixture of densities having independent variables. Here we go one step further since, by using the BCPD, we can relax the independence assumption. However, the question of the validity of the non-negative BCPD as an approximation tool is an open problem which deserves further studies. We did not pay further attention to this point, but as a first attempt to illustrate the regularization property of the non-negativity, we showed that rank $(L_k, L_k, 1)$ exact non-negative BCPD decomposition can be unique without any additional assumption. This is not the case for general BCPD unless some other constraints (such as orthogonality) are enforced. We showed in the previous paragraphs that NMF is involved in the considered models. In that respect, uniqueness of the NMF plays a central role in the uniqueness of the non-negative BCPD. In [Brie14a] we studied the uniqueness of the non-negative BCPD from the NMF perspective and gave some practical results. The interested reader is referred to the up-mentioned manuscript for details.

5.1.2.2 The data processing algorithm

The data processing method proposed in [Brie14a] consists of three steps. The first two steps deal with the processing of the flow cytometry data while the third one addresses the coupling (fusion) of the flow cytometry data with the cell fluorescence spectra. We present hereafter only the first two steps.

Estimation of the probability density functions The flow cytometry data *pdf* is performed by computing the N-D histograms of the data which correspond to hypercube Parzen window function. The developed algorithm requires to define the number of bins (one for each dimension) on which the histograms are calculated. Here, only 2-D histograms are considered. For all experiments, the number of bins was fixed to 50 along each dimension which results in *pdf* estimations showing a good tradeoff between resolution and accuracy. In fact, in practice the setting of this parameter value remains supervised and depends on the data at hand.

Non-negative BCPD of the data For the needs of this step, the non-negative BCPD algorithm developed in the tensorlab toolbox [Sorber13] can be used. We also developed a procedure to estimate the ranks L_k and K of the decomposition. It first consists in performing a non-negative high-rank CPD decomposition of the data. Thanks to the partial uniqueness properties of the CPD model [Guo12], the matrix containing the collinear loadings is ensured to be unique. In practice, because of the noise/error terms, the estimated loading may not be strictly collinear. The "most" collinear loadings, *i.e.* those presenting a correlation greater than a specified threshold (typically 0.9), are collapsed into a single loading by an averaging procedure. The K resulting averaged loadings are gathered into a full column rank matrix $\boldsymbol{A}$. The corresponding loadings of the other two modes are gathered to form matrices $\boldsymbol{E}_k$, $k = 1, \ldots, K$. This results in a non-negative BCPD decomposition which is used as initial solution for the non-negative BCPD algorithm of [Sorber13]. Once the non-negative BCPD

decomposition is achieved, the corresponding N-D *pdf*s are normalized to have unit sum and the normalization factor is then transferred on the corresponding loadings, representing the responses of the K sub-populations to the physical parameter.

5.1.2.3 Results on cell cytometry data

We illustrate the effectiveness of the proposed approach on real flow cytometry data.

Data acquisition The experimental study aims at validating the proposed data processing approach on real multicolor flow cytometry data corresponding to the response of a cell line to a widely used and well understood uncoupling agent (carbonyl cyanide p-chlorophenylhydrazone - CCCP, Sigma-Aldrich). CCCP is an ionophore used to uncouple oxidative phosphorylation in mitochondria. It causes a mitochondrial proton leak, leading to a depolarization of the mitochondrial membrane. Thus, it is frequently used as a negative control in mitochondrial membrane potential measurements by flow cytometry. Human ductal breast epithelial tumor cell line, T47D (from ATCC) was grown in RPMI 1640 medium supplemented with 10% fetal calf serum, 2mM L-glutamine and 5 μg/ml Gentamicin at 37 °C in a humidified atmosphere containing 5% CO_2. Thus, we chose 6 CCCP concentrations varying between 0 and 100 μM ($[CCCP] = 0,\ 5,\ 10,\ 25,\ 50,\ 100$) to ensure that the whole CCCP response range is observed. The fluorescence signal was measured in two channels : "green " ($515 - 545$nm) and "red" ($564 - 606$nm). Detailed information on the cell culture and data acquisition conditions is available in [Brie14a].

Sample processing results The first step of the processing consists in estimating the 2-D histograms. For each dimension, the number of bins is fixed to 50 resulting in a 50×50 data matrix. Then the six matrices are gathered into a 3-way array of dimension $(50 \times 50 \times 6)$. For this data set, the compensation was fixed to have a maximum separation along the horizontal axis.

Figure 5.8 shows the results of the non-negative BCPD corresponding to the dataset of figure 5.7. The number of block-component is expected to be equal to 2: a highly polarized cell sub-population whose response to CCCP is expected to decrease and a depolarized cell sub-population whose response is expected to increase with the CCCP concentration. The experimental parameters L_k of BCPD were determined after successive trials. The first block rank is $(3, 3, 1)$ while the second block rank is $(2, 2, 1)$. The responses of the two sub-populations are in very good agreement with what was expected.

From a biological point of view, the top left-hand side plot on figure 5.8 represents the distribution of cells with a low mitochondrial membrane potential (referred to as "green" fluorescence) while the top right-hand side figure corresponds to the cells with high mitochondrial membrane potential (referred to as "red" fluorescence). The associated responses show that the low mitochondrial membrane potential sub-population increases with the concentration of CCCP while the high mitochondrial membrane potential sub-population decreases. In

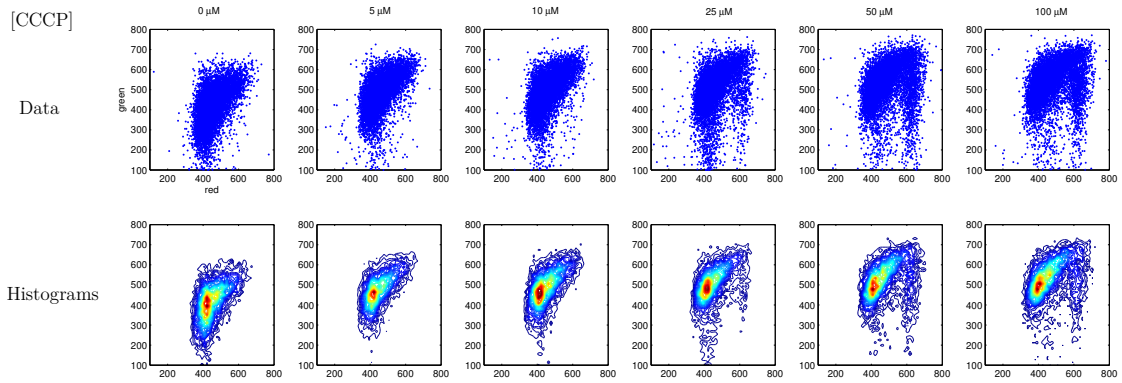


Figure 5.7: A sequence of flow cytometry data showing the response of T47D cells to CCCP. The upper line figures correspond to the data and the lower line figures are the corresponding 2-D histograms.

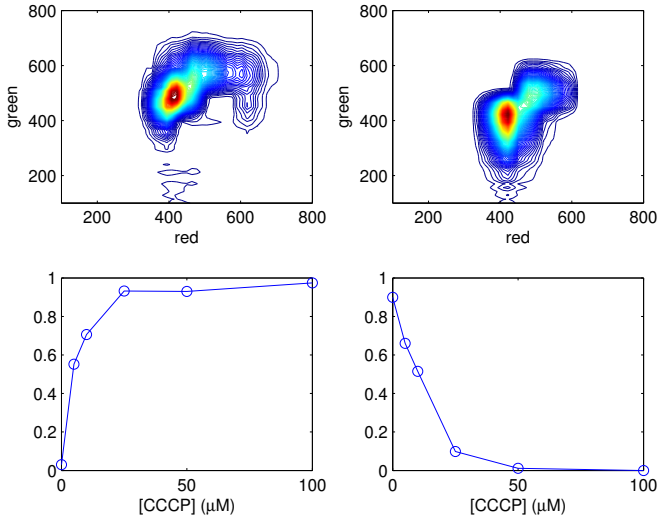


Figure 5.8: Non-negative BCP decomposition results of flow cytometry dataset corresponding to the first experiment

particular, it can be observed that the full cell population depolarization is reached after a concentration of 50 μM which is corresponding to the value generally accepted by practitioners. To evaluate the reproducibility of the experiments, we repeated it three times. The rank of the BCP decomposition were fixed to the same values as in the first experiment. From the results of the non-negative BCPD of the other two datasets it appears that the reproducibility of the results is good: not only the CCCP responses are quite similar but also the cell sub-population distributions are quite similar.

Further results and comments as well as a joint analysis of cytometry and spectroscopy data can be found in [Brie14a].

5.2 Tensors for polarized Raman spectroscopy signal processing

I present in this section a joint work with colleagues from the "Laboratoire de Chimie Physique et Microbiologie pour l'Environnement" (LCPME) on the used of polarization diversity in Raman spectroscopy.

Polarization analysis in Raman spectroscopy provides a powerful tool in chemical physics to identify the symmetry of the vibration modes of molecules, macromolecules and crystals, according to their crystallographic structure and local spatial orientation. In [Miron11] we proposed a source separation approach in polarized Raman spectroscopy for data with rotational and spatial diversity and showed that the joint use of parallel and crossed polarization data yields more accurate source separation results and improves the uniqueness properties of the solution in blind nonnegative source separation algorithms. Two signal processing models for the Raman polarized spectra, acquired with angular and spatial diversity, were introduced. I will only present hereafter the spatial diversity case, as it results in a tensor-based algorithm.

5.2.1 Some physical considerations

To develop the mathematical models for the polarized Raman data, it is necessary to briefly remind the physical phenomenon related to Raman scattering of polarized light. A useful quantity in this context is the *polarizability tensor*. The polarizability tensor is a 3×3 real-valued matrix $\boldsymbol{\alpha}$, relating the induced electric dipole moment $\boldsymbol{p}$ in the crystal to the electric field vector $\boldsymbol{e}$ of the incident light. In the crystal-fixed coordinate system (O, x, y, z) the polarizability tensor can be written as:

$$\boldsymbol{\alpha} = \begin{pmatrix} \alpha_{xx} & \alpha_{xy} & \alpha_{xz} \\ \alpha_{yx} & \alpha_{yy} & \alpha_{yz} \\ \alpha_{zx} & \alpha_{zy} & \alpha_{zz} \end{pmatrix}. \tag{5.13}$$

A Raman polarizability tensor is associated to a particular *vibrational mode* for a given crystal orientation [Long02, Turrell72]. The vibrational modes will be thought of as sources in the source separation problems presented in this manuscript. As the measurements are performed in the laboratory coordinates system (O, X, Y, Z), a rotation matrix $\boldsymbol{R}(\varphi, \theta, \chi)$ [Long02, Turrell72], accounts for the coordinate system change between the sample and the laboratory frames. Denoting the Euler angles describing the sample rotations in the coordinate system (O, X, Y, Z) by $\boldsymbol{\theta} = (\varphi, \theta, \chi)$, the induced electric vector in non polar crystal can be expressed by the tensorial change of basis:

$$\boldsymbol{p}(\boldsymbol{\theta}) = \boldsymbol{R}^{\mathsf{T}}(\boldsymbol{\theta})\, \boldsymbol{\alpha}\, \boldsymbol{R}(\boldsymbol{\theta})\, \boldsymbol{e} \quad . \tag{5.14}$$

If we denote by p_X, p_Y, p_Z, the three components of $\mathbf{p}$ along X, Y and Z axes, the intensity of the scattered light, measured by the spectrometer, is given by the following proportionality relationship:

$$\boldsymbol{i}(\boldsymbol{\theta}) = \begin{pmatrix} i_X & i_Y & i_Z \end{pmatrix} \mathsf{T} \propto \begin{pmatrix} p_X^2 & p_Y^2 & p_Z^2 \end{pmatrix}^{\mathsf{T}}. \tag{5.15}$$

In practical applications in backscattering geometry only i_X and i_Y are measured, corresponding to the crossed and parallel polarizations, respectively.

We will suppose in the following that the recorded scattered light is a mixture of N sources and that K wavelengths/wavenumbers are acquired for each spectrum in one point on the crystal surface.

5.2.2 The proposed approach

We present in this section a tensor model for Raman imaging data using both crossed and parallel polarized lights.

5.2.2.1 Data model

For the spatial diversity scheme, two polarized spectra are acquired in L different points denoted by: $d_1, \ldots, d_L$, on a regular spatial grid. Define the following matrices:

$$\boldsymbol{P} = \begin{pmatrix} i_{1,X} & \ldots & i_{N,X} \\ i_{1,Y} & \ldots & i_{N,Y} \end{pmatrix} \in \mathbb{R}^{2\times N} \quad , \tag{5.16}$$

$$\boldsymbol{B} = \begin{pmatrix} b_1(d_1) & \ldots & b_N(d_1) \\ \vdots & \ddots & \vdots \\ b_1(d_L) & \ldots & b_N(d_L) \end{pmatrix} \in \mathbb{R}^{L\times N} \quad . \tag{5.17}$$

The $\boldsymbol{B}$ matrix contains on its columns the spatial evolution of the source concentrations and $\boldsymbol{P}$ characterizes the behavior of the N sources for the two polarization orientations. The acquired data can thus be expressed as two $L \times K$ matrices:

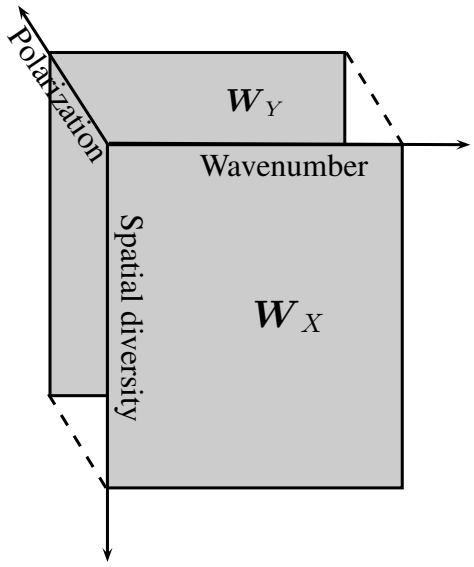


Figure 5.9: Illustration of the three-way data structure and the corresponding diversities.

$$\boldsymbol{W}_p = \boldsymbol{B}\boldsymbol{D}_p(\boldsymbol{P})\boldsymbol{S}^{\mathsf{T}} + \boldsymbol{E}_p, \quad p = X, Y \quad , \tag{5.18}$$

where $\boldsymbol{D}_p(\boldsymbol{P}) = \text{diag}\,(i_{1,p}, \ldots, i_{N,p})$ is a diagonal matrix which takes the first or the second row of $\boldsymbol{P}$ as its diagonal, and $\boldsymbol{E}_p \in \mathbb{R}^{L\times K}$ the noise matrix. Relation (5.18) clearly expresses a three-way CPD model for the polarized data with spatial diversity. Figure 5.9 illustrates the three-dimensional structure of the data.

The trilinear model validity is conditioned by the invariance of the source polarization behavior with respect to spatial displacement. In practice, this is the case if `i)` the analyzed objects have well-defined spatial orientation and their size is bigger than the laser spot or `ii)` they have random orientation and their size is largely inferior to the laser spot.

5.2.2.2 Identifiability issues for the trilinear model

Before presenting the algorithms for fitting the spatial diversity data, identifiability conditions for the trilinear mixture model expressed by Eq. (5.18) must be discussed. The spatial acquisition scheme yields a trilinear CPD model as one can see in Eq. (5.18). The inverse problem for this model is the estimation of $\boldsymbol{B}, \boldsymbol{S}$ and $\boldsymbol{P}$ from the observations $\boldsymbol{W}_p$. Applied to our case, Kruskal's condition (see Eq. 3.6) states that one can uniquely estimate the matrices $\boldsymbol{B}, \boldsymbol{S}$ and $\boldsymbol{P}$ in Eq. (5.18) if:

$$\text{krank}\{\boldsymbol{B}\} + \text{krank}\{\boldsymbol{S}\} + \text{krank}\{\boldsymbol{P}\} \geq 2(N+1) \quad . \tag{5.19}$$

We shall only consider next, the case where $\boldsymbol{S}$ is a tall matrix, meaning that there are more spectral points than sources (very common in practice). Given the spiky nature of Raman

spectra, $\boldsymbol{S}$ is full column rank with high probability, implying krank$\{\boldsymbol{S}\} = N$. Kruskal's condition can thus be re-written as:

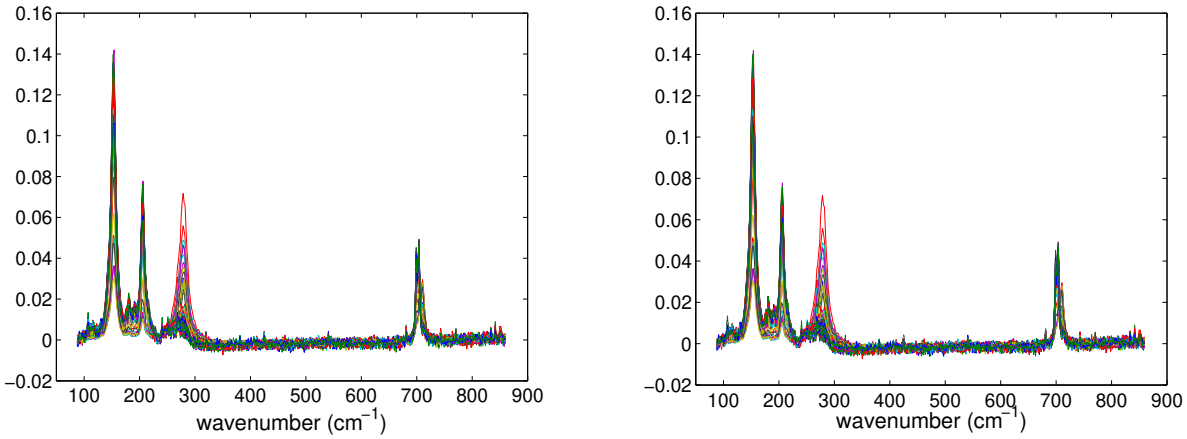
$$\text{krank}\{\boldsymbol{B}\} + \text{krank}\{\boldsymbol{P}\} \geq N + 2 \quad . \tag{5.20}$$

Given the size of matrices $\boldsymbol{B}$ and $\boldsymbol{P}$, and knowing that the Kruskal-rank of a matrix is upper bounded by its classical rank, the only possible solution to inequality (5.20) is krank$\{\boldsymbol{P}\} = 2$ and krank$\{\boldsymbol{B}\} = N$. This means that, in order to have an unique factorization, every two sources must have different polarization behaviors and $\boldsymbol{B}$ (with $L \geq N$) has to be full column rank. Thus, in practice, the CPD model (5.18) is identifiable in most cases, provided that more acquisition points than sources are used.

To fit the CPD model (5.18), we used an optimized nonnegative algorithm for three-way factorization from the MATLAB `N-way` toolbox [Andersson00].

5.2.3 Results on real data

We analyzed a mixture of two polymorphs of $CaCO_3$ having a distinct crystallographic structure. Though the Raman active modes involve similar atomic displacements, the change of the crystal symmetry produces significant separation of the Raman peaks for these two polymorphs. Some peaks are overlapped (for instance, around 700 cm^{-1}) but others are clearly well separated. Therefore, one source is classically viewed as the entire Raman spectrum of one polymorph in this section.



(a) Crossed polarization

(b) Parallel polarization

Figure 5.10: Polarized data for $CaCO_3$

For the illustration of the trilinear data model, we used a calcite crystal of several micrometers in diameter grown on an aragonite substrate. A spatial grid of 8×9 with a spatial step of 1 micrometer was read on crossed and parallel polarizations. The 2×72 resulting spectra are presented in Figure 5.10 (a) and (b). The spectral range was $100 - 900$ cm^{-1}.

In this case a source is no longer assimilated with one vibrational mode but rather with the totality of spectral peaks characterizing calcite or aragonite.

The estimated spectral signatures and mixing coefficients for the two crystalline varieties of $CaCO_3$ are presented on Figure 5.11 and Figure 5.12, respectively. As one can see in Figure 5.11 the spectral signatures of the two sources are well recovered, in perfect agreement with the literature [Carteret09]. The fact that there are a lot of zero-values in the estimated spectra is due to negative values present in the analyzed data (see Figure 5.10) as a side effect of the baseline removal in the preprocessing steps. The estimation of the mixing coefficients is also accurate, allowing to plot low resolution maps of the aragonite and calcite concentrations (Figure 5.12 (a) and (b)), matching the visual image obtained in the microscope oculars. The concentration maps reported in Figure 5.12 are also coherent with the fact that the diameter of the laser spot was around 2 micrometers and the size of the calcite crystal was a few micrometers. The size of the calcite single crystal is bigger than the laser spot and the crystal orientation is fixed. Therefore, this source obeys the model hypothesis stated in section 5.2.2. The second source corresponds to aggregates of nano-crystals of aragonite polymorph, which are randomly oriented, and of size quite inferior to the laser spot. Thus, this second source also satisfies the model hypothesis.

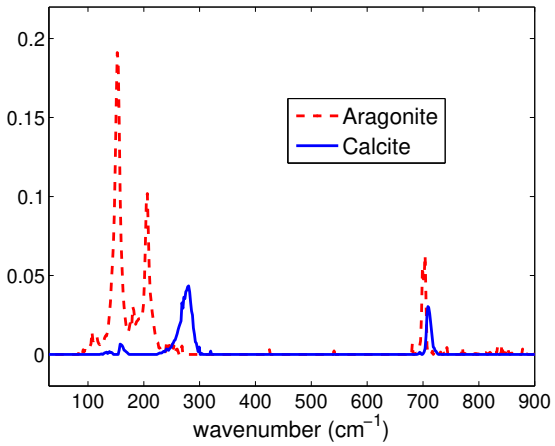
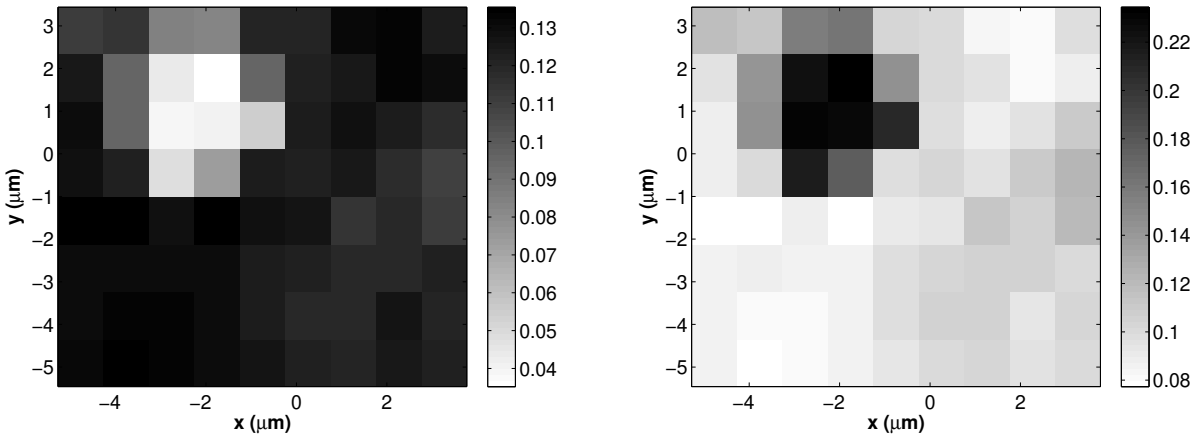


Figure 5.11: Spectra separation result $CaCO_3$

5.3 Conclusions

I presented in this chapter some of my contributions regarding the application of tensor models and methods to spectroscopy data processing. Two types spectroscopy signals were considered: cell fluorescence and polarized Raman for cristal analysis.

The first contribution is the development of a tensor representation for the fluorescence signal emitted by biosensors, that permits to jointly estimate biosensors' responses to several pollutants. Secondly, we derived a new approach for flow cytometry data analysis based



(a) Aragonite concentration map

(b) Calcite concentration map

Figure 5.12: Concentration maps for the $CaCO_3$ sample

on a non-negative block tensor decomposition. Unlike the classically employed clustering methods, strongly relying on the user expertise, our method requires only the knowledge of the number of sub-populations to be extracted from the data, and is very little sensitive to the overlapping of cell sub-population distributions. The third contribution of this chapter is the representation of the polarized Raman spectroscopy data with spatial diversity as a three-way CPD model. Besides improving source estimation accuracy, we proved that the use of the proposed model ensures mixture identifiability in Raman imaging.

Chapter 6

Constrained unmixing of multidimensional sources

This chapter is dedicated to my work on multidimensional source unmixing subject to various constraints. Section 6.1 gives results on binary data array decomposition developed in the Ph.D. of Mamadou Diop (see Section 1.6: D.3). In Section 6.2 I present a sequential algorithm for real-time unmixing of hyperspectral images, proposed in the Ph.D. of Ludivine Nus (see Section 1.6: D.4). Finally, In Section 6.3 I introduce a quaternion framework for hyperspectral source separation with polarization diversity, developed in the Post-doctoral work of Julien Flamant (see Section 1.6: pD.2).

6.1 Boolean decomposition of binary data arrays

Categorical data in general, and *binary* data in particular, are increasingly present in various domains. Binary data can be either generated by *two state processes/phenomena* (and naturally encoded on the binary $\{0, 1\}$ values), *e.g.* proximity sensors, push buttons, electric switches, yes/no answers in surveys, *etc.*, or artificially obtained by *high (1-bit) quantization* of real-valued data. Therefore, *binary matrix factorization (BMF)* and *binary tensor factorization (BTF)* gained a lot of interest lately and have already been successfully used in many domains such as digital telecommunications [Talwar96, VanderVeen97], role mining [Lu12], phenotype identification from gene expression data [Meeds06, Zhang10], protein complex prediction [Tu11], document classification [Li05], feature reconstruction in images [Meeds06], pattern discovery for gene expression images [Shen09] or recommendation systems [Nenova13].

While the BMF methods mentioned above use fundamentally different algorithmic approaches, they are all based on a mixture model defined using the classical real-matrix product. This implies (as explained in the following section) that these decompositions have the tendency to generate "decorrelated" factors (sources), as they strongly penalize the factors having overlapping supports. This represents a major drawback in a large class of applica-

tions where common behaviors of different "populations" are sought. Moreover, by forcing the non-overlapping of the factors, these methods yield an increased rank of the BMF decomposition, which is an undesired feature in low-rank approximation and data compression application

We addressed in [Miron21] the *binary matrix factorization* from a signal processing (*source separation*) perspective. We proposed a BMF approach based on a relaxation over the real field of the *Boolean matrix product*. Unlike the approach in [Zhang07], by using a *post-nonlinear mixture model*, the binary matrix product used by our method is equivalent to the *Boolean matrix product* in the exact binary case, and therefore, the correlation of the underlaying factors does no longer affect the algorithm performance. Another major contribution of [Miron21] is represented by the results on the uniqueness of the *Boolean* decomposition of a binary matrix; we provided a necessary condition, a sufficient condition and a necessary and sufficient condition for the uniqueness of the *Boolean* decomposition of a binary matrix.

To solve the BMF problem we proposed in [Miron21] a gradient descent algorithm with a multiplicative update rule, similar to the popular NMF algorithm introduced in [Lee99]. For the binary tensor framework, we proposed a more evolved algorithm for fitting the BTF model, based on an ADMM [1] procedure [Diop19]. We present next some of the results on BMF and BTF, with emphasis on the first one.

6.1.1 The Binary Matrix Factorization (BMF) model

The decomposition of a $N \times M$ matrix $\boldsymbol{X}$ with binary entries ($\boldsymbol{X}_{ij} \in \{0, 1\}$), into a sum of K (generally $K \ll \min\{M, N\}$) rank-1 terms:

$$\boldsymbol{X} = \boldsymbol{X}^{(1)} + \cdots + \boldsymbol{X}^{(K)}, \tag{6.1}$$

with $\boldsymbol{X}^{(k)} = \boldsymbol{w}_k \boldsymbol{h}_k^{\mathsf{T}}$, $(k = 1, \ldots, K)$ is generally known as *Binary Matrix Decomposition* or *Binary Matrix Factorization (BMF)*. The vectors $\boldsymbol{w}_k$ and $\boldsymbol{h}_k$ are oftenly termed as *factors*. The smallest K for which (6.1) holds exactly is called the *rank* of the decomposition.

If we adopt a *source separation* point of view, $\boldsymbol{X}$ represents the observed source mixtures (the measurements), while the rank-1 matrix $\boldsymbol{X}^{(k)}$ is the kth source, with the two dimensions corresponding to the two observed diversities; K is the number of sources that we seek to separate from the mixture.

Depending on the set over which this factorization is performed, two types of BMF and decomposition ranks can be defined:

- If the entries of $\boldsymbol{w}_k$ and $\boldsymbol{h}_k$ are restricted to the $\{0, 1\}$ set, the decomposition of (6.1) is generally known as the *binary* decomposition of $\boldsymbol{X}$ and the associated rank is the *binary rank* (see *e.g.*, [Watson16]), denoted $\text{rank}_{\{0,1\}}\{\boldsymbol{X}\}$. The *binary rank* can be interpreted as the minimum number of *disjoint* all-ones (rank-1) rectangles needed to cover all the 1's of $\boldsymbol{X}$ (after some row and column permutations).

[1] *Alternating Direction Method of Multipliers*

- If we keep the restriction of $\boldsymbol{w}_k$ and $\boldsymbol{h}_k$ to the $\{0,1\}$ set and replace the "+" in (6.1) by the logical OR operator "$\vee$", than eq. (6.1) can be seen as a decomposition of $\boldsymbol{X}$ over the *Boolean semiring* $(\mathbb{B}, \vee, \wedge)$. The symbol "$\wedge$" denotes the AND logical operator, which is the same as the real multiplication for binary values. This decomposition is oftenly termed as the *Boolean* factorization of the binary matrix $\boldsymbol{X}$, and the rank of the decomposition is known as the *Boolean rank* [Watts01] or the *Schein rank* [Kim82], and denoted $\text{rank}_{\mathbb{B}}\{\boldsymbol{X}\}$. The *Boolean rank* of $\boldsymbol{X}$ can be understood as the minimum number of all-ones (*not necessarily disjoint*) rectangles needed to cover all the 1's of $\boldsymbol{X}$ (after some row and column permutations). The problem of determining the *Boolean rank* of a binary matrix is NP-complete [Orlin77].

It is known [Watson16] that the *Boolean rank* of a binary matrix is smaller or equal to its *binary* rank:

$$\text{rank}_{\mathbb{B}}\{\boldsymbol{X}\} \leq \text{rank}_{\{0,1\}}\{\boldsymbol{X}\}. \tag{6.2}$$

We illustrate next, on a practical example, the utility of the *Boolean* factorization in data analysis. We consider a subset of the `UCI zoo dataset`[2], composed of a list of five animal species and five *Boolean*-valued attributes, as shown in Table 6.1.

	`airborne`	`backbone`	`toothed`	`aquatic`	`milk`
`crow`	**1**	**1**	0	0	0
`hawk`	**1**	**1**	0	0	0
`carp`	0	**1**	**1**	**1**	0
`dolphin`	0	**1**	**1**	**1**	**1**
`elephant`	0	**1**	**1**	0	**1**

Table 6.1: The considered subset of the UCI zoo dataset. Each row corresponds to a species and each column to an attribute.

The presence/absence of an attribute for a species is encoded by "1"/ "0", respectively. This results in a 5×5 binary-valued matrix $\boldsymbol{X}$, as shown in Table 6.1. The objective is to find classes of similar species, *i.e.,* species sharing the same attributes. This comes down to decomposing the binary attributes matrix into a sum of rank-1 binary terms $\boldsymbol{X}^{(k)} = \boldsymbol{w}_k \boldsymbol{h}_k^{\mathsf{T}}$. For each rank-1 term, $\boldsymbol{w}_k$ will have "1"s in the positions corresponding to the similar species and $\boldsymbol{h}_i$ will have "1"s in the positions corresponding to the shared attributes. For the data considered here, we obtain a rank-3 *Boolean* decomposition; the three rank-1 terms are highlighted in Table 6.1. Interestingly, one can see that, according to the decomposition, the dolphin belongs to two classes, as it has common features both with the `carp` and the `elephant`. This a reasonable result, as the dolphin is an aquatic mammal. This result is possible thanks to the fact that the *Boolean* decomposition allows overlapping supports of the rank-1 terms,

[2] Available on the UCI Machine learning repository website: `https://archive.ics.uci.edu/ml/datasets/Zoo` .

as illustrated in Table 6.1.

6.1.2 Uniqueness of the *Boolean* BMF

In the previous section we have seen that *Boolean* decomposition seems to present interesting features. We prove in this section a *necessary and sufficient* condition for the unique *Boolean* decomposition of a binary matrix, and some other uniqueness results. Consider the *Boolean decomposition* of a binary matrix $\boldsymbol{X}$ given by:

$$\boldsymbol{X} = \boldsymbol{X}^{(1)} \vee \cdots \vee \boldsymbol{X}^{(K)} = \bigvee_{k=1}^{K} \boldsymbol{X}^{(k)}, \tag{6.3}$$

with $\boldsymbol{X}^{(k)} = \boldsymbol{w}_k \boldsymbol{h}_k^\mathsf{T}$, and the entries of $\boldsymbol{w}_k$ and $\boldsymbol{h}_k$ restricted to the set $\{0, 1\}$.

Definition 6.1 (Full uniqueness)**.** *We say that the decomposition* (6.3) *is* fully unique *if, for any other set of K rank-1 binary matrices $\left\{\bar{\boldsymbol{X}}^{(1)}, \cdots, \bar{\boldsymbol{X}}^{(K)}\right\}$ satisfying* (6.3)*, we have $\boldsymbol{X}^{(k)} = \bar{\boldsymbol{X}}^{(k)}$, for all $k = 1, \cdots, K$, up to some permutation of the superscripts $(1, \cdots, K)$.*

This permutation of the superscripts is also known in the literature as the *order* indeterminacy. It is worth noting that, for the decomposition in binary terms there is no *scaling* indeterminacy, contrary to the real-value decompositions.

The condition for the full uniqueness of (6.3) as well as its proof, is based on an intermediary and more relaxed notion which is the *partial uniqueness*. Let us consider the *Boolean* decomposition of the rank-K binary matrix $\boldsymbol{X}$ given by (6.3), and re-write it as:

$$\boldsymbol{X} = \bigvee_{k \neq i} \boldsymbol{X}^{(k)} \vee \boldsymbol{X}^{(i)} = \boldsymbol{X}^{(\setminus i)} \vee \boldsymbol{X}^{(i)}, \tag{6.4}$$

where $\boldsymbol{X}^{(\setminus i)} = \bigvee_{k \neq i} \boldsymbol{X}^{(k)}$ is a binary matrix of *Boolean* rank $K - 1$.

Definition 6.2 (Partial uniqueness)**.** *We say that the decomposition* (6.4) *is* partially unique *with respect to $\boldsymbol{X}^{(i)}$ if, for a given $\boldsymbol{X}^{(\setminus i)}$, the only rank-1 binary matrix satisfying* (6.4) *is $\boldsymbol{X}^{(i)}$.*

We study next under which conditions the partial uniqueness of a *Boolean* decomposition can be achieved. First we prove a simple yet insightful *necessary* condition for partial uniqueness.

Theorem 6.1 (Partial uniqueness: a necessary condition [Miron21])**.** *The decomposition* (6.4) *is not partially unique with respect to $\boldsymbol{X}^{(i)} = \boldsymbol{w}_i \boldsymbol{h}_i^\mathsf{T}$ if there exists $j \in \{1, \cdots, K\} \setminus \{i\}$ such that* $\text{supp}\{\boldsymbol{w}_i\} \subseteq \text{supp}\{\boldsymbol{w}_j\}$ *or* $\text{supp}\{\boldsymbol{h}_i\} \subseteq \text{supp}\{\boldsymbol{h}_j\}$.

An illustration of this necessary condition is given on Figure 6.1. The gray and the dark gray rectangles represent the supports of $\boldsymbol{X}^{(j)}$ and $\boldsymbol{X}^{(i)}$, respectively, as defined in Theorem 6.1. One can observe that $\text{supp}\{\boldsymbol{w}_i\} \subset \text{supp}\{\boldsymbol{w}_j\}$. The dashed lines represent the bounds of three different possible support configurations for the rank-1 term $\boldsymbol{X}^{(i)}$, illustrating the non-uniqueness of this term in the decomposition. A key point for understanding the uniqueness results presented in this section is the fact that all these admissible configurations of $\boldsymbol{X}^{(i)}$ include the rank-1 term $\boldsymbol{w}_i(\boldsymbol{h}_i^{\not\subseteq})^\mathsf{T}$, which is outside the support of $\boldsymbol{X}^{(j)}$. This term represents the $\boldsymbol{X}^{(i)}$ with the minimum support, satisfying (6.4).

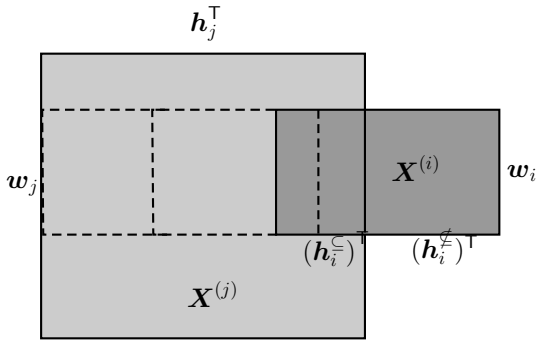


Figure 6.1: Illustration of the necessary condition for partial uniqueness given by Theorem 6.1

Based on these observations, we prove next a necessary and sufficient condition for partial uniqueness.

Theorem 6.2 (Partial uniqueness: a necessary and sufficient condition [Miron21])**.** *Consider the decomposition* (6.4) *and define the following subsets:*

- $\Omega_{\boldsymbol{w}_i}$ *: the smallest subset of* $\{1, \cdots, K\} \setminus \{i\}$ *such that* $\text{supp}\{\boldsymbol{w}_i\} \subseteq \bigcup\limits_{k \in \Omega_{\boldsymbol{w}_i}} \text{supp}\{\boldsymbol{w}_k\}$
- $\Omega_{\boldsymbol{h}_i}$ *: the smallest subset of* $\{1, \cdots, K\} \setminus \{i\}$ *such that* $\text{supp}\{\boldsymbol{h}_i\} \subseteq \bigcup\limits_{k \in \Omega_{\boldsymbol{h}_i}} \text{supp}\{\boldsymbol{h}_k\}$

The decomposition is partially unique with respect to $\boldsymbol{X}^{(i)} = \boldsymbol{w}_i\boldsymbol{h}_i^\mathsf{T}$ *iff:* $\bigcap\limits_{k \in \Omega_{\boldsymbol{h}_i}} \text{supp}\{\boldsymbol{w}_k\} = \emptyset$ *and* $\bigcap\limits_{k \in \Omega_{\boldsymbol{w}_i}} \text{supp}\{\boldsymbol{h}_k\} = \emptyset$.

Figure 6.2 illustrates the partial uniqueness conditions in the case of a rank-3 decomposition; the grey rectangles represent the supports of the three sources. In Figure 6.2 (a) it can be seen that the partial uniqueness condition with respect to $\boldsymbol{X}^{(2)}$ is not satisfied. Indeed, $\text{supp}\{\boldsymbol{h}_2\} \subseteq (\text{supp}\{\boldsymbol{h}_1\} \cup \text{supp}\{\boldsymbol{h}_3\})$ and $\text{supp}\{\boldsymbol{w}_1\} \cap \text{supp}\{\boldsymbol{w}_3\} \neq \emptyset$. Thus, by subtracting, for example, the hatched part from the support of $\boldsymbol{X}^{(2)}$, another admissible rank-1 term $\bar{\boldsymbol{X}}^{(2)} \neq \boldsymbol{X}^{(2)}$ can be obtained. On Figure 6.2 (b), $\text{supp}\{\boldsymbol{w}_1\} \cap \text{supp}\{\boldsymbol{w}_3\} = \emptyset$, and therefore, given $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(3)}$, the only possible rank-1 term satisfying the decomposition is $\boldsymbol{X}^{(2)}$.

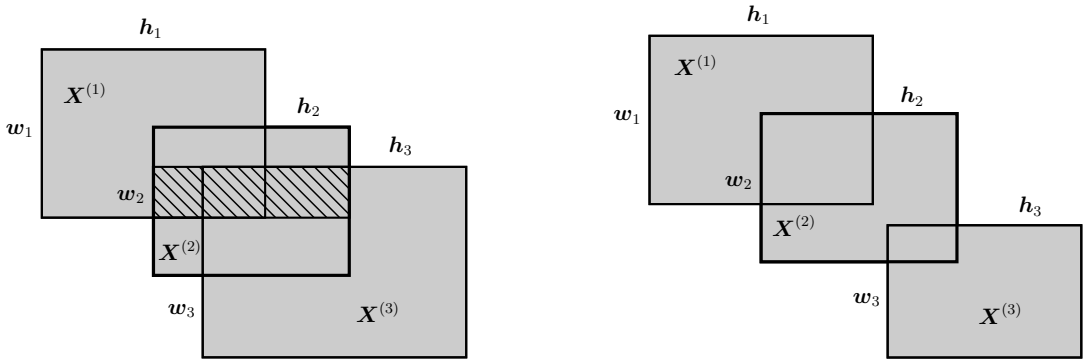


Figure 6.2: An illustration of partial uniqueness condition of Theorem 6.2, for a rank-3 decomposition. The grey rectangles represent the supports of the three rank-1 terms.
(a) *Partial non-uniqueness* with respect to $\boldsymbol{X}^{(2)}$: by subtracting the hatched part from the support of $\boldsymbol{X}^{(2)}$ we obtain another rank-1 term $\bar{\boldsymbol{X}}^{(2)}$, that satisfies $\boldsymbol{X} = \boldsymbol{X}^{(1)} \vee \bar{\boldsymbol{X}}^{(2)} \vee \boldsymbol{X}^{(3)}$.
(b) *Partial uniqueness* with respect to $\boldsymbol{X}^{(2)}$: for fixed $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(3)}$, $\boldsymbol{X}^{(2)}$ is the only rank-1 binary matrix satisfying $\boldsymbol{X} = \boldsymbol{X}^{(1)} \vee \boldsymbol{X}^{(2)} \vee \boldsymbol{X}^{(3)}$.

Theorem 6.3 (Full uniqueness : necessary and sufficient condition [Miron21])**.** *Let* $\boldsymbol{X}$ *be a binary matrix of* Boolean *rank* K*, following the* Boolean *decomposition given by* (6.3)*. The decomposition is* fully unique *(as defined in Definition 6.1) iff the partial uniqueness with respect to* $\boldsymbol{X}^{(i)}$ *is satisfied for all* $i = 1, \ldots, K$*.*

A direct consequence of Theorem 6.2 and Theorem 6.3 is the following sufficient condition which is of practical interest.

Corollary 6.1 (Full uniqueness: a sufficient condition [Miron21])**.** *The decomposition* (6.3) *is* fully unique *if its rank-1 terms* $\boldsymbol{X}^{(i)} = \boldsymbol{w}_i \boldsymbol{h}_i^{\mathsf{T}}$ *satisfy* $\text{supp}\{\boldsymbol{w}_i\} \nsubseteq \bigcup_{k \neq i} \text{supp}\{\boldsymbol{w}_k\}$ *and* $\text{supp}\{\boldsymbol{h}_i\} \nsubseteq \bigcup_{k \neq i} \text{supp}\{\boldsymbol{w}_k\}$*, for all* $i = 1, \ldots, K$*.*

Roughly speaking, the uniqueness results presented in this section express the fact that a *Boolean* decomposition with rank-1 terms having overlapping supports (correlated sources) has less chances of being unique than a decomposition with non-overlapping supports (non-correlated sources), and the "probability of non-uniqueness" increases with the correlation degree of the sources.

6.1.3 A post-nonlinear mixture model approach to *Boolean* BMF

Consider the *Boolean* decomposition of a binary matrix given by (6.3). By arranging the factors $\boldsymbol{w}_k$ and $\boldsymbol{h}_k$, $k = 1, \ldots, K$ of the rank-1 terms on the columns of matrices $\boldsymbol{W} = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_K]$ and $\boldsymbol{H} = [\boldsymbol{h}_1, \ldots, \boldsymbol{h}_K]$, of respective sizes $N \times K$ and $M \times K$, the *Boolean* BMF can be re-written as:

$$\boldsymbol{X} = \boldsymbol{W} \diamond \boldsymbol{H}^{\mathsf{T}}. \tag{6.5}$$

"$\diamond$" is called *Boolean matrix product* and is defined as $\boldsymbol{X}_{ij} = \bigvee_{k=1}^{K} (\boldsymbol{W}_{ik} \wedge \boldsymbol{H}_{jk})$, where "$\vee$" and "$\wedge$" are OR and AND logical operators, respectively. Until now we only considered the *exact* BMF, *i.e.,* $\boldsymbol{X}$ can be exactly recovered from the K rank-1 terms. In the sequel we suppose that the rank-K model can be corrupted by noise/errors. Depending on the set over which the decomposition of the binary matrix is performed, this noise term can take different forms. For the decomposition of $\boldsymbol{X}$ in binary terms, and for the *Boolean* BMF in particular, it is natural to assume that the entries of the "noise" matrix take values in the $\{0, 1\}$ set, and that we have

$$\boldsymbol{X} = \boldsymbol{W} \diamond \boldsymbol{H}^{\mathsf{T}} \veebar \boldsymbol{N}, \tag{6.6}$$

with "$\veebar$", the element-wise XOR logical operator. Intuitively, this means that the 1's of $\boldsymbol{N}$ change the corresponding entries of $\boldsymbol{W} \diamond \boldsymbol{H}^{\mathsf{T}}$ regardless of their respective values ("0" or "1").

Thus, the inverse problem to be solved in order to achieve the *Boolean* BMF is the following:

$$\{\boldsymbol{W}, \boldsymbol{H}\} = \arg \min_{\boldsymbol{W},\boldsymbol{H}\in\{0,1\}} \left\| \boldsymbol{X} - \boldsymbol{W} \diamond \boldsymbol{H}^{\mathsf{T}} \right\|_{\mathsf{F}}^2 = \arg \min_{\boldsymbol{w}_k,\boldsymbol{h}_k\in\{0,1\}} \left\| \boldsymbol{X} - \bigvee_{k=1}^{K} \boldsymbol{w}_k \boldsymbol{h}_k^{\mathsf{T}} \right\|_{\mathsf{F}}^2. \tag{6.7}$$

It is worth noting that, for a binary-valued matrix, the *Frobenius* norm and the "entry-wise" norms ℓ_0 and ℓ_1 are equivalent.

We showed in [Miron21] that most existing methods for the *Boolean* BMF yield good results only in the case of uncorrelated sources, because they are solving approximate, relaxed versions of the inverse problem (6.7). We propose in this section a BMF problem formulation based on a *post-nonlinear* mixture model which is equivalent to (6.7) when the matrices $\boldsymbol{W}$ and $\boldsymbol{H}$ are exactly binary. An algorithm for estimating $\boldsymbol{W}$ and $\boldsymbol{H}$ based on this formulation is also introduced. Therefore, this new approach is expected to provide good results even in the case of highly correlated sources.

The proposed Post NonLinear Penalty Function algorithm (PNL-PF) We proposed to replace the *Boolean* matrix product by the real matrix product with binary constraints on $\boldsymbol{W}$ and $\boldsymbol{H}$ followed by a nonlinear function that guarantees the binarity of the reconstructed data. In theory, we desire a peace-wise linear function such that:

$$\Phi_{th}(x) = \begin{cases} 1 & \text{for } x > (1-\delta) \\ x & \text{for } \delta \leq \frac{x-\delta}{1-2\delta} \leq 1-\delta \\ 0 & \text{for } x < \delta \end{cases}, \text{ with } 0 \leq \delta < 0.5. \tag{6.8}$$

Thus, the inverse problem we seek to solve, can be formulated as:

$$\{\boldsymbol{W}, \boldsymbol{H}\} = \arg \min_{\boldsymbol{W},\boldsymbol{H}\in\{0,1\}} \|\boldsymbol{X} - \Phi_{th}(\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}})\|_{\mathsf{F}}^2. \tag{6.9}$$

However, for algorithmic convenience, in this paper we approximate $\Phi_{th}(x)$ by the sigmoid function $\Phi(x) = \frac{1}{1+e^{-\gamma(x-0.5)}}$ as it is continuously differentiable; the parameter γ allows to adjust the slope of Φ.

Figure 6.3 plots $\Phi_{th}(x)$ for $\delta = 0.1$ and $\Phi(x)$ for $\gamma = 1, 5, 10, 100$. One can see that, as the value of γ increases, $\Phi(x)$ fits better $\Phi_{th}(x)$. However, for algorithmic reasons, the value of γ must not be big in order to facilitate the parameters' variations during the iterations. In practice, a good choice for γ is a value between 1 and 10, depending on the application.

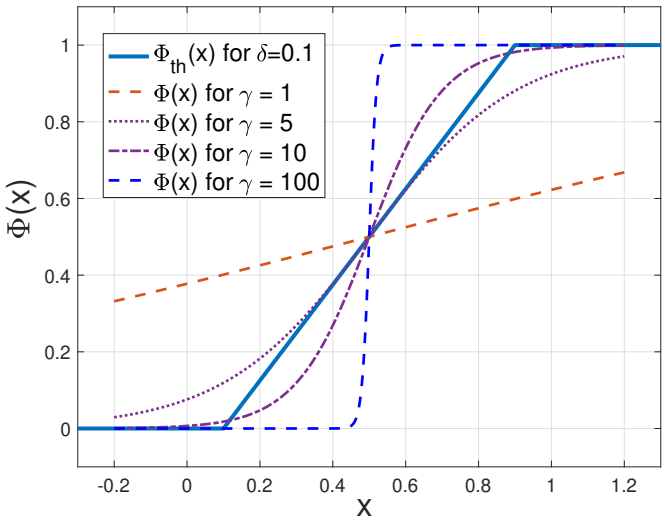


Figure 6.3: Function $\Phi_{th}(x)$ and its sigmoid approximations for various values of γ

To solve (6.9), we propose a gradient descent algorithm with a multiplicative update rule. The proposed algorithm minimizes the objective function (6.10); the fundamental difference compared to [Zhang07] is the integration of the non-linear function Φ to guarantee $\Phi(\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}) \approx \boldsymbol{W} \diamond \boldsymbol{H}^{\mathsf{T}}$:

$$G(\boldsymbol{W}, \boldsymbol{H}) = \frac{1}{2}\sum_{i,j}\left(\boldsymbol{X}_{ij} - \Phi(\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}})_{ij}\right)^2 + \frac{1}{2}\lambda\sum_{i,k}\left(\boldsymbol{W}_{ik}^2 - \boldsymbol{W}_{ik}\right)^2 + \frac{1}{2}\lambda\sum_{j,k}\left(\boldsymbol{H}_{jk}^2 - \boldsymbol{H}_{jk}\right)^2. \tag{6.10}$$

Algorithm 2: : Post NonLinear Penalty Function algorithm (PNL-PF)

Input: $\boldsymbol{X}$, K, Nb_{iter}, λ, ε, γ

Output: $\boldsymbol{W}$, $\boldsymbol{H}$

STEP 1: Initializations: $\boldsymbol{W} \leftarrow rand(N, K)$, $\boldsymbol{H} \leftarrow rand(N, K)$, $iter = 0$

STEP 2: Normalization

$\boldsymbol{W} \leftarrow \boldsymbol{W}\boldsymbol{D}_W^{-1/2}\boldsymbol{D}_H^{1/2}$, with $\boldsymbol{D}_W = diag\left(\max(\boldsymbol{w}_1), \max(\boldsymbol{w}_2), \cdots, \max(\boldsymbol{w}_K)\right)$

$\boldsymbol{H} \leftarrow \boldsymbol{H}\boldsymbol{D}_H^{-1/2}\boldsymbol{D}_W^{1/2}$, with $\boldsymbol{D}_H = diag\left(\max(\boldsymbol{h}_1), \max(\boldsymbol{h}_2), \cdots, \max(\boldsymbol{h}_K)\right)$

STEP 3: Updates of $\boldsymbol{W}$, $\boldsymbol{H}$

$iter \leftarrow iter + 1$

$$\boldsymbol{H} \leftarrow \boldsymbol{H} \boxdot \frac{\gamma((\boldsymbol{X} \boxdot \Omega(\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}))^{\mathsf{T}} \cdot \boldsymbol{W}) + 3\lambda\boldsymbol{H}^2}{\gamma(\Psi(\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}})^{\mathsf{T}} \cdot \boldsymbol{W}) + 2\lambda\boldsymbol{H}^3 + \lambda\boldsymbol{H}}$$

$$\boldsymbol{W} \leftarrow \boldsymbol{W} \boxdot \frac{\gamma((\boldsymbol{X} \boxdot \Omega(\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}})) \cdot \boldsymbol{H}) + 3\lambda\boldsymbol{W}^2}{\gamma(\Psi(\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}) \cdot \boldsymbol{H}) + 2\lambda\boldsymbol{W}^3 + \lambda\boldsymbol{W}}$$

STEP 4: Stop criterion

if $iter \geq Nb_{iter}$ or $G(\boldsymbol{W}, \boldsymbol{H}) < \varepsilon$ **then**

  break

else

  return to *STEP 3*

end if

In *STEP 3*, $\boldsymbol{W}$ and $\boldsymbol{H}$ are normalized in order to confine the values of $\boldsymbol{W}_{ij}$ and $\boldsymbol{H}_{ij}$ within the interval $[0, 1]$. Ω and Ψ are two element-wise functions that associate to each entry $\boldsymbol{Z}_{ij}$ of a matrix $\boldsymbol{Z}$ (given in argument) the real values $\frac{e^{-\gamma \cdot \left(\boldsymbol{Z}_{ij}-0.5\right)}}{\left(1+e^{-\gamma \cdot \left(\boldsymbol{Z}_{ij}-0.5\right)}\right)^2}$ and $\frac{e^{-\gamma \cdot \left(\boldsymbol{Z}_{ij}-0.5\right)}}{\left(1+e^{-\gamma \cdot \left(\boldsymbol{Z}_{ij}-0.5\right)}\right)^3}$, respectively. In Algorithm 2, the power operation for a matrix $\boldsymbol{Z}$ is defined as $\boldsymbol{Z}^2 = \boldsymbol{Z} \boxdot \boldsymbol{Z}$, and the matrix division is performed element-wise. The derivation of the update rules for $\boldsymbol{W}$ and $\boldsymbol{H}$ are detailed in [Miron21].

6.1.4 A post-nonlinear mixture model approach to *Boolean* BTF

Using the same post-nonlinear function idea as for BMF, we developed in [Diop19] a BTF algorithm based on an Alternating Optimization (AO)-ADMM procedure. Unlike the BMF algorithm described in the previous section, this new BTF algorithm presents convergence guarantees under mild conditions, as explained in [Huang16].

Consider a three-way binary data array (tensor) $\boldsymbol{\mathcal{X}}$ of size $N \times M \times P$ such that its elements $\boldsymbol{\mathcal{X}}_{nmp} \in \{0, 1\}$ (with $n = 1, \ldots, N$, $m = 1, \ldots, M$ and $p = 1, \ldots, P$) can be expressed as:

$$\boldsymbol{\mathcal{X}}_{nmp} = \bigvee_{k=1}^{K} (w_{nk} \wedge h_{mk} \wedge v_{pk}), \tag{6.11}$$

with $w_{nk}, h_{mk}, v_{pk} \in \{0,1\}$. Equation (6.11) expresses a third-order *Boolean Canonical Polyadic Decomposition (BCPD)* of rank K. As $w_{nk}, h_{mk}, v_{pk} \in \{0,1\}$, the logical "AND" operator in (6.11) can be equivalently replaced by the classical real numbers product, and thus $\boldsymbol{\mathcal{X}}_{nmp} = \bigvee_{k=1}^{K} w_{nk} h_{mk} v_{pk}$. By regrouping the elements w_{nk}, h_{mk} and v_{pk} on the columns of matrices $\boldsymbol{W} = [\boldsymbol{w}_1 \dots \boldsymbol{w}_K]$ $(N \times K)$, $\boldsymbol{H} = [\boldsymbol{h}_1 \dots \boldsymbol{h}_K]$ $(M \times K)$ and $\boldsymbol{V} = [\boldsymbol{v}_1 \dots \boldsymbol{v}_K]$ $(P \times K)$, respectively, the BCPD (6.11) can be expressed as:

$$\boldsymbol{\mathcal{X}} = \bigvee_{k=1}^{K} (\boldsymbol{w}_k \circ \boldsymbol{h}_k \circ \boldsymbol{v}_k) = [\![\boldsymbol{W}, \boldsymbol{H}, \boldsymbol{V}]\!], \tag{6.12}$$

where "$\circ$" denotes the vector outer product and where the logical operation "$\vee$" is performed element-wise. Thus, in order to estimate the BCPD of $\boldsymbol{\mathcal{X}}$, one must solve the following inverse problem:

$$\{\hat{\boldsymbol{W}}, \hat{\boldsymbol{H}}, \hat{\boldsymbol{V}}\} = \arg \min_{\boldsymbol{W},\boldsymbol{H},\boldsymbol{V} \in \{0,1\}} \left\| \boldsymbol{\mathcal{X}} - \bigvee_{k=1}^{K} \boldsymbol{w}_k \circ \boldsymbol{h}_k \circ \boldsymbol{v}_k \right\|_{\mathsf{F}}^2. \tag{6.13}$$

A classical way to perform the canonical polyadic decomposition in the real-valued case is to alternatingly estimate the three loading matrices using the three n-mode unfoldings of $\boldsymbol{\mathcal{X}}$ [Kolda09]. A similar strategy can be used in the Boolean case, based on the three unfoldings hereafter:

$$\boldsymbol{X}_{(1)} = \boldsymbol{W} \diamond (\boldsymbol{V} \odot \boldsymbol{H})^{\mathsf{T}}, \tag{6.14}$$

$$\boldsymbol{X}_{(2)} = \boldsymbol{H} \diamond (\boldsymbol{V} \odot \boldsymbol{W})^{\mathsf{T}}, \tag{6.15}$$

$$\boldsymbol{X}_{(3)} = \boldsymbol{V} \diamond (\boldsymbol{H} \odot \boldsymbol{W})^{\mathsf{T}}. \tag{6.16}$$

We base our algorithm for solving the BCPD problem (6.13) on a *post-nonlinear* mixture model approach, similar to the one used for the matrix case. Instead of solving directly (6.13), we solve a relaxed version of it:

$$\{\hat{\boldsymbol{W}}, \hat{\boldsymbol{H}}, \hat{\boldsymbol{V}}\} = \arg \min_{\boldsymbol{W},\boldsymbol{H},\boldsymbol{V} \in \{0,1\}} \left\| \boldsymbol{\mathcal{X}} - \Phi\left(\sum_{k=1}^{K} \boldsymbol{w}_k \circ \boldsymbol{h}_k \circ \boldsymbol{v}_k \right) \right\|_{\mathsf{F}}^2, \tag{6.17}$$

where $\Phi(x)$ is the sigmoid function in Figure 6.3, applied element-wise. The proposed *AO-ADMM* algorithm for solving (6.17) can be summarized as follows:

Repeat

$$\min_{\boldsymbol{W},\bar{\boldsymbol{W}}} \frac{1}{2} \left\| \boldsymbol{X}_{(1)} - \Phi\left(\bar{\boldsymbol{W}} (\boldsymbol{V} \odot \boldsymbol{H})^{\mathsf{T}} \right) \right\|_{\mathsf{F}}^2 + \frac{\lambda}{2} \left\| \boldsymbol{W} - \boldsymbol{W} * \boldsymbol{W} \right\|_{\mathsf{F}}^2$$

subject to $\boldsymbol{W} = \bar{\boldsymbol{W}}$

$$\min_{\boldsymbol{H},\bar{\boldsymbol{H}}} \frac{1}{2}\left\|\boldsymbol{X}_{(2)} - \Phi\left(\bar{\boldsymbol{H}}\left(\boldsymbol{V}\odot\boldsymbol{W}\right)^{\mathsf{T}}\right)\right\|_{\mathsf{F}}^{2} + \frac{\lambda}{2}\left\|\boldsymbol{H} - \boldsymbol{H} * \boldsymbol{H}\right\|_{\mathsf{F}}^{2}$$
$$\text{subject to } \boldsymbol{H} = \bar{\boldsymbol{H}}$$

$$\min_{\boldsymbol{V},\bar{\boldsymbol{V}}} \frac{1}{2}\left\|\boldsymbol{X}_{(3)} - \Phi\left(\bar{\boldsymbol{V}}\left(\boldsymbol{H}\odot\boldsymbol{W}\right)^{\mathsf{T}}\right)\right\|_{\mathsf{F}}^{2} + \frac{\lambda}{2}\left\|\boldsymbol{V} - \boldsymbol{V} * \boldsymbol{V}\right\|_{\mathsf{F}}^{2}$$
$$\text{subject to } \boldsymbol{V} = \bar{\boldsymbol{V}}$$

until convergence.

The second term in the expressions to minimize is used to constrain the entries of $\boldsymbol{W}, \boldsymbol{H}, \boldsymbol{W}$ to binarity. Using the results of [Huang16], update rules can be obtained for the three minimization problems. For example, for the update of $\boldsymbol{W}$, the following expressions are obtained:

$$\bar{\boldsymbol{W}} \leftarrow \arg\min_{\bar{\boldsymbol{W}}} \frac{1}{2}\left\|\boldsymbol{X}_{(1)} - \Phi\left(\bar{\boldsymbol{W}}\left(\boldsymbol{V}\odot\boldsymbol{H}\right)^{\mathsf{T}}\right)\right\|_{\mathsf{F}}^{2} + \frac{\rho}{2}\left\|\boldsymbol{W} - \bar{\boldsymbol{W}} + \boldsymbol{A}\right\|_{\mathsf{F}}^{2}$$
$$\boldsymbol{W} \leftarrow \arg\min_{\boldsymbol{W}} \frac{\lambda}{2}\left\|\boldsymbol{W} - \boldsymbol{W} * \boldsymbol{W}\right\|_{\mathsf{F}}^{2} + \frac{\rho}{2}\left\|\boldsymbol{W} - \bar{\boldsymbol{W}} + \boldsymbol{A}\right\|_{\mathsf{F}}^{2}$$
$$\boldsymbol{A} \leftarrow \boldsymbol{A} + \boldsymbol{W} - \bar{\boldsymbol{W}},$$

with ρ, a regularization parameter. The two minimization problems below are solved by gradient descent steps. The resulting algorithm, that we called *Boolean Tensor - ADMM (BT-ADMM)*, is explicitly given in [Diop19].

6.1.5 Performance evaluation on synthetic and real data

We briefly illustrate hereafter the performance of the proposed BMF approach on synthetic and real data; the conclusions are similar for the BTF method. More results can be found in [Miron21, Diop19].

Evaluation on synthetic data We study in the following experiment the behavior of our approach in the presence of XOR binary noise. We simulated the following model: $\boldsymbol{X} \veebar \boldsymbol{N}$, where the noise matrix $\boldsymbol{N}$ is a random binary matrix that follows a Bernouilli distribution of parameter b, and $\boldsymbol{X} = \boldsymbol{W} \diamond \boldsymbol{H}^{\mathsf{T}}$, is a rank-$K$ binary matrix. We computed the average reconstruction error for $\boldsymbol{X}$, $\text{Err}_{\boldsymbol{X}} = \frac{\|\boldsymbol{X}-\hat{\boldsymbol{X}}\|_{\mathsf{F}}^{2}}{MN}$ (where $\hat{\boldsymbol{X}}$ is the reconstructed matrix) over 40 trials, as a function of the added noise rate b. $\boldsymbol{W}$ and $\boldsymbol{H}$ were generated according to a Bernoulli distribution of parameter p. Thus, we designed an experiment to compare the performance of the proposed PNL-PF algorithm with state-of-the-art methods, in the case of $K = 3$ "sparse" sources with $p = 0.25$ (which corresponds to a 1's to 0's ratio of ≈ 0.0625 for each rank-1 term). We plotted on Figure 6.4 the reconstruction error $\text{Err}_{\boldsymbol{X}}$ with respect to the noise rate b, for PNL-PF, NMF [Lee99], PF [Zhang07], FC [Belohlavek10], and

ASSO [Miettinen08]. One can see that, in the noiseless case ($b = 0$), all algorithms estimate

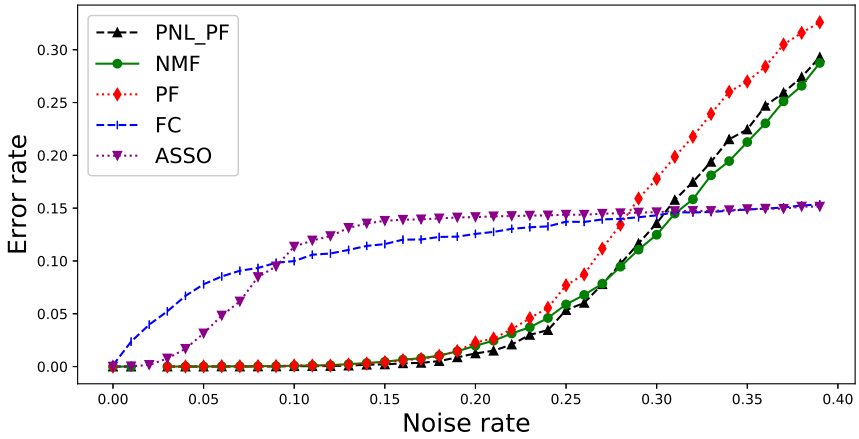


Figure 6.4: Comparison of the reconstruction error for PNL-PF with state-of -the-art algorithms, in the case of a rank-3 "sparse" binary matrix $\boldsymbol{X}$, for increasing additive XOR noise rates.

$\boldsymbol{X}$ exactly. When the noise power increases, PNL-PF, PF and NMF still yield accurate estimations, while the two greedy algorithms (FC and ASSO) completely fail to reconstruct $\boldsymbol{X}$. This is an expected behavior as the greedy algorithms were designed for exact binary factorizations, and do not perform well for low-rank approximation problems. Moreover, one can observe that the proposed PNL-PF algorithm exhibits performances very similar to NMF; this is somewhat surprising, as NMF is not constraint to binary values, and therefore one may expect NMF to better fit the low-rank decomposition of $\boldsymbol{X}$. However, for high noise rates (> 0.3), all algorithms fail, as the low-rank structure of $\boldsymbol{X}$ is destroyed by the non-linear addition of the XOR noise.

The second experiment is very similar to the second one, with the difference that matrices $\boldsymbol{W}$ and $\boldsymbol{H}$ were generated according to a Bernoulli distribution of parameter $p = 0.6$, which implies much more "dense" rank-1 terms (with a 1's to 0's ratio of ≈ 0.36). It can be seen on Figure 6.5 that, not only the performances of ASSO et FC degrade further, but also the gap between PNL-PF and PF increases. The reason for this is the strong overlapping of the rank-1 term supports (*i.e.,* the strong correlation of the sources). In this case, the algorithms based on the classical matrix product of binary matrices, such as PF, no longer produce good results. Meanwhile, the proposed PNL-PF method still achieves good performances, close to NMF, thanks to its underlying post-nonlinear mixture model.

The results of these experiments also illustrates the fact that, for a binary-valued matrix, its binary decomposition (associated with the PF, FC and ASSO algorithms) requires a higher rank to achieve the same approximation error as compared to its *Boolean* decomposition (performed by PNL-PF).

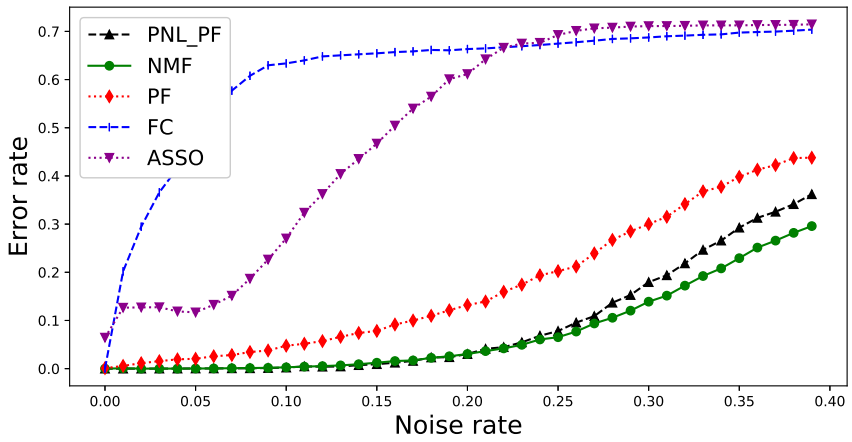


Figure 6.5: Comparison of the reconstruction error for PNL-PF with state-of -the-art algorithms, in the case a rank-3 "dense" binary matrix $\boldsymbol{X}$, for increasing additive XOR noise rates.

Application to real/ realistic data We analyze next the `UCI zoo dataset`[3] that was partially seen in Section 6.1.1. This dataset consists of 101 animal species from a zoo. There are 16 variables with various traits to describe the animals. We considered only the 15 *Boolean* valued variables: `airborne, feathers, breathes, backbone, tail, eggs, predator, aquatic, fins, toothed, milk, catsize, hair, venomous, domestic`, in this order. This results in a 101×15 binary-valued matrix that can be seen in Figure 6.6. We applied our PNL-PF algorithm to the data matrix from Figure 6.6 for different decomposition ranks. The highest rank for which the decomposition result was stable (the same result was obtained each time over 10 consecutive runs, with random initializations) is rank 3. The result of the rank-3 *Boolean* decomposition by PNL-PF is plotted in Figure 6.7(a). We obtain three animal classes sharing the following common features:

- features class #1: `airborne, feathers, breathes, backbone, tail, eggs`;
- features class #2: `backbone, tail, eggs, predator, aquatic, fins, toothed`;
- features class #3: `breathes, backbone, tail, toothed, milk, catsize, hair`.

An animal from an identified class has most of the class features, but not necessarily all of them. Generally speaking, the first class represents the "birds", the second one, the "fishes" and the third one, the "mammals":

It can be seen that there is an overlapping between class #2 and class #3; they have the following animals in common: `dolphin, sea lion, seal, platypus, porpoise`. This is a reasonable result as these particular animals present feature belonging to both classes. Six other animals (`clam, flea, scorpion, slug, termite, worm`) were not included in any of

[3] `https://archive.ics.uci.edu/ml/datasets/Zoo.`

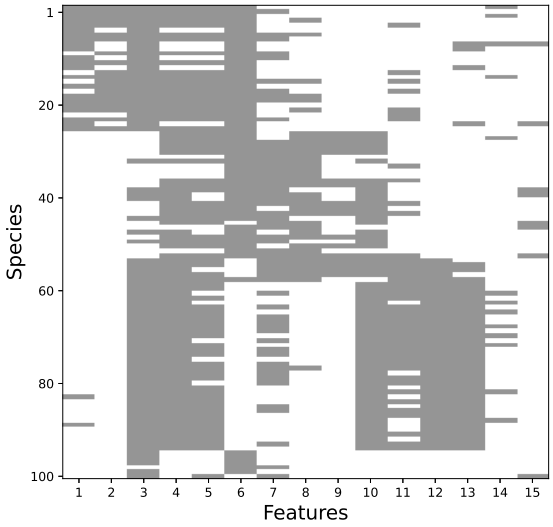


Figure 6.6: The initial zoo dataset

these three classes because they do not present sufficient characteristics of any of them. Also, one can see that the algorithm did not take into account the last two attributes (`venomous`, `domestic`), as they were not discriminant for this rank-3 decomposition.The results for the three other binary algorithms are also given in Figure 6.7. As expected, PF yields a similar result, but with smaller overlapping between the last two terms. The results for FC and ASSO are more difficult to interpret; moreover, they exclude additional animals from the classification.

6.2 Real-time unmixing of hyperpectral images

This section presents our contributions to the problem of sequential (or on-line) unmixing of hyperspectral Near InfraRed (NIR) images acquired by a *pushbroom* imager [Lu14], by means of Non-negative Matrix Factorization (NMF)-like approaches. This problem is encountered *e.g.,* in real-time industrial systems, for product quality control applications. In particular, the approach proposed in this paper is part of a project that aims at predicting and classifying in real-time, the rendering quality of pieces of wood in an industrial process. This is to be done by *on-line* unmixing method of hyperspectral images of wood surface, which presents several advantages over the off-line methods: *i)* it is well-adapted to real-time data processing for on-line industrial acquisition systems; *ii)* it allows to alleviate computational burden and reduce memory requirements for big hyperspectral data cubes; *iii)* it permits to track the spatial/time variability of the endmembers in a hyperspectral imaging application.

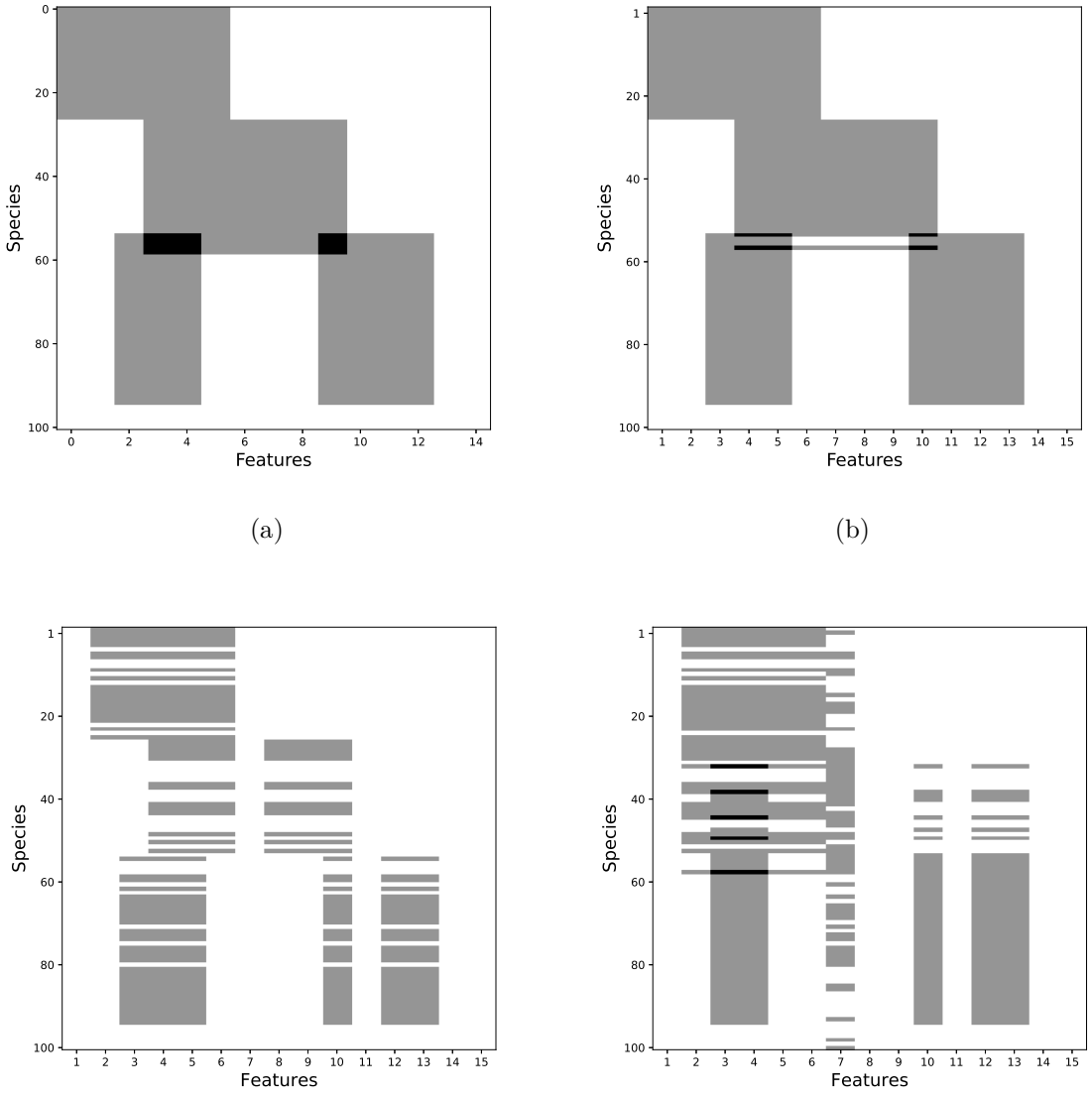


Figure 6.7: Results of the rank-3 decomposition of the zoo dataset from Figure 6.6: (a) PNL-PF algorithm, (b) PF algorithm, (c) FC algorithm, (d) ASSO algorithm.

6.2.1 Pushbroom acquisition scheme

In pushbroom imaging systems, hyperspectral data cubes are acquired slice by slice, sequentially in time. Each slice is an image, characterized by a spectral dimension and a spatial dimension (one line of the scene). Figure 6.8 illustrates the test bench that we consider for the real-time characterization of the wood material. Hyperspectral images are acquired line by line, as the pieces of wood are carried under the imager, via a conveyor.

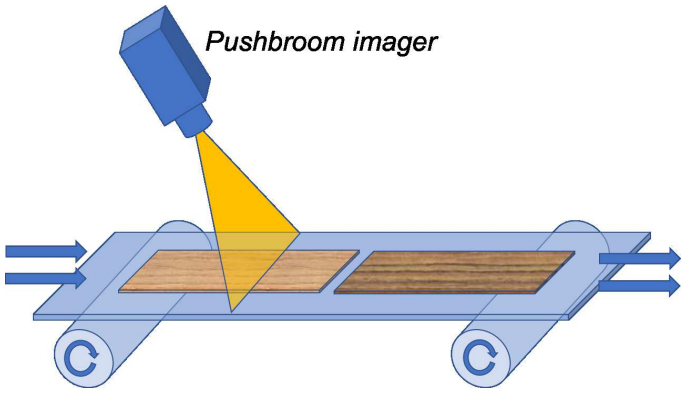


Figure 6.8: Acquisition of hyperspectral images of wood

The stream of spectral-spatial data arrays is then stacked to form the hyperspectral data cube. For each acquisition time k ($k = 1, ..., K$), the new slice (represented by a dotted line in Figure 6.9) is a matrix of dimensions $L \times P$, where L denotes the spectral dimension (wavelengths) and P, the across track spatial dimension (one line of the scene).

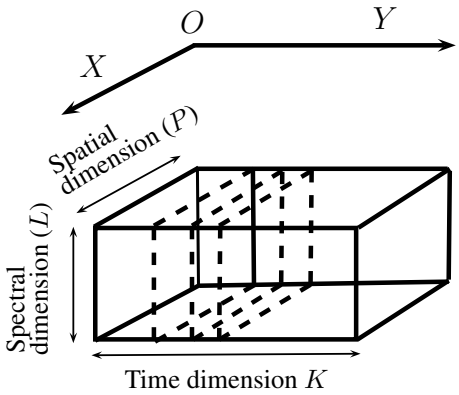


Figure 6.9: Data structure for pushbroom acquisition

The goal of the on-line blind hyperspectral unmixing is to produce real-time estimates of the endmember (source) and abundance matrices for each new incoming slice, at the pushbroom system acquisition rate.

6.2.2 On-line NMF methods

On-line NMF algorithms sequentially update the endmembers and abundances as the data size increases. In that respect, they are perfectly adapted to the processing of hyperspectral data streaming as they allow to maintain a low and controlled computational complexity. These algorithms can be gathered into two main categories, depending on the considered assumptions on the endmembers. In [Mairal10, Wang11, Guan12, Zhao16], the endmembers do not vary from one sample to another, while in [Bucak09, Lefèvre11, Zhou11, Wang13a, Wu14, Nus18], the endmembers may evolve between successive samples. In particular, the Incremental NMF (INMF) [Bucak09] considers that the endmembers evolve slowly between two consecutive acquisitions; this is now the most widely used assumption adopted in on-line NMF algorithms. In the context of hyperspectral unmixing, it allows to account for the so-called spectral variability. However, as in the off-line case, the uniqueness of the solution is not guaranteed, which led to the development of on-line regularized NMF such as [Zhou11, Wang13a, Wu14].

Main contributions We proposed two on-line unmixing algorithms, with two different regularizations, specially designed for pushbroom hyperspectral imaging systems, namely the on-line Minimum Volume Regularization NMF (OMVR-NMF) [Nus18] and the on-line Minimum Dispersion Regularization ADMM (OMDR-ADMM) [Nus20]. The Minimum Dispersion Regularization (MDR) can be seen as a surrogate for the minimum volume regularization; it aims at minimizing the sum of the distances between each endmember and the endmembers' centroid, which indirectly minimizes the volume of the convex hull determined by the endmembers. Although this regularization could be considered as a rough one, it has some interesting properties, in particular for the considered on-line context:

- Using MDR results in lower computational complexity compared to MVR: only matrix multiplications are required for MDR while MVR needs matrix inversions (required in OMVR-NMF algorithm). This is of upmost importance in the context of fast on-line unmixing.
- Using MDR makes the endmemebers estimation subproblem convex and thus, offers interesting optimization properties by enabling explicit updates of the parameters.
- Using MDR results in faster convergence of the algorithm compared to the non-convex regularization approach.
- Using MDR enables preservation of the decomposition rank for a wide range of the regularization hyperparameter, which is essential for the correct performance of the proposed approach.

We will only present hereafter the OMDR-ADMM algorithm as we showed that it outperforms OMVR-NMF in terms of reconstruction accuracy and convergence rate.

6.2.3 Data model for the pushbroom acquisition scheme

The principle of the proposed on-line method is to alternatively update the endmember and abundance matrices estimated at time instant k when a new sample (slice) arrives at time instant $k+1$. One way to handle the problem is to unfold the hyperspectral image as shown in Figure 6.10, where $\widetilde{\boldsymbol{X}}^{(1)} = \boldsymbol{X}^{(1)}$ is the first slice of the hyperspectral image and $\widetilde{\boldsymbol{X}}^{(k)}$ is the k^{th} slice. The entire dataset at time instant $k+1$, *i.e.*, $\boldsymbol{X}^{(k+1)}$, can be represented as the concatenation of the first k samples with the new incoming sample *i.e.*, $\boldsymbol{X}^{(k+1)} = \left[\boldsymbol{X}^{(k)}\ \widetilde{\boldsymbol{X}}^{(k+1)}\right]$. Similarly, we define $\boldsymbol{S}^{(k+1)} = \left[\boldsymbol{S}^{(k)}\ \widetilde{\boldsymbol{S}}^{(k+1)}\right]$ and $\boldsymbol{A}^{(k+1)} = \left[\boldsymbol{A}^{(k)}\ \widetilde{\boldsymbol{A}}^{(k+1)}\right]$, where $\boldsymbol{S}$ denotes the endmembers (sources) matrix and $\boldsymbol{A}$ contains on its rows the abundances for the recorded samples.

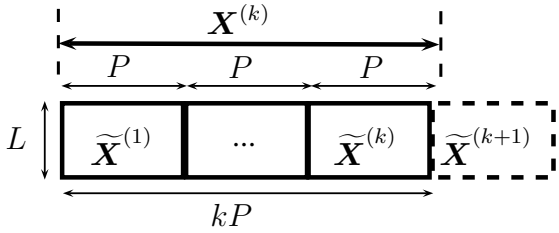


Figure 6.10: Unfolded pushbroom hyperspectral image

Then, the on-line NMF model is given by:

$$\widetilde{\boldsymbol{X}}^{(k+1)} \approx \widetilde{\boldsymbol{S}}^{(k+1)} \widetilde{\boldsymbol{A}}^{(k+1)}. \tag{6.18}$$

6.2.4 Cost function

The simplest way to fit the NMF model to the data is to minimize the least square distance between the data and the model. Let $\mathcal{J}^{(k)}$ denote the cost function corresponding to the first k samples:

$$\mathcal{J}^{(k)}\left(\boldsymbol{S}^{(k)}, \boldsymbol{A}^{(k)}\right) = \frac{1}{2}\sum_{\ell=1}^{k}\left\|\widetilde{\boldsymbol{X}}^{(\ell)} - \widetilde{\boldsymbol{S}}^{(\ell)}\widetilde{\boldsymbol{A}}^{(\ell)}\right\|_{\mathsf{F}}^{2}. \tag{6.19}$$

When the $k+1^{th}$ sample, $\widetilde{\boldsymbol{X}}^{(k+1)}$ arrives, the corresponding cost function can be decomposed as follows:

$$\mathcal{J}^{(k+1)}\left(\boldsymbol{S}^{(k+1)}, \boldsymbol{A}^{(k+1)}\right) = \frac{1}{2}\sum_{\ell=1}^{k}\left\|\widetilde{\boldsymbol{X}}^{(\ell)} - \widetilde{\boldsymbol{S}}^{(\ell)}\widetilde{\boldsymbol{A}}^{(\ell)}\right\|_{\mathsf{F}}^{2} + \frac{1}{2}\left\|\widetilde{\boldsymbol{X}}^{(k+1)} - \widetilde{\boldsymbol{S}}^{(k+1)}\widetilde{\boldsymbol{A}}^{(k+1)}\right\|_{\mathsf{F}}^{2}. \tag{6.20}$$

Without further assumptions, (6.20) is just a set of k independent least squares problems and in this case, the on-line setup has no particular interest. However, a natural assumption is

that the endmembers vary only slightly between consecutive samples *i.e.*, $\widetilde{\boldsymbol{S}}^{(k+1)} \approx \widetilde{\boldsymbol{S}}^{(k)}$, $\forall k$. Thus, the cost function (6.20) can be expressed as:

$$\mathcal{J}^{(k+1)}\left(\widetilde{\boldsymbol{S}}^{(k+1)}, \widetilde{\boldsymbol{A}}^{(k+1)} \middle| \boldsymbol{A}^{(k)}\right) = \mathcal{J}^{(k)}\left(\widetilde{\boldsymbol{S}}^{(k+1)} \middle| \boldsymbol{A}^{(k)}\right) + \widetilde{\mathcal{J}}^{(k+1)}\left(\widetilde{\boldsymbol{S}}^{(k+1)}, \widetilde{\boldsymbol{A}}^{(k+1)}\right), \tag{6.21}$$

with

$$\mathcal{J}^{(k)}\left(\widetilde{\boldsymbol{S}}^{(k+1)} \middle| \boldsymbol{A}^{(k)}\right) = \frac{1}{2}\sum_{\ell=1}^{k}\left\|\widetilde{\boldsymbol{X}}^{(\ell)} - \widetilde{\boldsymbol{S}}^{(k+1)}\widetilde{\boldsymbol{A}}^{(\ell)}\right\|_{\mathsf{F}}^{2}, \tag{6.22}$$

$$\widetilde{\mathcal{J}}^{(k+1)}\left(\widetilde{\boldsymbol{S}}^{(k+1)}, \widetilde{\boldsymbol{A}}^{(k+1)}\right) = \frac{1}{2}\left\|\widetilde{\boldsymbol{X}}^{(k+1)} - \widetilde{\boldsymbol{S}}^{(k+1)}\widetilde{\boldsymbol{A}}^{(k+1)}\right\|_{\mathsf{F}}^{2}. \tag{6.23}$$

In order to add some tracking capability to the algorithm, a weighting coefficient α ($0 \leq \alpha \leq 1$) is incorporated into the cost function as:

$$\mathcal{J}^{(k+1)}\left(\widetilde{\boldsymbol{S}}^{(k+1)}, \widetilde{\boldsymbol{A}}^{(k+1)}\right) = \alpha\mathcal{J}^{(k)} + (1-\alpha)\widetilde{\mathcal{J}}^{(k+1)}. \tag{6.24}$$

In others words, α controls the trade-off between the contribution of the old and the new samples. Nevertheless, without additional regularization, the solution of (6.24) is not unique. The uniqueness of the NMF relies on the sparsity of the underlying latent variables. In particular, if either $\boldsymbol{S}$ and/or $\boldsymbol{A}$ has only non-zero entries, the NMF factorization is not unique. To reduce the size of the set of admissible solutions, we propose to impose on the endmember matrix $\widetilde{\boldsymbol{S}}^{(k+1)}$, a geometric regularization which forces the simplex bounded by the endmembers to circumscribe the data as closely as possible. For now, we denote this penalty by $\text{Vol}\left(\widetilde{\boldsymbol{S}}^{(k+1)}\right)$ and integrate it into the cost function as follows:

$$\mathcal{J}_{\text{Vol}}^{(k+1)}\left(\widetilde{\boldsymbol{S}}^{(k+1)}, \widetilde{\boldsymbol{A}}^{(k+1)}\right) = \alpha\mathcal{J}^{(k)} + (1-\alpha)\widetilde{\mathcal{J}}^{(k+1)} + \mu\ \text{Vol}\left(\widetilde{\boldsymbol{S}}^{(k+1)}\right), \tag{6.25}$$

where $\mathcal{J}_{\text{Vol}}$ becomes the volume regularized criterion and $\mu \geq 0$ controls the trade-off between the data fitting term and the volume regularizer. Thus, the approach proposed in this paper aims at solving the following optimization problem:

$$\begin{aligned} &\underset{\widetilde{\boldsymbol{S}}^{(k+1)}, \widetilde{\boldsymbol{A}}^{(k+1)}}{\text{minimize}}\ \mathcal{J}_{\text{Vol}}^{(k+1)}\left(\widetilde{\boldsymbol{S}}^{(k+1)}, \widetilde{\boldsymbol{A}}^{(k+1)}\right), \\ &\text{subject to } \widetilde{\boldsymbol{S}}^{(k+1)} \geq \boldsymbol{0}, \widetilde{\boldsymbol{A}}^{(k+1)} \geq \boldsymbol{0}, \widetilde{\boldsymbol{A}}^{(k+1)^{\top}}\boldsymbol{1} = \boldsymbol{1}, \end{aligned} \tag{6.26}$$

for a particular choice of the volume regularizer term. The choice of $\text{Vol}\left(\widetilde{\boldsymbol{S}}^{(k+1)}\right)$ is largely discussed in [Nus20]. The Minimum Volume Regularization (MVR) was used in the NMF-like algorithm proposed in [Nus18], and it can be formulated as: $\log\left(\det\left(\widetilde{\boldsymbol{S}}^{\top}\widetilde{\boldsymbol{S}} + \varepsilon\boldsymbol{I}\right)\right)$, with

$\varepsilon > 0$ is a specified small value and $\boldsymbol{I}$ is the identity matrix.

In the following, we only consider the Minimum Dispersion Regularization (MDR), for the reasons mentioned above. The MDR is given by: trace $\left(\widetilde{\boldsymbol{S}}\boldsymbol{D}\widetilde{\boldsymbol{S}}^\top\right)$, where $\boldsymbol{D} = \boldsymbol{I} - \frac{1}{R}\mathbf{1}\mathbf{1}^\top$ ($\mathbf{1}$ is an all-ones column vector of size $R \times 1$). As illustrated in Figure 6.11, geometrically speaking, MDR computes the sum of the distances between each endmember and the endmembers' centroid. It can be thus interpreted as a measure of the dispersion of the endmembers around their centroid, and related indirectly to the volume of the convex hull of $\widetilde{\boldsymbol{S}}$.

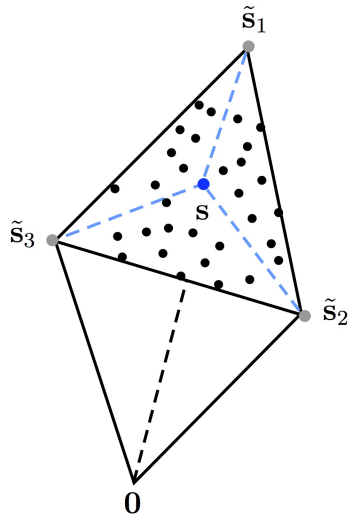


Figure 6.11: Geometric intuition of the minimum dispersion regularization for $R = 3$ endmembers

6.2.5 The proposed OMDR-ADMM algorithm

We proposed in [Nus20] an algorithm for on-line unmixing of hyperspectral images using the MDR and the ADMM optimization procedure, that we called *On-line MDR-ADMM* (OMDR-ADMM).

Let us define:

$$\mathcal{J}_{\text{Vol}}^{(k+1)}\left(\widetilde{\boldsymbol{S}}^{(k+1)}, \widetilde{\boldsymbol{A}}^{(k+1)}\right) = \alpha\mathcal{J}^{(k)} + (1-\alpha)\widetilde{\mathcal{J}}^{(k+1)} + \mu \text{ trace}\left(\widetilde{\boldsymbol{S}}^{(k+1)}\boldsymbol{D}\widetilde{\boldsymbol{S}}^{(k+1)^\top}\right). \quad (6.27)$$

By introducing two auxiliary variables $\widetilde{\boldsymbol{U}}$ and $\widetilde{\boldsymbol{V}}$, problem (6.26) is equivalent to:

$$\underset{\widetilde{\boldsymbol{S}}^{(k+1)}, \widetilde{\boldsymbol{A}}^{(k+1)}, \widetilde{\boldsymbol{V}}^{(k+1)}, \widetilde{\boldsymbol{U}}^{(k+1)}}{\text{minimize}} \mathcal{J}_{\text{Vol}}^{(k+1)}\left(\widetilde{\boldsymbol{S}}^{(k+1)}, \widetilde{\boldsymbol{A}}^{(k+1)}\right) + \mathbb{I}_{\mathbb{R}_+}\left(\widetilde{\boldsymbol{V}}^{(k+1)}\right) + \mathbb{I}_{\mathbb{R}_+}\left(\widetilde{\boldsymbol{U}}^{(k+1)}\right),$$
$$\text{subject to } \widetilde{\boldsymbol{S}}^{(k+1)} = \widetilde{\boldsymbol{U}}^{(k+1)}, \widetilde{\boldsymbol{A}}^{(k+1)} = \widetilde{\boldsymbol{V}}^{(k+1)} \text{and } \widetilde{\boldsymbol{A}}^{(k+1)^\top}\mathbf{1} = \mathbf{1}, \quad (6.28)$$

where $\mathbb{I}_{\mathbb{R}_+}$ is the indicator function of $\mathbb{R}_+$ ensuring the non-negativity of the endmembers and abundances. For algorithmic convenience, we use the scaled form of ADMM [Boyd11] in which the linear and quadratic terms are combined in the augmented Lagrangian and the dual variables are scaled.

Algorithm 3 summarizes the proposed OMDR-ADMM method for solving (6.28).

Algorithm 3: OMDR-ADMM

Inputs: $\boldsymbol{X}$; $\boldsymbol{D}$; R; α; μ; ρ; N_1; N_2;
Initialization: $k = 0$; $\boldsymbol{N} = \text{zeros}(L, R)$; $\boldsymbol{M} = \text{zeros}(R, R)$; $\widetilde{\boldsymbol{S}} = \text{rand}(L, R)$; $\widetilde{\boldsymbol{V}} = \text{zeros}(R, P)$; $\widetilde{\boldsymbol{U}} = \text{zeros}(L, R)$; $\widetilde{\boldsymbol{\Pi}} = \text{zeros}(R, P)$; $\widetilde{\boldsymbol{\Lambda}} = \text{zeros}(L, R)$; $\boldsymbol{A} = [\,]$; $\boldsymbol{S} = [\,]$;
Outputs: $\boldsymbol{A}$; $\boldsymbol{S}$;

while New sample $k+1$ available **do**
  $\widetilde{\boldsymbol{X}} = \widetilde{\boldsymbol{X}}^{(k+1)}$;
  $t_1 = 1$;
  while $t_1 < N_1$ **do**
    $t_2 = 1$;
    while $t_2 < N_2$ **do**
      $\widetilde{\boldsymbol{A}} \leftarrow \underset{\widetilde{\boldsymbol{A}}^\top \mathbf{1}=\mathbf{1}}{\arg\min} \frac{1}{2}(1-\alpha)\left\|\widetilde{\boldsymbol{X}} - \widetilde{\boldsymbol{S}}\widetilde{\boldsymbol{A}}\right\|_{\mathsf{F}}^2 + \frac{\rho}{2}\left\|\widetilde{\boldsymbol{A}} - \widetilde{\boldsymbol{V}} + \widetilde{\boldsymbol{\Pi}}\right\|_{\mathsf{F}}^2$;
      $\widetilde{\boldsymbol{V}} = \max\left(\mathbf{0}, \widetilde{\boldsymbol{A}} + \widetilde{\boldsymbol{\Pi}}\right)$;
      $\widetilde{\boldsymbol{\Pi}} \leftarrow \widetilde{\boldsymbol{\Pi}} + \widetilde{\boldsymbol{A}} - \widetilde{\boldsymbol{V}}$;
      $t_2 \leftarrow t_2 + 1$;
    end while
    $\widetilde{\boldsymbol{N}} = \alpha\boldsymbol{N} + (1-\alpha)\left(\widetilde{\boldsymbol{X}}\widetilde{\boldsymbol{A}}^\top\right)$;
    $\widetilde{\boldsymbol{M}} = \alpha\boldsymbol{M} + (1-\alpha)\left(\widetilde{\boldsymbol{A}}\widetilde{\boldsymbol{A}}^\top\right)$;
    $\widetilde{\boldsymbol{S}} = \left(\widetilde{\boldsymbol{N}} + \rho\left(\widetilde{\boldsymbol{U}} - \widetilde{\boldsymbol{\Lambda}}\right)\right)\left(\widetilde{\boldsymbol{M}} + 2\mu\boldsymbol{D} + \rho\boldsymbol{I}\right)^{-1}$;
    $\widetilde{\boldsymbol{U}} = \max\left(\mathbf{0}, \widetilde{\boldsymbol{S}} + \widetilde{\boldsymbol{\Lambda}}\right)$;
    $\widetilde{\boldsymbol{\Lambda}} \leftarrow \widetilde{\boldsymbol{\Lambda}} + \widetilde{\boldsymbol{S}} - \widetilde{\boldsymbol{U}}$;
    $t_1 \leftarrow t_1 + 1$;
  end while
  $\boldsymbol{N} = \widetilde{\boldsymbol{N}}$; $\boldsymbol{M} = \widetilde{\boldsymbol{M}}$;
  $\boldsymbol{A} \leftarrow \left[\boldsymbol{A}\ \widetilde{\boldsymbol{A}}\right]$; $\boldsymbol{S} \leftarrow \left[\boldsymbol{S}\ \widetilde{\boldsymbol{S}}\right]$;
end while

It includes three loops. The inner loop alternately estimates $\widetilde{\boldsymbol{A}}$, $\widetilde{\boldsymbol{V}}$ and $\widetilde{\boldsymbol{\Pi}}$ using a fixed number of iterations N_2. These estimates are iteratively refined in an outer loop with N_1 iterations, in which the parameters $\widetilde{\boldsymbol{S}}$, $\widetilde{\boldsymbol{U}}$ and $\widetilde{\boldsymbol{\Lambda}}$ are cyclically updated. This procedure is repeated for each new sample $\widetilde{\boldsymbol{X}}^{(k+1)}$.The inner loop of N_2 iterations ensures that the

objective function decreases between each iteration of the outer loop. In addition, and this was unexpected, including this additional loop allows to accelerate the convergence speed of the algorithm and consequently to reduce N_1.

It can be shown, using a similar approach to [Boyd04, Huang16], that any stationary point generated by a sequence of iterations of OMDR-ADMM satisfies the Karush-Kuhn-Tucker conditions. This statement is valid only in the case of our working assumption *i.e.* $\widetilde{\boldsymbol{S}}^{(k+1)} \approx \widetilde{\boldsymbol{S}}^{(k)}$, in other words, when the steady state is reached. No theoretical claims can be made regarding the transient behavior.

6.2.6 Experimental results on real data

In order to evaluate the performance of OMDR-ADMM for hyperspectral unmixing, as well as the influence of the various parameters on the algorithm performance, we conducted in [Nus20] extensive experiments on simulated and real hyperspectral images. In this manuscript we only illustrate the performance of the algorithm on two real datasets recorded by a push-broom acquisition system (which is the target application of our method) at CRAN (Centre de Recherche en Automatique de Nancy) laboratory.

Wood is a natural material whose rendering is appreciated in many applications. However, its surface heterogeneity and variability makes it difficult to control the quality of the final product. These technical difficulties generate significant non-quality costs and manual sorting in wood industries. Current industrial technical solutions only detect very pronounced defects, such as knots, using color vision systems. There are other types of defects, which have high colorimetric variability, and thus, are not systematically detected by the existing systems; this is the case of the sapwood (material located between the bark and the heartwood) and of the dark grain (dark color wood fibers), for example. The sapwood is generally undesirable in many applications due to its porosity, while dark grain-type defects are undesirable, particularly in the parquet industry, because they degrade the aesthetic appearance of the product. Thus, the detection of these defects represents an important economical issue, in many wood industry branches. A relevant technique to overcome the limitations of color vision systems is NIR hyperspectral imaging, which is the target application of this work. For the next experiment, we considered the wood samples shown in Figure 6.12, measuring 120 cm $\times$ 15 cm.

The sample 1 had a sapwood-type defect in the bottom left corner. The sample 2 had a dark grain area in the middle, over its entire length. These defects were hardly distinguishable by the colorimetric systems and were manually labeled by experts. The two pieces of wood were scanned by a NIR pushbroom imager working in the spectral range of 900-1600 nm, resulting in hyperspectral images of dimension $224 \times 270 \times 1514$ (for sample 1) and $224 \times 200 \times 1100$ (for sample 2), where 224 represents the number of wavelengths and the other two numbers, the (spatial $\times$ time) dimensions. The acquired images were processed sequentially by OMDR-ADMM, slice by slice. The number of endmembers to be extracted was set to $R = 3$, after several preliminary tests. The parameters of the algorithm were set as follows:

(a) Sample 1

(b) Sample 2

Figure 6.12: Selected wood samples (visible range color image)

$\widetilde{\mu} = 10^{-5}$, $\alpha = 0.99$, $\rho = 0.01$, $N_1 = 15$ and $N_2 = 10$. The endmember and abundance matrices were randomly initialized from a continuous uniform distribution on the interval $[0, 1]$.

The three abundance maps generated by OMDR-ADMM for the samples 1 and for the sample 2 are presented in Figure 6.13 and 6.14, respectively. On the first abundance map, we observe that the areas corresponding to the sapwood and to the dark grain are much more contrasted than the rest of the image. Moreover, it is interesting to highlight, particularly for sample 1, that all of the knots present on its surface are grouped on the third abundance map. These are very promising results, as, for the considered samples, the color of the sapwood/dark grain was very close to the color of heartwood, and therefore, the wood planks were considered as compliant with industrial quality standards by color vision systems. For the considered datasets, the processing time by OMDR-ADMM was of about 3.22 s, for the sample 1, and 2.24 s, for the sample 2. The processing time can be largely reduced by low-level implementation of our algorithm directly on the production line to satisfy the real-time industrial constraint of 1 linear m of wood /s.

We compared the results generated by OMDR-ADMM with the state-of-the-art volume-regularization algorithm RVolMin [Fu16] (see [Nus20]).

With a few differences, the abundance maps and the endmembers generated by RVolMin are very similar to those generated by OMDR-ADMM, which validates the effectiveness of the proposed approach. However, the major difference between the two methods is the processing time; *e.g.*, for sample 1, RVolMin processed the entire hyperspectral image in 2 min 15 s (compared to 3.22 s for OMDR-ADMM). This shows that there is a real interest in performing on-line blind hyperspectral unmixing.

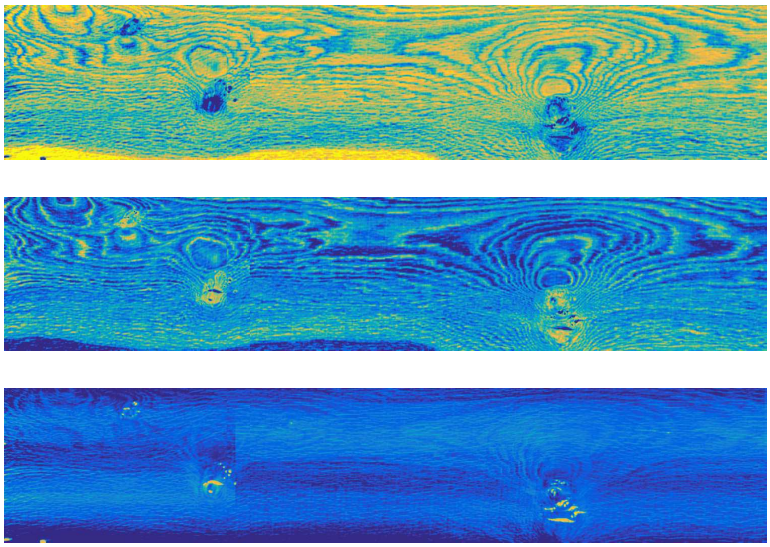


Figure 6.13: Abundance maps estimated by OMDR-ADMM (for the wood sample 1)

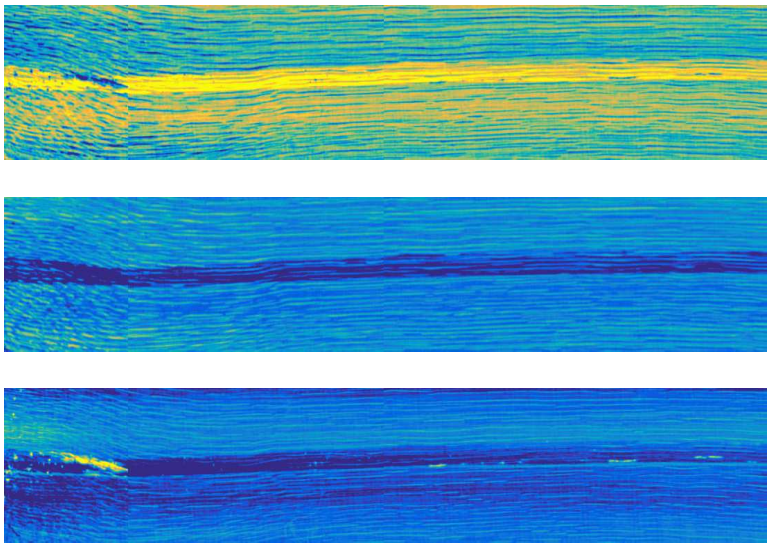


Figure 6.14: Abundance maps estimated by OMDR-ADMM (for the wood sample 2)

6.3 Quaternion source unmixing with non-negativity contraints

I present in this section a *quaternion non-negative matrix factorization (QNMF)* that extends the concept of non-negative matrix factorization (NMF) to polarized sources.

Spectro-polarimetric imaging exploits polarization diversity in hyperspectral imaging systems [Mu12, Zhao09, Miron11]. In full generality, it consists in acquiring, for a collection of wavelengths and locations, a 4-dimensional real vector gathering the 4 Stokes parameters. These energetic parameters are widely used in optics [Gil16] to describe the complete properties of light (intensity and polarization). The joint acquisition of spatial, spectral and polarization diversities raises important and original challenges, in particular regarding blind polarized source separation problems.

We proposed in [Flamant20] an approach for this source separation problem that relies on two key ingredients: *(i)* the algebraic representation of polarization information using quaternions and *(ii)* the exploitation of physical constraints on Stokes parameters, which generalize non-negativity to the case of polarized signals. Non-negativity is to be understood here in a *vectorial* sense: it encodes the non-negativity, or positive semi-definiteness, of the 2-by-2 complex covariance matrix associated with each polarized source, as will be explained later.

This *generalized* non-negativity constraint on the source factor makes QNMF fundamentally different from NMF. In fact, the three Stokes parameters describing polarization are not required to be real non-negative: instead, Stokes parameters must lie inside a particular 4D second-order convex cone. This means that QNMF cannot be simply reformulated as an augmented real NMF problem by stacking Stokes parameters in a taller matrix. Instead, QNMF preserves the intrinsic structure of Stokes parameters by leveraging their quaternion representation: it allows the associated generalized non-negativity constraints to be handled straightforwardly.

We believe that quaternions, beyond offering an elegant way to handle polarization diversity, permit numerous theoretical (*e.g.* uniqueness conditions) and methodological (*e.g.* algorithmic design) developments that would have been cumbersome to obtain otherwise.

6.3.1 Basic notions on quaternions

Quaternions $\mathbb{H}$ forms a 4-dimensional normed division algebra over the real numbers $\mathbb{R}$ with canonical basis $\{1, \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ where $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ are imaginary units $\boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = -1$ such that $\boldsymbol{ij} = \boldsymbol{k}$, $\boldsymbol{ij} = -\boldsymbol{ji}$ and $\boldsymbol{ijk} = -1$. Importantly, quaternion multiplication is noncoMiroZBDFBC20ietspmmutative, *i.e.* for $q, q' \in \mathbb{H}$, $qq' \neq q'q$ in general. Any quaternion $q \in \mathbb{H}$ reads in Cartesian form $q = a + b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k}$ where $a, b, c, d \in \mathbb{R}$. The real or scalar part of q is $\mathrm{Re}\, q = a$ and its imaginary or vector part is $\mathrm{Im}\, q = b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k}$. We note $\mathrm{Im}_{\boldsymbol{i}}\, q = b$, $\mathrm{Im}_{\boldsymbol{j}}\, q = c$ and $\mathrm{Im}_{\boldsymbol{k}}\, q = d$. When $\mathrm{Re}\, q = 0$, q is said to be *pure imaginary*, or simply *pure*. The quaternion conjugate of q is denoted by $\overline{q} = \mathrm{Re}\, q - \mathrm{Im}\, q$. The modulus of q is

$|q| = \sqrt{q\overline{q}} = \sqrt{\overline{q}q}$.

A quaternion matrix $\boldsymbol{Q} \in \mathbb{H}^{M\times N}$ has elements $\boldsymbol{Q}_{mn} = q_{mn} \in \mathbb{H}$. Notation $\overline{\boldsymbol{Q}}$ denotes the entry-wise conjugate of $\boldsymbol{Q}$ whereas its conjugate-transpose is denoted by $\boldsymbol{Q}^\dagger = \overline{\boldsymbol{Q}}^\top$. Usual matrix operations require special care due to non-commutativity of quaternion multiplication: *e.g.* given two matrices $\boldsymbol{Q}_1 \in \mathbb{H}^{M\times N}$ and $\boldsymbol{Q}_2 \in \mathbb{H}^{N\times P}$, in general one has $\overline{\boldsymbol{Q}_1\boldsymbol{Q}_2} \neq \overline{\boldsymbol{Q}}_1\,\overline{\boldsymbol{Q}}_2$ and $(\boldsymbol{Q}_1\boldsymbol{Q}_2)^\top \neq \boldsymbol{Q}_2^\top\boldsymbol{Q}_1^\top$. However, $(\boldsymbol{Q}_1\boldsymbol{Q}_2)^\dagger = \boldsymbol{Q}_2^\dagger\boldsymbol{Q}_1^\dagger$ is always true. More details on quaternions and their properties can be found in [Conway03]. See also [Brenner51, Lee48, Zhang97, Le Bihan04a, Rodman14a] and references therein for more on quaternion matrix algebra.

6.3.2 Polarization and Stokes parameters

For electromagnetic (EM) waves such as light traveling in free space, the instantaneous electric field vector $\boldsymbol{E}(t)$ lives in the 2D plane orthogonal to the direction of propagation. In such transverse waves, the end tip of $\boldsymbol{E}(t) \in \mathbb{R}^2$ describes a two-dimensional elliptical trajectory whose properties can vary with time, space or frequency. This phenomenon is called *polarization*: it describes the geometric nature of wave oscillations. This behaviour is not limited to electromagnetic radiation: transverse elastic waves and gravitational waves also exhibit polarization.

In optics [Born80, Brosseau98, Gil16], the polarization state of light is commonly described by four Stokes parameters S_0, S_1, S_2, S_3. These real-valued parameters are energetic quantities that can be measured in experiments [Schaefer07, Berry77]. The first Stokes parameter $S_0 \geq 0$ is classical and measures the *total* intensity of light, *i.e.* the sum of intensities from the *polarized* and *unpolarized* parts of light. The three remaining Stokes parameters S_1, S_2, S_3 describe the polarized part of light, *i.e.* the properties of the polarization ellipse encoded by the trajectory of the electric field vector. The *degree of polarization* $\Phi \in [0, 1]$ rules the relative contribution of the polarized part to the total intensity:

$$\Phi = \frac{\text{intensity of polarized part}}{\text{total intensity}} = \frac{\sqrt{S_1^2 + S_2^2 + S_3^2}}{S_0}\,. \tag{6.29}$$

When $\Phi = 1$, light is said to be *fully polarized*, whereas for $\Phi = 0$ it is said *unpolarized*. For $\Phi \in (0, 1)$, light is said to be *partially polarized*.

6.3.3 Quaternion representation of Stokes parameters

The Stokes vector $(S_0, S_1, S_2, S_3)^\top \in \mathbb{R}^4$ can be conveniently expressed using quaternion algebra as

$$w = S_0 + \boldsymbol{i}S_3 + \boldsymbol{j}S_1 + \boldsymbol{k}S_2 \in \mathbb{H}\,. \tag{6.30}$$

This algebraic representation of Stokes parameters has been long standing in the field of optics [Tudor10b, Tudor10a, Whitney71, Pellat-Finet84, Richartz49]. More recently [Flamant17,

Flamant18], it has been shown that (6.30) exhibits a general relevance for arbitrary bivariate signals. Our choice for the ordering of Stokes parameters in (6.30) adopts that of [Flamant17].

The quaternion representation of Stokes parameters enables straightforward physical and geometric interpretations while simplifying computations [Flamant18]. Eq. (6.30) can be rewritten for convenience as

$$w = I + I\Phi\boldsymbol{\mu} \tag{6.31}$$

where $I := S_0 \geq 0$ is the total intensity and $\boldsymbol{\mu}$ is the *polarization axis* – a pure unit quaternion such that $\boldsymbol{\mu}^2 = -1$. With this notation, the quantity $\Phi\boldsymbol{\mu}$ is a pure quaternion that can be identified with a vector of $\mathbb{R}^3$ inside the Poincaré sphere [Born80] which completely encodes the polarization state of light. The expression (6.31) also emphasizes one of the benefits of the quaternion formalism: it allows a straightforward separation between pure energetic information, conveyed by $\operatorname{Re} w = I$, and the geometric vector information encoded by $\operatorname{Im} w = I\Phi\boldsymbol{\mu}$.

6.3.4 Polarization constraint and quaternion non-negativity

Stokes parameters are energetic quantities which are subject to admissibility conditions. A vector $(S_0, S_1, S_2, S_3)^\top \in \mathbb{R}^4$ is admissible as a Stokes vector if it satisfies the following constraints $(\mathcal{S})$

$$S_0 \geq 0 \quad \text{and} \quad S_1^2 + S_2^2 + S_3^2 \leq S_0^2 \,. \tag{$\mathcal{S}$}$$

Those two constraints bear a strong physical interpretation. The first one is classical and indicates that the total intensity S_0 is a non-negative real quantity. The second one means that the intensity of the polarized part cannot be larger than the total intensity. As a result, these constraints equivalently encode the range $0 \leq \Phi \leq 1$ of the degree of polarization (6.29). Note that $(\mathcal{S})$ does not impose at all that S_1, S_2 and S_3 are non-negative.

From a mathematical perspective, $(\mathcal{S})$ defines a closed second-order convex cone in $\mathbb{R}^4$, denoted by $\mathbb{R}^4_{\mathcal{S}}$. Conditions $(\mathcal{S})$ are referred to as *non-negativity* constraints for polarized signals by the following argument. Given $(S_0, S_1, S_2, S_3)^\top \in \mathbb{R}^4$, consider the 2-by-2 Hermitian matrix[4]:

$$\boldsymbol{J} = \frac{1}{2}\begin{bmatrix} S_0 + S_1 & S_2 + \boldsymbol{i}S_3 \\ S_2 - \boldsymbol{i}S_3 & S_0 - S_1 \end{bmatrix} \in \mathbb{C}^{2\times 2}. \tag{6.32}$$

This mapping is bijective. Then, a quick check shows that imposing positive semi-definiteness of $\boldsymbol{J}$ – or simply, *non-negativity* of $\boldsymbol{J}$, is equivalent to $(\mathcal{S})$, *i.e.*

$$(\mathcal{S}) \Leftrightarrow \begin{cases} \operatorname{trace}\{\boldsymbol{J}\} & \geq 0 \\ \det \boldsymbol{J} & \geq 0 \end{cases} . \tag{6.33}$$

[4]In optics, the matrix $\boldsymbol{J}$ is known as the *polarization* or *coherency* matrix [Gil16, Section 1.4] associated with the bivariate electric field $\boldsymbol{E}$. It can be statistically defined as a covariance matrix [Gil16, Section 1.4.1], from which Hermitian positive semi-definiteness follows.

This shows that $\mathbb{R}^4_{\mathcal{S}}$ and the set of 2-by-2 Hermitian non-negative matrices are isomorphic. This result will prove to be essential later on for imposing ($\mathcal{S}$) in practice, see Section 6.3.8. Since the proposed approach relies on the quaternion representation (6.30) of Stokes parameters, we define by extension the set of *non-negative quaternions* $\mathbb{H}_{\mathcal{S}} \subset \mathbb{H}$ such that

$$\mathbb{H}_{\mathcal{S}} \triangleq \left\{ q \in \mathbb{H} \middle| \operatorname{Re} q \geq 0 \text{ and } |\operatorname{Im} q|^2 \leq (\operatorname{Re} q)^2 \right\} . \tag{6.34}$$

By definition, the sets $\mathbb{H}_{\mathcal{S}}$ and $\mathbb{R}^4_{\mathcal{S}}$ are isomorphic. They share the same properties: in particular, this means that $\mathbb{H}_{\mathcal{S}}$ is a closed second-order convex cone of $\mathbb{H}$.

6.3.5 The proposed quaternion non-negative matrix factorization

Without loss of generality, let us consider the setting of a typical spectro-polarimetric experiment. In such experiment, the four Stokes parameters $S_i(\nu_m, u_n)$, $i = 0, 1, 2, 3$ are acquired for a collection of sampled frequencies (or wavelengths) $(\nu_m)_{1 \leq m \leq M}$ and spatial locations $(u_n)_{1 \leq n \leq N}$. Using (6.30), the collected data can be written in quaternion form as

$$x(\nu_m, u_n) = S_0(\nu_m, u_n) + \boldsymbol{i} S_3(\nu_m, u_n) + \boldsymbol{j} S_1(\nu_m, u_n) + \boldsymbol{k} S_2(\nu_m, u_n) \in \mathbb{H}_{\mathcal{S}} \tag{6.35}$$

for $1 \leq m \leq M$, $1 \leq n \leq N$. Note that spectro-polarimetric imaging includes standard hyperspectral imaging as a special case, in which only intensity measurements are performed *i.e.* $x(\nu_m, u_n) = S_0(\nu_m, u_n)$.

A standard assumption in hyperspectral imaging is to consider that, at a given pixel u_n, data can be represented as a linearly weighted combination of P elementary sources (or endmembers) [Bioucas-Dias12, Ma14]. While being simple, the linear mixing model appears reasonable for many real-world settings [Keshava02, Dobigeon16]. Generalizing this model to the case of polarized sources, we decompose spectro-polarimetric data (6.35) as the linear superposition of P elementary polarized sources $w_p(\,\cdot\,) \in \mathbb{H}_{\mathcal{S}}$:

$$x(\nu_m, u_n) = \sum_{p=1}^{P} w_p(\nu_m) h_p(u_n) . \tag{6.36}$$

The quantity $h_p(\,\cdot\,) \geq 0$ denotes the spatial activation coefficients attached to the source p. Considering that $h_p(\,\cdot\,)$ is a non-negative real quantity has two key implications. First, it means that $h_p(u_n)$ can be interpreted as the abundance of the source p at location u_n. Second, it ensures that, for any elementary source $w_p(\,\cdot\,) \in \mathbb{H}_{\mathcal{S}}$ the linear combination $\sum_{p=1}^{P} w_p(\nu_m) h_p(u_n)$ is in $\mathbb{H}_{\mathcal{S}}$ since $\mathbb{H}_{\mathcal{S}}$ is a convex cone.

Rewriting (6.36) into matrix form introduces *Quaternion Non-Negative Matrix Factorization* (QNMF) like

$$\boldsymbol{X} = \boldsymbol{W}\boldsymbol{H} \tag{6.37}$$

where $\boldsymbol{X} \in \mathbb{H}_{\mathcal{S}}^{M \times N}$ is the *data matrix* with coefficients $(\boldsymbol{X})_{mn} = x_{mn} = x(\nu_m, u_n)$. The

matrix $\boldsymbol{W} \in \mathbb{H}_{\mathcal{S}}^{M\times P}$ is the *source matrix* with coefficients $(\boldsymbol{W})_{mp} = w_{mp} = w_p(\nu_m)$, whereas $\boldsymbol{H} \in \mathbb{R}_+^{P\times N}$ is called the *activation matrix* such that $(\boldsymbol{H})_{pn} = h_{pn} = h_p(u_n)$.

6.3.6 Relation with NMF

The QNMF problem extends the well-known NMF problem to the case of polarized signals. Compared to NMF, QNMF features a quaternion-valued sources factor $\boldsymbol{W}$ which exploits the constraint ($\mathcal{S}$) instead of the usual non-negativity constraint. Using quaternion algebra rules, the QNMF (6.37) can also be rewritten as

$$\boldsymbol{X} = \boldsymbol{W}\boldsymbol{H} \Leftrightarrow \begin{cases} \operatorname{Re}\boldsymbol{X} = [\operatorname{Re}\boldsymbol{W}]\,\boldsymbol{H} & \text{(NMF)} \\ \operatorname{Im}\boldsymbol{X} = [\operatorname{Im}\boldsymbol{W}]\,\boldsymbol{H} & \text{(polarization)} \end{cases}. \tag{6.38}$$

Eq. (6.38) shows that QNMF can be seen as a co-factorization problem with common activation factor $\boldsymbol{H}$. The first factorization problem is an usual NMF on the real part of $\boldsymbol{X}$, *i.e.* on intensity data (Stokes parameter $S_0 \geq 0$) only. The second one corresponds to a factorization problem on the imaginary part of $\boldsymbol{X}$ encoding polarization properties (Stokes parameters S_1, S_2, S_3). These two factorization problems are not independent for two reasons: *(i)* the activation factor $\boldsymbol{H}$ appears in both and *(ii)* for each coefficient (m, p) of the source factor $\boldsymbol{W}$, the constraint ($\mathcal{S}$) connects the modulus of the imaginary part (polarization factorization problem) to its real part (NMF problem) by an inequality constraint.

The relationship (6.38) provides another illustration of how QNMF extends NMF to account for polarization diversity. It allows a precise quantification of the role played by the polarization information and its associated constraint ($\mathcal{S}$). This will be particularly useful for comparing uniqueness conditions of QNMF and NMF, see Section 6.3.7.

6.3.7 Uniqueness conditions for QNMF

As for tensors, a fundamental question regarding QNMF is: upon which conditions on the source $\boldsymbol{W}$ and activation $\boldsymbol{H}$ factors is the QNMF $\boldsymbol{X} = \boldsymbol{W}\boldsymbol{H}$ unique? In [Flamant20] we studied in detail the $P = 2$ sources case and provided a sufficient condition for QNMF uniqueness. We then, considered the general $P \geq 2$ case and gave a necessary uniqueness condition that relaxes the classical NMF requirements for sources $\boldsymbol{W}$ to have some zero entries. These conditions are summarized next.

It is worth mentioning that QNMF presents the same trivial ambiguities as NMF: those can be handled by standard techniques, *e.g.* by imposing normalization on the source coefficients to get rid of the scale indeterminacy. As explained in Section 6.3.6, QNMF generalizes NMF to the case of polarized signals. The co-factorization perspective given by (6.38) shows that QNMF inherits sufficient uniqueness conditions from NMF.

Two sources case: range of admissible solutions and sufficient uniqueness condition. We first consider the simpler case of two sources ($P = 2$), which is particularly instrumental in understanding the role played by the polarization constraint ($\mathcal{S}$). The approach follows closely the one presented in [Moussaoui05] for the NMF case.

For $1 \leq m \leq M$ and $1 \leq n \leq N$, the corresponding entries of each of the factors $\boldsymbol{W}$ and $\boldsymbol{H}$ are given by

$$(\boldsymbol{W})_{mp} = w_{mp} \in \mathbb{H}_{\mathcal{S}} \text{ and } (\boldsymbol{H})_{pn} = h_{pn} \geq 0 \tag{6.39}$$

where $p = 1, 2$ denotes the source index. We use the convenient quaternion form (6.31) for w_{mp}, *i.e.*

$$w_{mp} = I_{mp} + I_{mp}\Phi_{mp}\boldsymbol{\mu}_{mp} \,. \tag{6.40}$$

Proposition 6.1 (Sufficient uniqueness condition[Flamant20])**.** *If the following conditions are satisfied:*

- $\exists\, m_1, m_2 \in \{1, 2, \dots M\}$ *s.t.*

$$\begin{cases} \Phi_{m_1 1} = 1,\ \Phi_{m_1 2}\boldsymbol{\mu}_{m_1 2} \neq \boldsymbol{\mu}_{m_1 1} \\ I_{m_1 1} \geq \dfrac{1}{2}\dfrac{1 - \Phi_{m_1 2}^2}{1 - \Phi_{m_1 2}\left\langle \boldsymbol{\mu}_{m_1 1}, \boldsymbol{\mu}_{m_2 2} \right\rangle} I_{m_1 2} \\ \Phi_2(k_2) = 1,\ \Phi_{m_2 1}\boldsymbol{\mu}_{m_2 1} \neq \boldsymbol{\mu}_{m_2 2} \\ I_{m_2 2} \geq \dfrac{1}{2}\dfrac{1 - \Phi_{m_2 1}^2}{1 - \Phi_{m_2 1}\left\langle \boldsymbol{\mu}_{m_2 2}, \boldsymbol{\mu}_{m_2 1} \right\rangle} I_{m_2 1} \end{cases} \tag{C1}$$

- $\exists\, n_1, n_2 \in \{1, 2, \dots N\}, n_1 \neq n_2$ *s.t.*

$$\begin{cases} h_{1 n_1} > 0 \ \textit{and}\ h_{2 n_1} = 0 \\ h_{2 n_2} > 0 \ \textit{and}\ h_{1 n_2} = 0 \end{cases} \tag{C2}$$

then the QNMF $\boldsymbol{X} = \boldsymbol{W}\boldsymbol{H}$ is essentially unique.

On the one hand, condition (C2) is identical to the one found for the activation factor in the standard NMF case [Moussaoui05, Laurberg08, Huang14]. On the other hand, condition (C1) illustrates the key role played by polarization information. Compared to the usual NMF sufficient conditions for the 2 sources case [Moussaoui05, Laurberg08, Huang14], it does not require each source to vanish alternatively.

General case ($P \geq 2$). The study of the uniqueness of QNMF $\boldsymbol{X} = \boldsymbol{W}\boldsymbol{H}$ for an arbitrary number of P sources is much more cumbersome than the $P = 2$ sources case, as the previous explicit modeling of indeterminacies is no longer amenable.

However, we showed that QNMF relaxes the NMF zero-entry necessary condition on $\boldsymbol{W}$, stated by Proposition 6.2 below.

Proposition 6.2 (Necessary condition for uniqueness with non-vanishing sources [Flamant20])**.** *Suppose that the QNMF* $\boldsymbol{X} = \boldsymbol{W}\boldsymbol{H}$ *is essentially unique such that* $\mathrm{Re}\, w_{mp} > 0$ *for every* m, p*. Then the following conditions are satisfied:*

- $\forall (p,q),\ p \neq q,$
$$\exists m \ s.t.\ \Phi_{mp} = 1,\ \Phi_{mq}\boldsymbol{\mu}_{mq} \neq \boldsymbol{\mu}_{mp} \tag{A1}$$
- $\forall (p,q),\ p \neq q,$
$$\exists n \ s.t.\ h_{pn} = 0 \ and \ h_{qn} > 0 \tag{A2}$$

Proposition 6.2 shows that a necessary condition for uniqueness of the QNMF with non-vanishing sources is that, for any distinct pair of sources (p, q), two criteria are satisfied: (A1) there exists an index m such that the source p is fully polarized and the source q exhibits a different polarization state (it can be fully polarized but with a different polarization axis); (A2) there is an index n such that the source q is active while the source p is not. Finally, note that condition (A2) on the activation factor is exactly the same as in the NMF case [Moussaoui05, Laurberg08, Huang14].

6.3.8 An algorithm for QNMF

We formulated QNMF as the following optimization problem:

$$\min_{\substack{\boldsymbol{W}\in\mathbb{H}_{\mathcal{S}}^{M\times P} \\ \boldsymbol{H}\in\mathbb{R}_{+}^{P\times N}}} \|\boldsymbol{X} - \boldsymbol{W}\boldsymbol{H}\|_{\mathsf{F}}^{2}\,. \tag{6.41}$$

The formulation of the resolution of the QNMF problem (6.41) appears much alike the standard NMF with Euclidean cost. However, two fundamental questions need to be answered: *(i)* is the constraint ($\mathcal{S}$) easy to implement? and *(ii)* can we optimize w.r.t. $\boldsymbol{W}$ directly in the quaternion domain?

Fortunately, the answers to these two questions are affirmative. For *(i)*, the answer relies on the key link between ($\mathcal{S}$) and the set of non-negative Hermitian 2-by-2 matrices. A positive answer to *(ii)* is made possible by the recent advent of the theory of quaternion-domain derivatives [Xu15b, Xu16, Xu15a, Mandic11] – the so-called *generalized* $\mathbb{HR}$ *calculus*. A detailed description of the implementation of the constraints on $\boldsymbol{W}$ and $\boldsymbol{H}$ factors can be found in [Flamant20].

The proposed algorithm adopts a popular strategy for solving the QNMF problem, based on the alternating constrained minimization of (6.41) w.r.t. $\boldsymbol{H}$ and $\boldsymbol{W}$. This choice is motivated by the fact that, whereas the Euclidean cost (6.41) is not convex in both $\boldsymbol{W}$ and $\boldsymbol{H}$, it is convex in each variable separately. After initialization of the factors, the iteration $r > 0$

reads

$$\boldsymbol{H}_{r+1} \leftarrow \underset{\boldsymbol{H} \in \mathbb{R}_+^{P\times N}}{\arg\min} \|\boldsymbol{X} - \boldsymbol{W}_r \boldsymbol{H}\|_{\mathsf{F}}^2 \tag{6.42}$$

$$\boldsymbol{W}_{r+1} \leftarrow \underset{\boldsymbol{W} \in \mathbb{H}_{\mathcal{S}}^{M\times P}}{\arg\min} \|\boldsymbol{X} - \boldsymbol{W} \boldsymbol{H}_{r+1}\|_{\mathsf{F}}^2 \,. \tag{6.43}$$

Equations (6.42)-(6.43) describe a two-block coordinate descent (CD) scheme in the quaternion domain. Two-block real-domain CD schemes are known to converge to a stationary point under mild conditions [Bertsekas99, Grippo00, Kim14]. These convergence results can be readily transposed to the quaternion case by identifying $\mathbb{H}$ with $\mathbb{R}^4$. Formally, since $\mathbb{R}^+$ and $\mathbb{H}_{\mathcal{S}}$ are closed convex sets, every limit point of the sequence $(\boldsymbol{W}_r, \boldsymbol{H}_r)$ generated by the two-block CD framework (6.42)-(6.43) is a stationary point of (6.41).

We proposed a simple, closed-form algorithm that approximately solves each subproblem (6.42), (6.43) at each iteration, at the price of losing convergence guarantees. The main idea is the following: at a given iteration r, for each factor, one solves the unconstrained least-squares problem and projects the obtained solution onto the corresponding constraint:

$$\boldsymbol{H}_{r+1} \leftarrow \Pi_{\mathbb{R}_+} \left[\underset{\boldsymbol{H}}{\arg\min} \|\boldsymbol{X} - \boldsymbol{W}_r \boldsymbol{H}\|_{\mathsf{F}}^2 \right] \tag{6.44}$$

$$\boldsymbol{W}_{r+1} \leftarrow \Pi_{\mathbb{H}_{\mathcal{S}}} \left[\underset{\boldsymbol{W}}{\arg\min} \|\boldsymbol{X} - \boldsymbol{W} \boldsymbol{H}_{r+1}\|_{\mathsf{F}}^2 \right] \,. \tag{6.45}$$

Due to its resemblance with the usual alternating least squares (ALS) algorithm [Paatero94] for NMF, we call this strategy *quaternion alternating least squares* (QALS). Derivation of explicit updates for unconstrained least-squares problems requires special care because of the quaternion nature of $\boldsymbol{X}$ and $\boldsymbol{W}$ (see [Flamant20]). As a result, one gets the explicit updates:

$$\boldsymbol{H}_{r+1} \leftarrow \Pi_{\mathbb{R}_+} \left[\left(\mathrm{Re} \left[\boldsymbol{W}_r^\top \overline{\boldsymbol{W}}_r \right] \right)^{-1} \mathrm{Re} \left[\boldsymbol{W}_r^\top \overline{\boldsymbol{X}} \right] \right] \tag{6.46}$$

$$\boldsymbol{W}_{r+1} \leftarrow \Pi_{\mathbb{H}_{\mathcal{S}}} \left[\boldsymbol{X} \boldsymbol{H}_{r+1}^\top \left(\boldsymbol{H}_{r+1} \boldsymbol{H}_{r+1}^\top \right)^{-1} \right] \,. \tag{6.47}$$

Projections implementations onto constraints sets are detailed in [Flamant20].

The proposed algorithm is remarkably simple and cheap. Performing optimization directly in the quaternion-domain yields updates expressions that are very much alike the standard ALS algorithm for NMF [Berry07]. Such computations would not have been amenable without using the powerful theory of quaternion derivatives introduced recently [Xu15b, Xu16, Xu15a, Mandic11]. Despite the lack of theoretical guarantees on its convergence, the quaternion ALS algorithm provides a good baseline for the resolution of QNMF with reasonably good results in most situations. Thus, it paves the way to further developments of more sophisticated algorithms for the QNMF problem.

6.3.9 A numerical experiment

To illustrate the relevance of the proposed approach, we present in this section a numerical experiment on synthetic data. More experimental results can be found in [Flamant20].

We simulated a realistic simulated spectro-polarimetric data consisting in a linear mixture of $P = 3$ sources imaged at $M = 128$ wavelengths and $N = 512 \times 512$ spatial locations. Each source exhibits broad-band behavior, featuring non-vanishing intensity S_0 spectral profile together with spectrally varying polarization $S_1/S_0, S_2/S_0, S_3/S_0$ parameters. For simplicity, we assume that each source is fully polarized. Two-dimensional activations maps of each source are designed to represent a typical spectro-polarimetric imaging setting, *e.g.* solar spectro-polarimetry [McCauley19, Orozco Suárez07], and such that the resulting activation matrix $\boldsymbol{H}$ satisfies the separability condition [Huang14]. As a result, the designed $\boldsymbol{W}$ and $\boldsymbol{H}$ matrices satisfy the necessary conditions for QNMF identifiability stated in Proposition 6.2. Fig. 6.15A shows simulated spectro-polarimetric data $\boldsymbol{X} = \boldsymbol{W}\boldsymbol{H}$ for two distinct wavelengths indices $m = 38$ and $m = 76$.

We show that QALS permits to effectively recover the QNMF factors. We adopt a random initialization strategy. The initial activation $\boldsymbol{H}_0$ is chosen as a matrix of i.i.d. entries drawn from $\mathcal{U}([0, 1])$, the uniform distribution on $[0, 1]$. The initial source factor $\boldsymbol{W}_0$ has i.i.d. entries drawn from the quaternion circular unit Gaussian distribution [Le Bihan17] projected onto the constraint ($\mathcal{S}$). We then alternate Eqs. (6.46)–(6.47) until convergence. The latter was assessed by monitoring the relative error $\varepsilon_r = \|\boldsymbol{X} - \boldsymbol{W}_r\boldsymbol{H}_r\|_{\mathsf{F}}^2/\|\boldsymbol{X}\|_{\mathsf{F}}^2$ and stopping whenever the improvement was below a given threshold, *i.e.* $|\varepsilon_r - \varepsilon_{r-1}| \leq 10^{-5}$ in our case. Repeating this procedure for $K = 100$ independent initializations of the QALS algorithm, it took on average 29 iterations for the algorithm to converge. Out of 100 initializations, we observed that the QALS algorithm recovered each time identical sources and activation factors, strongly indicating that the QNMF of the data $\boldsymbol{X}$ may be essentially unique.

Fig. 6.15B depicts the 3 sources and corresponding activations factors recovered by the QALS algorithm for one arbitrary initialization. Excellent match between reconstructed and ground truth factors is indicated by very small relative errors, *i.e.* $\varepsilon_{\boldsymbol{H}} = \|\hat{\boldsymbol{H}} - \boldsymbol{H}\|_{\mathsf{F}}^2/\|\boldsymbol{H}\|_{\mathsf{F}}^2 = 4.91 \cdot 10^{-5}$ and $\varepsilon_{\boldsymbol{W}} = \|\hat{\boldsymbol{W}} - \boldsymbol{W}\|_{\mathsf{F}}^2/\|\boldsymbol{W}\|_{\mathsf{F}}^2 = 3.92 \cdot 10^{-5}$. Looking at sources intensities S_0, we observe that they never vanish, *i.e.* $S_0 > 0$, plus that the first and second sources exhibit very similar peaks. This is a very challenging situation for NMF-based methods, which would require additional regularizations such as minimum volume [Lin15, Fu15, Fu18] to ensure uniqueness of the solution. It is thus remarkable that the QNMF model with the QALS algorithm is able to solve this blind source problem in a simple manner by leveraging polarization information. One of the advantages of polarization in source separation problems can be directly seen by inspecting $\boldsymbol{X}_{76:}$ in Fig. 6.15A: while it is not possible to distinguish 2 sources using S_0 only, the distinctive polarization pattern on S_2 makes it easy to identify source 1 and source 2.

These simulation results demonstrate the ability of the proposed QALS algorithm to effectively solve the QNMF problem. Despite no associated provable convergence guaran-

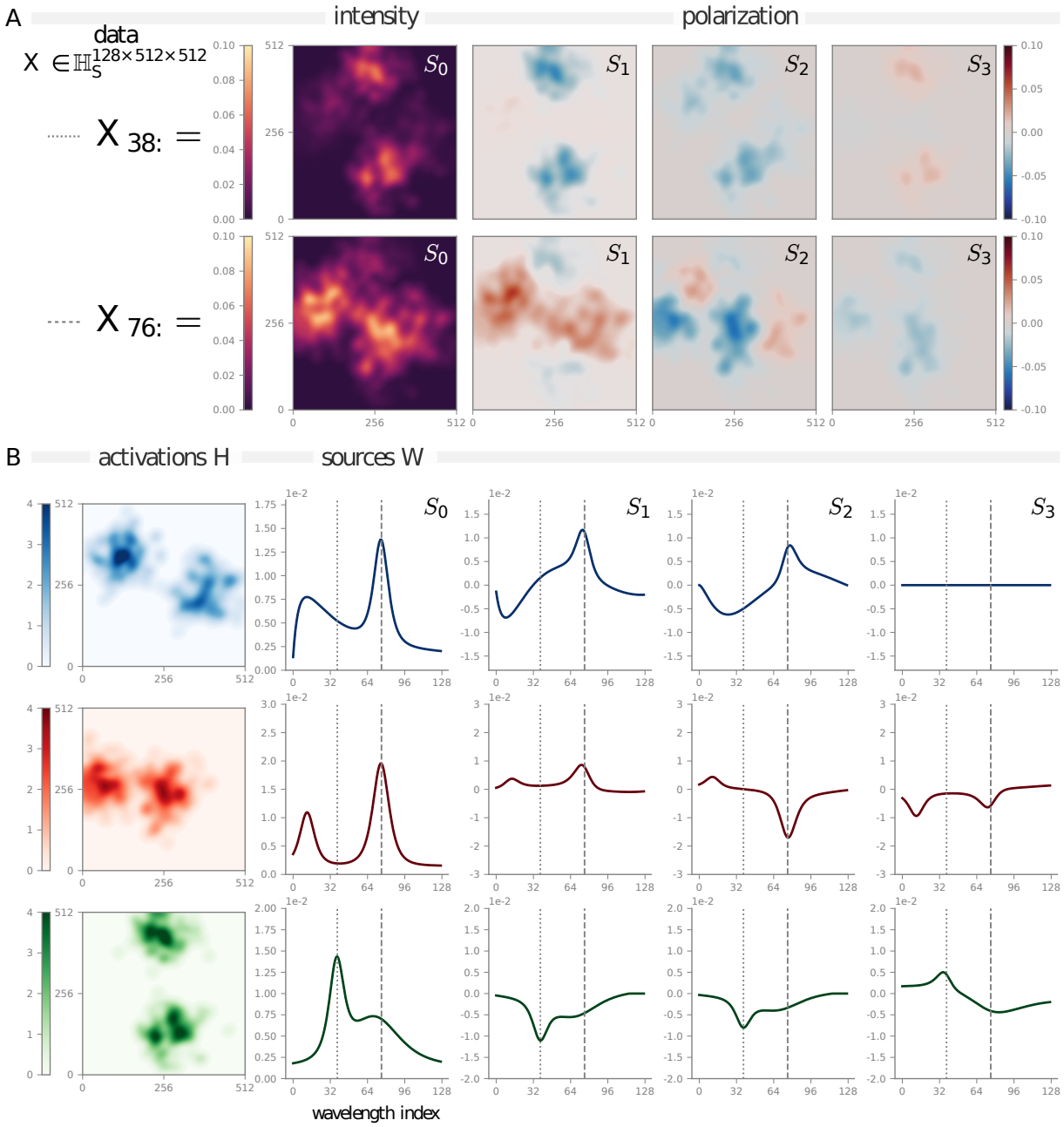


Figure 6.15: Blind source separation of spectro-polarimetric data using QALS. **A** 2D intensity (S_0) and polarization (S_1, S_2, S_3) maps for two wavelength indices $m = 38$ and $m = 76$, corresponding to local intensity maxima of sources. **B** reconstructed 2D activations maps and sources Stokes spectral profiles for each of the three factors. Dotted and dashed lines correspond to wavelength indices $m = 38$ and $m = 76$, respectively.

tees, QALS offers a simple, computationally efficient algorithm. We developed recently in [Flamant21] a general framework for solving constrained quaternion-optimization problems, directly in the quaternion domain. These new results pave the way to developing more evolved algorithms presenting convergence guarantees.

6.4 Conclusions

I presented in this chapter my contributions to the domain of constrained multidimensional source unmixing, in three parts.

The first part concerns the *Boolean* factorization of binary-valued matrices and tensors. We proposed an approach based on a post-nonlinear mixture model which is equivalent to the *Boolean* mixture when the factors are exactly binary. This model explicitly accounts for the correlation of the mixed sources and therefore it yields interpretable results in the case of "overlapping" factors, unlike most state-of-the-art algorithms. A simple, yet efficient algorithm to estimate the parameters of this new mixture model was proposed, based on NMF-like multiplicative update rules. An important contribution is represented by the uniqueness conditions for the *Boolean* factorization of binary matrices, which, to the best of our knowledge, are the first of this kind in the binary matrix factorization literature.

In the second part I presented an approach for the on-line unmixing of pushbroom hyperspectral images. We proposed an algorithm (OMDR-ADMM) that outperforms, in terms of convergence speed, the state-of-the-art methods based on multiplicative update rules. OMDR-ADMM, using the minimum dispersion regularization yield results similar to the minimum volume approach, but for a much smaller number of iterations, which demonstrates the interest of this type of penalty for fast on-line hyperspectral unmixing. We proved, on real data, that OMDR-ADMM represents an interesting solution for on-line unmixing of pushbroom hyperspectral images, compliant with the real-time constraints of the industrial wood sorting systems.

The third part of this presentation introduced a new powerful signal processing tool, called quaternion non-negative matrix factorization (QNMF). First, we introduced the QNMF problem from a spectro-polarimetric linear mixing model perspective and demonstrated how it generalizes the usual NMF. Second, we studied the uniqueness conditions for the QNMF problem and revealed the key disambiguating role played by polarization. QNMF is shown to encompass known NMF sufficient uniqueness conditions, while relaxing NMF necessary uniqueness conditions. Finally, we provided a simple yet efficient algorithm in the quaternion domain that effectively solves the QNMF problem.

Part III

Research perspectives

Chapter 7

Research project

I present in this chapter some future research directions organized in two parts: short-term and mid-term perspectives. The first part gathers scientific problems for which reflection is already in progress and which could be addressed in the months to come. The second part contains perspectives with a higher degree of scientific uncertainty that require further investigation.

Globally speaking, this research projet encompasses two components:

- a methodological component, that revolves around tensor models and decompositions;
- an applicative component, that focuses on multisensor data analysis for environnemental surveillance and network analysis.

7.1 Short-term research perspectives

7.1.1 Decompositions of Q-way arrays with categorial entries

This perspective is in the continuity of the work on the *Boolean* decomposition of binary arrays presented in Section 6.1. *Categorical* data arrays are arrays with entries drawn from a J-element alphabet $\mathcal{S} \coloneqq \{s_1, \ldots, s_J\}$. They are encountered in various applications such as recommender systems, where the the ratings take integer values from 1 to 5, or in health care applications, where the results are *positive*, *negative* or *uncertain*. The aim is to find low-parametric decompositions, on a categorical set, of a Q-way array $\boldsymbol{\mathcal{X}}$ of size $(N_1 \times \cdots \times N_Q)$. An important case is represented by the *categorical CPD* that can be formulated as

$$\boldsymbol{\mathcal{X}} \approx \sum_{r=1}^{R} \boldsymbol{p}_1(r) \circ \cdots \circ \boldsymbol{p}_Q(r) = [\![\boldsymbol{P}_1, \cdots, \boldsymbol{P}_Q]\!], \quad \text{with} \quad \{\boldsymbol{P}_q\}_{q=1}^{Q} \in \mathcal{S}. \tag{7.1}$$

When the entries of $\boldsymbol{\mathcal{X}}$ also belong to $\mathcal{S}$, the categorical decomposition intends to find interpretable latent factors underlying the phenomena generating $\boldsymbol{\mathcal{X}}$. If the entries of $\boldsymbol{\mathcal{X}}$ are real-valued, the goal is data compression. In this case, the categorical vectors forming the

columns of $\{\boldsymbol{P}_q\}_{q=1}^{Q}$ can be seen as quantized versions of certain analog-valued data vectors, that belong (or lie close) to the linear low-dimensional subspaces spanned by the columns of the Q unfoldings of $\boldsymbol{\mathcal{X}}$.

Binary arrays A case of particular interest is represented by the *Boolean* CPD (BCPD) of binary tensors, for which the arithmetical sum in (7.1) is replaced by the logical OR operator and $\mathcal{S} = \{0, 1\}$. In [Diop19] we proposed an ADMM-based algorithm for solving the (BCPD) problem for third-order tensors, that we briefly presented in Section 6.1.4. Meanwhile, there are no available uniqueness conditions for BCPD. The study of the uniqueness of Q-order BCPD is one of my short-term research perspectives. We conjectured in [Diop19] that uniqueness conditions for BCPD can be obtained by applying the matrix uniqueness results from [Miron21] to the unfolded versions of the binary tensor, $\text{unfold}_q\boldsymbol{\mathcal{X}}$. By using the Q possible unfoldings, I expect to obtain uniqueness conditions that are less restrictive than those that we proved for the matrix case. Several issues are of high interest in this context:

- How do the uniqueness conditions scale with the tensor order Q?

  It it is well-known that for the real-valued tensors, uniqueness conditions become less restrictive as the tensor order is increasing. As the *Boolean* sum is nonlinear, we expect to observe different behaviors compared to the real case. It would be interesting to see, for example, if there is a maximum value of Q beyond which these conditions are no longer improving.

- Can we formulate more intuitive, "user friendly" conditions as compared to the ones derived in [Miron21] for the matrix case?

  Indeed, for the *Boolean* decomposition of binary matrices, the conditions are based on the analysis of the intersections between the supports of the columns of the factor matrices. This is sometimes complex to check in practice and difficult to apprehend. Moreover, in real-valued tensors literature, most uniqueness conditions are formulated in terms of ranks of the loading matrices $\boldsymbol{P}_q$. For comparison reasons, it would be interesting to have similar conditions in the *Boolean* case. A promising way to achieve this is to introduce a novel notion: the *Boolean dependence/independence* of binary vectors:

  Definition 7.1 (Boolean dependence)**.** *Consider a set of K binary-valued vectors $V = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_K\}$. We will say that a non-zero binary vector $\boldsymbol{u}$ is* Boolean dependent *of the vectors in V if*

  $$\boldsymbol{u} \vee (\bigvee_{k=1}^{K} \boldsymbol{v}_k) = \bigvee_{k=1}^{K} \boldsymbol{v}_k. \tag{7.2}$$

  Otherwise, we will say that $\boldsymbol{u}$ is Boolean independent *of the vectors in V.*

  Definition 7.1 could be used to define *Boolean* notions of rank and Kruskal-rank for a set of binary vectors, similar to the real-valued vector case. It is important to mention that

these new rank notions are different from the notion of *Boolean* rank of a binary matrix, as defined in [Watts01, Miron21]. This is mainly because, in this non-linear framework, the *Boolean* rank of a matrix is different from the *Boolean* rank of its column vectors.

Categorical arrays The general case where the entries of the array belong to a J-element alphabet is more challenging than the binary one and very few results on this topic are available in the literature. Several works extended the well-known independent component analysis (ICA) framework to the case of matrices taking values on *finite fields* [Gutch12, Silva13]. Extensions of the matrix completion algorithms to the case of categorical/discrete data have also been proposed in [Huang13, Cao15, Huo16, Nguyen18].

However, none of the up-mentioned works is solving the problem (7.1) or its matrix version. Some interesting aspects regarding (7.1) are the following:

- How to define the sum and the product operations? If we represent the J values of the alphabet by the integers $\{0, \ldots, J-1\}$, two main strategies can be considered:

  1. Performing addition and multiplication modulo J as it is done *e.g.*, in [Gutch12]. This definition guarantees good algebraic properties but in general it generates non-interpretable results in data analysis applications.
  2. Thresholding the result to the operations. This strategy is less appealing from an algebraic point of view but it facilitates the interpretation of the decomposition results.

- What are the uniqueness properties of the decomposition (7.1)? As for the binary case, restraining the decomposition values to a subset of discrete values should improve uniqueness of the matrix and tensor decomposition. We should be able to derive uniqueness conditions similar to the ones that we have for the binary case. These conditions should depend on the way the operations are defined and on the cardinality of $\mathcal{S}$.

- How to design efficient algorithms for computing the (7.1)? For the algorithmic design, a fundamental question is how to impose the constraint $\{\boldsymbol{P}_q\}_{q=1}^{Q} \in \mathcal{S}$ into the decomposition? In the case of integers, a solution would be to use a quantizer with J levels applied entry-wise on matrices $\boldsymbol{P}_q$. This quantizer could then be transformed (for computational reasons) into a piecewise smooth function by replacing the Heaviside step functions by logistic functions of the form:

$$\Phi(x)_{\gamma,c} = \frac{1}{1 + e^{-\gamma(x-c)}}, \tag{7.3}$$

  where c denotes the center of the sigmoid and γ is a scalar that controls its slope. Various strategies could then be tested to build and minimize cost functions involving (7.3).

Collaborations and funding The binary tensor decomposition is the extension of the work from the Ph.D. of Mamadou Diop, that I co-supervised with David Brie. I am planning to continue the collaboration with David Brie for the first part of this perspective. For the second part (categorical arrays) I am in the process of setting up a collaboration with Rodrigo Cabral Farias (Maître de conférences, Université Côte d'Azur) who has a strong expertise in constrained tensor decompositions and algorithms for quantized measurements estimation. I am also planning to recrute a Master 2 internship student in 2022 to work on this topic.

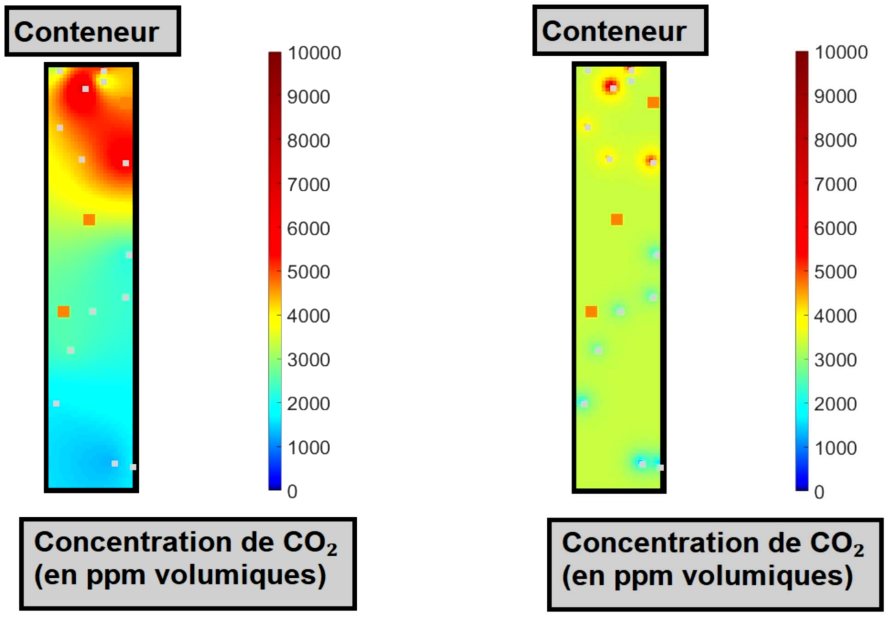
7.1.2 Spatio-temporal cartography of pollutants using tensor completion methods

This project represents the continuation of the Master 2 internship work of Philippe Florés, in collaboration with the INRS, started in 2020. The objective of this collaboration is to produce dynamic pollution maps of working environnements, from a small number of measurements.

INRS disposes of a synchronized network of sensors able to mesure locally the concentration of different pollutants (carbon dioxide, carbon monoxide, NH_3, etc.), with a time sampling step of the order of several seconds. However, deploying a large number of sensors in a working environnement is cumbersome and costly. Thus, the objective is to generate high-resolution, time varying, pollution maps using a small number of spatial measurements (sensors).

If we define a $I \times J$ spatial grid, the problem can be formulated as the estimation of a temporal sequence of matrices $\boldsymbol{X}_k \in \mathbb{R}^{I \times J}$ with $k \in \{1, \dots, K\}$, knowing a subset of their (spatial) entries $\boldsymbol{X}_k(i,j)$, with $(i,j) \in \Omega_s \subseteq \Big\{\{1, \dots, I\} \times \{1, \dots, J\}\Big\}$. In the Master internship thesis from 2020, we proposed a method for pollution cartography based on the kriging interpolation method [Stein12]. The results were globally satisfactory, as they allowed to obtain interpretable dynamic maps of CO_2 with a spatial resolution of 119×23 points, using only 14 sensors. However, the processing was performed independently for each snapshot of the sensor array, without taking into account the strong temporal coherence between successive time samples. This lead, in several scenarios, to temporal discontinuities in the reconstructed pollution cartographies, as illustrated on Figure 7.1.

The objective of this work is to alleviate this problem by explicitly taking into account the temporal coherence of the data in the cartography processing algorithms. To this end, I propose to use the recently developed framework on tensor sampling and reconstruction [Kanatsoulis18, Kanatsoulis19, Prévost20, Zhang20]. From this point of view, pollution cartography can be seen as a tensor completion issue. If we discretize the time dimension $k \in \{1, \dots, K\}$, the problem can be formulated as the reconstruction of a tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I \times J \times K}$ from a subset $\boldsymbol{\mathcal{X}}_s$ of spatially sampled fibers along the time dimension, with the elements $\boldsymbol{\mathcal{X}}_s(i,j,k)$, such that $(i,j) \in \Omega_s$ and $k \in \{1, \dots, K\}$. According to the results of



(a) Estimated CO_2 map at time instant t (b) Estimated CO_2 map at time instant $t+1$

Figure 7.1: Example of temporal discontinuity of the reconstructed pollution maps (real data) when interpolation is performed at each array snapshot independently. The two images represent CO_2 maps estimated at two consecutive time instants using the kriging interpolation method. The small squares indicate the positions of the pollution sensors.

[Kanatsoulis19], $\boldsymbol{\mathcal{X}}$ can be reconstructed from $\boldsymbol{\mathcal{X}}_s$ only if the samples are taken from every row i and column j of the tensor. This is because completely unobserved slabs are impossible to recover. In our case, this means that we need at least $\max(I, J)$ sensors. Given the small number of sensors available, it is impossible to recover the pollution map with an acceptable spatial resolution $I \times J$, by simply using the results of tensors fiber sampling theory.

To tackle this problem, I propose to artificially increase the number of measurements using 2D spatial interpolation of the measurements, at a small set of time instants k. This would allow us to use the recent results on coupled tensor factorization, to achieve super-resolution [Kanatsoulis18, Prévost20]. Thus, the problem can be cast as the coupled factorization of two degraded versions of $\boldsymbol{\mathcal{X}}$: a high temporal resolution tensor $\boldsymbol{\mathcal{X}}_t \in \mathbb{R}^{I_s \times J_s \times K}$, with $I_s \leq I$ and $J_s \leq J$ and a high spatial resolution tensor $\boldsymbol{\mathcal{X}}_t \in \mathbb{R}^{I \times J \times K_t}$, with $K_t \leq K$. The frontal slabs of $\boldsymbol{\mathcal{X}}_t$, $\boldsymbol{\mathcal{X}}_t(:, :, k)$, with $k \in \{1, \ldots, K_t\}$, could be "artificially" reconstructed using a 2D spatial interpolation method, such as the kriging. The "full" resolution tensor $\boldsymbol{\mathcal{X}}$ could then be recovered from $\boldsymbol{\mathcal{X}}_s$ and $\boldsymbol{\mathcal{X}}_t$ by one of the coupled low-rank tensor factorization algorithmes proposed in the up-mentioned papers. In the considered case, the use of the low-rank tensor approximation framework is fully justified, as the pollution diffusion patterns present intrinsic low-rank spatial and temporal structure.

Several aspects deserve investigation:

- What is the best way to perform spatial interpolation? In our previous work we used the kriging to interpolate the spatial data as it provides the best linear unbiased prediction for a Gaussian process. However, if the the Gaussian assumptions are not satisfied, other interpolation methods, such as polynomial kernel interpolation [Rasmussen03], could provide more interesting alternatives.

- What is the optimal placement of the pollution sensors? The coupled factorization approaches require a rank-1 spatial sampling grid. Meanwhile, we observed in the conducted simulations, that a random sampling grid offers more accurate interpolation results by kriging, as it allows measurements of a wider range of spatial correlation lags. Therefore, an idea would be to decompose a random spatial grid in a sum of R rank-1 grids, and reformulate the problem as the coupled decomposition of a set of $R+1$ tensors: $\boldsymbol{\mathcal{X}}_s^{(1)}, \ldots, \boldsymbol{\mathcal{X}}_s^{(R)}, \boldsymbol{\mathcal{X}}_t$. If we model the spatial grid by a binary matrix $\boldsymbol{G} \in \{0, 1\}^{I \times J}$ (where the "1"s indicate the positions of the sensors), the grid decomposition problem boils down to the decomposition of the binary matrix $\boldsymbol{G}$ in a sum of rank-1 binary matrices $\{\boldsymbol{G}^{(r)}\}_{r=1}^{R}$. This is in line with the binary matrix decompositions problems exposed in Section 6.1.

  An interesting question is wether it is possible to specifically design the R grids in order to maximize some reconstruction accuracy criterion.

- Is it possible to include the 2D spatial interpolation step directly in the factorization method? Including the spatial interpolation directly in the algorithm would result in

more efficient parameter updates, allowing (hopefully) optimality claims for the algorithm.

- Can we generate real-time pollution maps? Indeed, the ultimate objective of the project is to produce real-time pollution cartographies of the working environnements. Therefore, a perspective of this work is to study the possibility of adapting this algorithm to an "online" context. The goal is to efficiently update the maps with each new incoming array snapshot. To this end, a similar procedure to the one developed in [Nus20] could possibly be used.

Collaborations and funding This perspective was proposed as a Master 2 internship project. The student who was recruited, Maxime Desplanches, works on this topic since March 2021, under the joint supervision of Philippe Duquenne (INRS) and mine. The internship period ends in September, but the unfinished developments could be pursued in the Ph.D. starting in autumn 2021 (see Section 7.2.3). The internship is fully funded by the lNRS.

7.1.3 Dynamic graph analysis using tensors: application to Software-Defined Networking (SDN)

The starting point for this research perspective is the recently initiated collaboration with the department ISET (Ingénierie des Systèmes Eco-Techniques) from CRAN, on the multivariate analysis of Software-Defined Networking (SDN) data.

Recent works proved that tensors are highly successful in identifying directed graph topologies and tracking dynamic networks [Shen17], for learning from data associated with a collection of graphs [Ioannidis20], or for graph dimensionality reduction and clustering [Fernandes18, An18, Al-Sayouri20]. Inspired by these findings, we aim at developing a tensor approach to analyse and optimize the SDN.

SDN is a set of innovative technologies designed to enable centralized control of network resources, improved programmability and orchestration of these resources, as well as virtualization of these resources by separating them from the physical elements of the network. The aim of these innovations is to simplify network administration and, following the example of what virtualization has achieved in the world of servers, to make the consumption of network resources by applications more flexible. The practical aim of SDNs is to make networks programmable via a centralized controller. Today's switches and routers program their forwarding tables locally, which means that network devices make their own internal decisions about how best to route traffic. These decisions are based on distributed information collected by routing protocols such as OSPF and BGP or protocols such as Spanning Tree. However, these protocols are not very flexible. In order to work together, all the equipment in the network must follow the rules defined by the standards. This leaves little room for creativity or unusual business requirements.

With SDNs, there is a clear separation between the control plane (which defines how an equipment forwards traffic) and the data plane (the part of the switches and routers that actually directs the data). In SDNs, the control plane is placed in a centralized controller that has visibility into the entire network, including the hosts connecting to it, and has a complete view of the network topology. An important advantage of the SDNs is the fact that we have real-time access to a wide range of network metrics such as traffic flows, package loss, energy consumption, etc., which facilitates the application of advanced data analysis methods.

Several tensor-based approaches for SDNs have been proposed over the last years. They make use, in majority, of the tensor framework for data compression or dimensionality reduction *via* Higher Order-SVD, [Kuang16b, Kuang16a, Kaur18] or Tensor Train-SVD [Nie19, Singh19]. The aim of this project is to use graph representations of data networks and the tensor framework for graph analysis, to develop methods for analyzing, learning and optimizing the network topology and traffic.

A data network can be represented as a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ represents the set of nodes (the network equipements) connected together by a set of edges $\mathcal{E}$. Depending on the application, the graph can be undirected/directed and unweighted/weighted. For a dynamic graph, the set of nodes $\mathcal{V}$ and/or the set of edges $\mathcal{E}$ are depending on the time t. The graph connectivity is captured by an adjacency matrix $\boldsymbol{A}(t)$ that is also time dependent. If we consider that the network topology and traffic is stationary over times slots t_k, $k \in \{1, \ldots, K\}$, that the set $\mathcal{V}$ has N elements, we can built the following $N \times N \times K$ tensor $\boldsymbol{\mathcal{A}}$ such that: $\boldsymbol{\mathcal{A}}(:, :, k) = \boldsymbol{A}(t_k)$. This tensor tracks the evolution of the network topology over time. We list next several network issues that can be addressed via the tensor network analysis:

- *Time-evolving network summarization.* The goal here is to summarize the network connectivity over a period of time and to unveil latent structural patterns. In this case, the adjacency matrices $\boldsymbol{A}(t_k)$ are binary-valued and the latent structures are estimated by low-rank decomposition of the binary tensor $\boldsymbol{\mathcal{A}}$. This topic is in line with the categorical tensor factorization perspective presented in Section 7.1.1. The objective is to find "super-nodes" in the dynamic graph, *i.e.*, groups of nodes that exhibit similar connection patterns over time. This could allow the controller to optimize the traffic by adding direct links between the equipments of a super-node, or to simplify network topology by suppressing unused (or little used) connections.

- *Network attacks detection.* This is problem of utmost importance in the nowadays socio-economical context. In the dynamic graph framework, network attacks detection amounts to finding abnormal traffic behaviors in the network. By using valued graphs representations (where the edges represent data traffic flows between equipments), the normal behavior of a network can be characterized by analyzing the factors of the low-rank decomposition of tensor $\boldsymbol{\mathcal{A}}$ over a network operating cycle (hour, day, week, etc.). Strong deviations from this normal behaviors are generally associated to network attacks.

- *Equipment failure prediction.* An equipment failure due *e.g.* to traffic overload, is generally preceded by typical sequences of equipment commutations. Tensor analysis of data traffic and network topology, could allow to identify these typical sequences and to predict the failure of an equipment. This would allow the controller to prevent the network crash by redirecting the traffic.

- *Traffic prediction.* Measuring the traffic between all the network equipments and passing the information to the controller is not always possible, especially for large networks. Thus, using tensor representations of the network and the tensor completion algorithms [Song19] could help filling in the missing information and allow the controller to optimize network traffic.

In addition to the practical interest of the up-mentioned issues for network administration, this topic also presents theoretical and methodological challenges from a data analysis point of view:

- What is the "best" low-rank decomposition of the dynamic adjacency tensor? CPD is highly useful for unveiling network structure but it lacks flexibility. For large and complex networks, CPD may not be the way to go, as it requires a large number of rank-1 terms to describe the underlying tensor structure. This over-factorization often results in collinear loadings, which are difficult to interpret. For complex networks one may want to "zoom in" on some of the super-nodes (clusters) while keeping block representations for the others. Developing hierarchical tensor decompositions [Salmi09] seems to be a more appropriate way to tackle this problem.

- How to include node features in dynamic graph analysis? Beside their connectivity pattern, the network equipments are also characterized by a certain number of feature such as energy consumption, the type of transported data (audio, video, ...), the protocole, etc. When determining the super-nodes, one may want to cluster nodes sharing not only similar connectivities but also similar features. An idea would be to define a node feature matrix of size $N \times F$, where F denotes the number of features. This matrix indicates by "1" the presence of a feature for a node and by "0" its absence. Thus, network summarization could be performed by the coupled factorization of the binary adjacency tensor of size $N \times N \times K$ and of the binary feature matrix. To the best of my knowledge there are no available algorithms in the literature for performing coupled factorization of binary data arrays.

- How to handle time variations of the cardinality of $\mathcal{V}$? In time-evolving networks, new nodes (equipments) may be added to the graph, while other existing nodes may be removed. This a challenging problem for the tensor analysis of a network because it results in adjacency matrices with time-varying sizes and meanings. For the case of node removal, setting the corresponding adjacency column/row of the adjacency matrix to zero may not be the best thing to do, as it could cause numerical problems. Particular

tensor models such as Parafac2 [Kiers99] can handle different size tensor slices, but they have worse uniqueness properties and explanatory power compared to CPD. An interesting prospect is to develop a tensor framework allowing to efficiently handle this scenario.

- How to cope with heterogeneous data? When using weighted graphs to model data network, the value of the traffic flow between nodes is generally used as weight. However, in nowadays networks, describing the communication between two nodes by a traffic value is reductive. We have now access to a wider range of information, such as source/destination ports, routing information, the nature of the data packages, etc. How can we use this information to improve the results of dynamic graph analysis? A solution would be to use an pre-processing step to code the available information in an enhanced homogenous format and then embed it into a real number. The obtained values could be used as the new graph weights. Another option would is to build multiple homogenous graphs and then use coupled tensor factorizations algorithms to analyze them jointly.

- Can we study graph causality *via* tensor analysis? A fundamental question in network analysis is how does an event occurring at one node (equipment failure, traffic peak, etc.) influence another node later on? Causality can be modeled by a directed graph for which the $N \times N$ non-symmetric adjacency matrix $\mathbf{A}$ has non-zero entries $a_{n_1 n_2}$ if the node n_1 has an influence on node n_2, and zero otherwise. As nodes relationships may evolve over time, studying causality could be seen as detecting particular time patterns between nodes in the decomposition of the directed adjacency tensor.

The methodological perspective on dynamic graphs, presented here for the SDN framework, should be easily adaptable to other applications involving dynamic graphs, such as the electroencephalogram (EEG) analysis in neuroscience. In this case, the neurons are the nodes and the synapses represent the edges. The graph is oriented: outgoing edges are synapses to downstream neurons (through the axon) and incoming edges are synapses to upstream neurons (through the dendrites). The dynamic graph issues presented here for data networks have different physical interpretations in neuroscience, but similar mathematical formalisms.

Collaboration and funding This project has two main parts, strongly interconnected. The first part concerns the methodological and algorithmic aspects. I plan to collaborate on this part with Rémy Boyer (Université de Lille) and David Brie (CRAN) - on the tensor developpements - and with Marianne Clausel (Institute Elie Cartan De Lorraine) - on the multivariate analysis of dynamic graphs.

The second part covers the applications. A collaboration with Jean-Philippe Georges and Thierry Divoux (ISET, CRAN), on tensor analysis of SNDs was initiated in 2021. A student in Erasmus Mundus PERCCOM Master will work on this topic for his Master thesis during 2021-2022. Also, I am currently setting up a collaboration with Radu Ranta and Steven Le

Cam, from the research team NEURO (CRAN), on the analysis of dynamic graphs applied to neuroscience.

7.2 Mid-term research perspectives

7.2.1 Low-rank decomposition of quaternion-valued tensors

As illustrated in Section 6.3.1, a quaternion is a hypercomplex number $q = a+b\boldsymbol{i}+c\boldsymbol{j}+d\boldsymbol{k}$, where $a, b, c, d \in \mathbb{R}$ and $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ are imaginary units $\boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = -1$ such that $\boldsymbol{ij} = \boldsymbol{k}$, $\boldsymbol{ij} = -\boldsymbol{ji}$ and $\boldsymbol{ijk} = -1$. The set of quaternions $\mathbb{H}$ forms a 4-dimensional normed division algebra over the real numbers $\mathbb{R}$ with canonical basis $\{1, \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$.

Quaternions, and quaternion matrix calculus has been successfully used in signal processing and data analysis for more than two decades (see *e.g.* [Sangwine98, Le Bihan04b, Miron06, Chen19, Flamant20]. Most applications use quaternions to model multicomponent data such as color images, polarized signals or multi-channel data and store them in bi-dimensional arrays (matrices). The increase in data complexity over the last years calls for the use of Q-way arrays of quaternions ($Q > 2$) to account for the multiple diversities. For example, quaternion-valued tensors have been used recently for color videos and image recovery [Miao20]. However, unlike for quaternion matrices, the literature and the mathematical tools for quaternion-valued tensors are very limited. The existing works on this topic consider a quaternion tensor as a Q-way array of quaternions without any underlying tensor space structure, and mainly deals with the matrix unfoldings of these arrays [Miao20, He17]. This lack of results on quaternion tensors is mainly caused by the fact that quaternion product is non-commutative, which makes the manipulation of quaternion quantities an arduous task. Some definitions of quaternion tensor product have been proposed in the context of quantum mechanics (see [Razon91] and references therein), but they are not transposable to the signal processing domain. The objective of this research perspective is to introduce a solid framework for the use of quaternion-valued tensors for signal and data analysis, and to develop efficient algorithms for computing the quaternion tensor decompositions.

A Q-order quaternion tensor can be seen as an element of the outer (tensor) product of Q quaternion vector spaces: $\mathbb{H}^{N_1} \circ \cdots \circ \mathbb{H}^{N_Q}$. A question of central importance is how to define these vector spaces $\mathbb{H}^{N_1}, \cdots, \mathbb{H}^{N_Q}$, in order to have nice properties of the quaternion tensor space. A fundamental property of the tensor product [Comon14] is that, if $\boldsymbol{x}_1 \circ \cdots \circ \boldsymbol{x}_Q$ is an element of $\mathbb{H}^{N_1} \circ \cdots \circ \mathbb{H}^{N_Q}$, and $\alpha_1, \cdots, \alpha_Q$ are scalars from the associated field $\mathbb{K}$, such that $\alpha_1\alpha_2 \cdots \alpha_Q = 1$, than $(\boldsymbol{x}_1 \circ \cdots \circ \boldsymbol{x}_Q)$ and $(\alpha_1\boldsymbol{x}_1 \circ \cdots \circ \alpha_Q\boldsymbol{x}_Q)$ represent the same objects. Its well established now that $\mathbb{H}^N$ is a right quaternion vector space [Rodman14b] over $\mathbb{H}$. However, by choosing $\mathbb{K} = \mathbb{H}$, the property of the tensor product mentioned above does not hold because of the noncommutativity of quaternion multiplication. A reasonable choice seems to be $\mathbb{K} = \mathbb{R}$, meaning that we consider the vector spaces $\mathbb{H}^{N_q}$ over $\mathbb{R}$, $(\mathbb{H}^{N_q}, \mathbb{R})$. Knowing that the dimension of $(\mathbb{H}^{N_q}, \mathbb{R})$ is $4N_q$, the vector space $\mathbb{H}^{N_1} \circ \cdots \circ \mathbb{H}^{N_Q}$ will be of

dimension $4^Q N_1 \cdots N_Q$, and an element of it can be expressed as:

$$\boldsymbol{\mathcal{X}} = \sum_{n_1}^{4N_1} \cdots \sum_{n_Q}^{4N_Q} x_{n_1 \cdots n_Q} \boldsymbol{e}_1(n_1) \circ \cdots \circ \boldsymbol{e}_Q(n_Q), \tag{7.4}$$

where $\{\boldsymbol{e}_q(n_q)\}_{n_q=1}^{4N_Q}$ is a basis of the $\mathbb{H}^{N_q}$ vector space, and $x_{n_1 \cdots n_Q}$ are the real coordinates of $\boldsymbol{\mathcal{X}}$ with respect to these bases. Multi-linearity also holds for this definition of quaternion tensors. That is, if we consider a change of basis in the vector spaces $\mathbb{H}^{N_q}$, such that $\widetilde{\boldsymbol{e}}_q(n_q) = \boldsymbol{M}^{(q)} \boldsymbol{e}_q(n_q)$ for some given matrices $\{\boldsymbol{M}^{(q)}\}_{q=1}^{Q}$, then the new coordinates of tensor $\boldsymbol{\mathcal{X}}$ can be expressed as a function of the original ones as:

$$\widetilde{x}_{n_1' \cdots n_Q'} = \sum_{n_1}^{4N_1} \cdots \sum_{n_Q}^{4N_Q} M^{(1)}_{n_1' n_1} \cdots M^{(Q)}_{n_Q' n_Q} x_{n_1 \cdots n_Q}. \tag{7.5}$$

This is a direct consequence of the commutativity between quaternions and real numbers. However, in most signal processing and data analysis applications, we aim at decomposing a tensor in a sum of rank-1 terms, leading in this case to the *Quaternion CPD (QCPD)*:

$$\boldsymbol{\mathcal{X}} = \sum_{r=1}^{R} \boldsymbol{p}_1(r) \circ \ldots \circ \boldsymbol{p}_Q(r) = [\![\boldsymbol{P}_1, \ldots, \boldsymbol{P}_Q]\!], \tag{7.6}$$

where $\{\boldsymbol{P}_q\}_{q=1}^{Q}$ are quaternion-valued matrices. Several interesting questions concerning (7.6) require investigation:

- What is the rank R of the decomposition (7.6)? Two quaternion vectors can be linearly dependent over $\mathbb{H}$ but linearly independent over $\mathbb{R}$. Is there a *generic* rank for a $N_1 \times \cdots \times N_Q$ quaternion-valued tensor, as it is the case for complex-valued tensors [Comon09]? If this is the case, how can this rank be computed?

- An important issue for the interpretability of the decomposition result is its uniqueness. If a minimum rank-R decomposition of $\boldsymbol{\mathcal{X}}$ exists, is this decomposition unique? What are the, in this case, the uniqueness conditions for QCPD? Does the non-commutativity of quaternion product allow to relax the uniqueness conditions compared to the real case?

- A fundamental problem in data analysis is the low-rank approximation of a tensor:

$$\boldsymbol{\mathcal{X}} \approx \sum_{r=1}^{R} \boldsymbol{p}_1(r) \circ \ldots \circ \boldsymbol{p}_Q(r). \tag{7.7}$$

  It was shown that, for real-valued tensors, the low-rank approximation problem is ill-posed [De Silva08]. Thus, real tensors of order 3 or higher can fail to have best rank-R approximations, while this is not the case for complex-valued tensors [Qi20]. Given the hypercomplex nature of quaternions, a legit question is whether quaternion tensors

admit a best rank-R approximation, and if so, under which conditions?

- If this best rank-R approximation exists, how can it be computed? Experience shows that, even if the theoretical existence conditions are not met, it is still possible to obtain good low-rank tensor approximation, with high practical utility. Therefore, we aim at developing efficient algorithms for computing the QCPD decomposition of a quaternion tensor and apply them to real data and signal analysis. This is not a trivial task, because the real/complex CPD estimation strategies cannot directly apply to quaternion tensors. Indeed, most CPD estimation algorithms, and in particular ALS (the most popular one) are based on the matrix unfoldings of the tensor, defined as:

$$\text{unfold}_q \boldsymbol{\mathcal{X}} = \boldsymbol{P}_q \left(\boldsymbol{P}_Q \odot \cdots \odot \boldsymbol{P}_{q+1} \odot \boldsymbol{P}_{q-1} \odot \cdots \odot \boldsymbol{P}_1 \right)^{\mathsf{T}} . \tag{7.8}$$

  Because of the non-commutativity of quaternion multiplication, expression (7.8) may hold for $q = Q$, but, in general, it does not hold for all the values of q. How can this problem be solved in order to design efficient ALS-like algorithms? Is it possible to define other tensor unfolding strategies having the desired properties? For the third order tensor, this problem could be overcome by considering joint diagonalization schemes of the tensor "slices" [De Lathauwer06b].

  Another promising algorithmic lead is represented by the recent advances in the domain of convex quaternion optimization [Flamant21]. This allows to formulate the QCPD problem as an optimization problem and solve it directly in the quaternion domain. Furthermore, constraints such as *pure* quaternion or *non-negativity* [Flamant20] of quaternion factors can be easily imposed and efficiently solved in this framework.

- In a large family of applications only one of the factor matrices $\boldsymbol{P}_q$ is quaternion-valued while the others are real. This is the case, for example, in polarized source unmixing, where the hyperpectral signatures of the sources can be modeled by quaternion sequences, while the other diversities (space, time, etc.) are real-valued. This induces an important simplification of the tensor model, as in this case, commutativity is preserved. It would be interesting to study how this affects the QCPD uniqueness, rank as well as the QCPD estimation algorithms.

Another tensor decomposition largely used in signal and image processing applications is the *Higher-Order Singular Value Decomposition (HOSVD)* [De Lathauwer00], that estimates orthonormal bases in the linear spaces associated with the Q modes of the tensor. The extension of this decomposition to the quaternion case is a problem that deserves attention. As the quaternion SVD is a mature method, the HOSVD of a quaternion-valued tensor should be easier to compute compared to its CPD. The quaternion HOSVD could eventually be used as a pre-processing step for quaternion CPD, as it is done in the real-valued case [Bro98b].

The main application targeted by this quaternion tensor framework is the polarized data analysis in hyperspectral imaging. For application on real data we will use the pushbroom acquisition system (with polarized filters) at our disposal at CRAN as well as the recently acquired polarized microscope for analyzing biological samples.

Another intended application is the compression of convolutional layers in quaternion neural networks (QNN). QNN has recently become an active field of research [Parcollet20], in particular for multicomponent data processing (color image, audio, etc.). The convolutional layers of a QNN are Q-way arrays of quaternions. It has been recently shown [Lebedev14], in the case of real-valued neural networks, that the convolutional layers can be highly compressed using CPD. As a consequence, the required computational power can be strongly reduced while preserving the network performance, making these NN implementable on embedded systems. It would be interesting to see how this type of approach applies to QNN.

Collaborations and funding I have already initiated a collaboration on this topic with Xavier Luciani (Maître de conférences, Université de Toulon) - an expert in algorithms for tensor decompositions - and with Julien Flamant (Chargé de recherche CNRS, CRAN) - an expert in quaternions. For the QNN layers compression, I plan to collaborate with Konstantin Usevich (Chargé de recherche CNRS, CRAN) and with Titouan Parcollet (Maître de conférences, Université d'Avignon). To financially support this project I am considering to apply for a PEPS (Projets Exploratoires Premier Soutien) grant for 2022.

7.2.2 Tensors methods for frequency diverse array (FDA) processing

The project presented in this section is a continuation of my collaboration with Kainam Thomas Wong (Beihang University, China), initiated in 2011. It is, at the same time, in line with a constant focus of my research activity since my Ph.D.: the sensor array processing.

The "classical" phased-array (PA) has been successfully used since the late 1930s in a wide range of applications (radar, electronic warfare, radio astronomy, airport safety, etc.), because of its capability to electronically steer the beam with high effectiveness. For a PA, the beam steering is usually implemented by applying a linear phase shift across the array elements. However, PA has a limitation in that the beam steering is fixed in an angle for all ranges. Controlling range-dependent energy distribution has become an increasingly important requirement in many applications nowadays. To overcome this drawback, a flexible beam scanning array, referred to as frequency diverse array (FDA), was introduced in [Antonik06]. The basic FDA scheme consists in an uniform linear array (ULA) composed of M transmitting elements, radiating narrow-band signals:

$$s_m(t) = S_m(t) \exp(j2\pi f_m t), \tag{7.9}$$

where $S_m(t)$ is the base-band waveform for the mth sensor. The f_m frequency radiated by the Mth element is such that:

$$f_m(t) = f_0 + (m-1)\Delta f, \quad m = 1, \ldots, M, \tag{7.10}$$

where f_0 is a fundamental frequency and Δf is a small frequency increment. Assuming a point target in space, the generated FDA beam pattern at the target is a function $P_{FDA}(t, \Delta f, r, \theta)$, where r and θ represent the range and direction (angle) of the target, respectively. Inspired by this novel antenna setup, FDA-based applications have been proposed over the last few years, mostly in the field of radar and navigation (see *e.g.,* [Wang16b, Cui18]).

In the classical FDA radar setup, at a given far-field point target $\{\theta, r\}$, the phase difference $\Delta\varphi_m$ between the received signals from first and the mth sensor of the FDA, can be expressed as ([Wang14]):

$$\Delta\varphi_m = -\frac{2\pi\ f_0\ m\ d\ \sin\theta}{c_0} + \frac{2\pi\ m\ \Delta f\ r}{c_0} - \frac{2\pi\ m^2\ \Delta f\ d\ \sin\theta}{c_0}, \tag{7.11}$$

where c_0 represents the speed of light and d is the inter-sensor spacing. The first term in (7.11) is the classical PA factor. The second term is important because is shows that the phase-shift pattern depends on the range and frequency increment. In FDA systems, the third term is small and can be generally ignored.

As one can see in expression (7.11), a major problem in FDA radar is that the phase-shifts associated to the range and to angle are coupled. This results in ambiguities in the beamformer output, *i.e.,* we may get multiple angle-range estimate pairs for one target. To address this problem, several approaches have been proposed, such as the use of two transmitted pulses with zero and non-zero frequency increments [Wang13b], the use of of two subarrays with different carrier frequency increments [Wang14], or the use of random [Liu16] and co-prime [Qin16] frequency increments. However, for these methods the computation complexity and localization accuracy rely on the search/predefined grid. To overcome this drawback, tensor-based methods represent an interesting alternative thanks to their efficient use of diversities and to their identifiability properties. Several FDA schemes using the tensor framework have been introduced recently. In [Cui18], a third-order CPD data model is built for bistatic FDA-MIMO radar by leveraging the spatial diversities offered by the transmitting and receiving arrays as well as the time (pulse diversity). This method allows joint direction of departure (DOD), direction of arrival (DOA) and range estimation. For monostatic FDA-MIMO radar, a similar method was reported in [Xu20]. Another tensor decomposition, the HOSVD, was used for subspace-based multiple-target 3D localization [Li18b]. It makes use of planar frequency diverse subarrays in monostatic MIMO radar. In [Li18a], electromagnetic wave polarization is used in bistatic MIMO radar to create an additional diversity. The transmitting array is composed of ordinary elements while the receiving array is formed of crossed-dipoles. A third-order CPD model for the data is then constructed, but the parame-

ters estimation problem is finally solved using an ESPRIT-like algorithm.

Nevertheless, FDA is an emerging technology and there is still a lot of room for new antenna configurations and algorithmic developments, for improving localization accuracy and reducing computational complexity. I present next several research ideas based on the tensor formalism.

The majority of existing FDA radar setups consider that the target is fixed, meaning that its location (angle and range) does not change between the transmitted pulses. However, in fast moving-target applications, this assumption is no longer valid and the up-mentioned methods can no longer be utilized. The existing solutions for this scenario use complex closed-loop beamforming techniques to compensate the target movement [Wang16a]. Therefore, it is worth investigating wether tensor-based models can provide more efficient algorithms for moving target localization. The idea would be to be able to estimate the range and angle for multiple moving targets using a single transmitted pulse, in order to avoid considering time-dependent parameters. A solution could be the use of multiple collocated phased-subarrays with a frequency increment between them. In practice, this scheme could be implemented, for example, as an array of M elements, where each element is composed of 3 collocated crossed-dipoles. A subarray is then composed of the M identically oriented dipoles. The use of orthogonally polarized transmitted signals, as well as orthogonal wave-forms should alleviate the electromagnetic coupling problems between the array elements. Thus, the transmitting steering vector of size $3M$ could be expressed as $\boldsymbol{a}_t(\theta, \Delta f, r) = \boldsymbol{a}_{PA}(\theta) \otimes \boldsymbol{a}_{\Delta f}(\Delta f, r)$, where $\boldsymbol{a}_{PA}(\theta)$ is the classical steering vector of a phased-array and $\boldsymbol{a}_{\Delta f}(\Delta f, r)$ is a three-element vector containing the phase-shifts due to the frequency increments between the three dipoles of an element. The receiving steering vector $\boldsymbol{a}_r(\theta)$ in this case is the same as $\boldsymbol{a}_{PA}(\theta)$. The joint steering vector can then be expressed as:

$$\boldsymbol{a}(\theta, \Delta f, r) = \boldsymbol{a}_t(\theta, \Delta f, r) \otimes \boldsymbol{a}_r(\theta) = \boldsymbol{a}_{PA}(\theta) \otimes \boldsymbol{a}_{\Delta f}(\Delta f, r) \otimes \boldsymbol{a}_{PA}(\theta). \tag{7.12}$$

Supposing that K target signals impinge upon the receiving array, the received data, after the matched filtering, can be expressed as a third-order CPD model of rank K. The different steering vectors could then be estimated by CPD decomposition of the data. The proposed setup presents several interesting advantages over the existing approaches:

- it only requires a single pulse for angle-range estimation, which makes it suitable for tracking moving targets;
- it directly performs decoupling of the range r and the angle θ, and does not require a disambiguation step. Indeed, unlike in the existing FDA tensor-based models, it can be observed from equation (7.12) that the two parameters of interest are present in separate factors of the CPD model and are therefore estimated independently;
- the resulting CPD data model presents a symmetric structure, that can be taken into account in the decomposition algorithm to improve method accuracy.

The proposed approach can be easily adapted to planar array for estimating 2D angles (azimuth and elevation) of the targets. Other diversities, such as pulse polarization, can be utilized to develop similar tensor approaches in FDA monostatic/bistatic MIMO radar. We can also imagine using two different FDA radar systems sharing a common diversity, *e.g.,* the frequency increment diversity. This would lead to interesting tensor-tensor or tensor-matrix factorization problems, and result (hopefully) in improved estimation accuracy and parameters identifiability. All these research directions need to be extensively tested in simulations to asses their advantages/drawbacks.

Collaborations and funding I initiated a collaboration in 2019 with Kainam Thomas Wong (Beihang University, China), on methods for FDA signal processing. A Master student from Ecole Centrale de Pékin was planned to undertake a visit to CRAN in 2020 for his Master thesis, to work on this topic under our joint supervision. However, because of the Covid-19 crisis, this project was prematurely stopped. We have resumed this collaboration this year and, if the global health situation allows, a Master student from the Chinese university will take over this project in 2022. His subsistence, accommodation and travel expenses to France will be fully covered by Ecole Centrale de Pékin.

7.2.3 Joint analysis of multivariate, heterogenous data for pollutant exposure risk evaluation in working environnements

This topic will be developed in the Ph.D. thesis that I co-promote with Philippe Duquenne (Head of ASTEC laboratory, INRS, Nancy). The ultimate objective of this project is to develop a method for estimating the exposure of workers to different type of pollutants, using data acquired by different types of sensors deployed in the working environnement.

The first phase of this project is to produce reliable dynamic (time dependent) pollution cartographies of the observed industriel environnement. This goes through the deployment of sensors capable of measuring the concentration in the air of chemical compounds such as carbon dioxide, carbon monoxide, NH_3, ethanol, dietyl-ketone, etc. INRS also has at its disposal sensors able to measure the pollution produced by micro-particles. Such a sensor outputs a vector of measurements containing the distribution of particles with respect to their sizes. All these sensors produce highly accurate measurements of one particular substance in one point in space. This information, however, is insufficient to estimate the spatial distribution of the pollutants. Because of their high cost and size, deploying spatial networks of such sensors is unrealistic. Instead, we propose to use spatial networks of low-cost compact sensors, such as the portable spectrometers. These sensor are less accurate but they are sensitive to a wide class of pollutants (chemical compounds and micro-particles). A first question is how to efficiently merge these low-quality spatially spread non-specific measurements with high quality spatially localized substance-specific measurements, in order to produce a high

quality fine grid cartography of the pollutants. This will probably require the development of methods combining spatial interpolation and source separation algorithms. A promising lead in this direction is represented by the coupled data decomposition algorithms developed in the last decade [Acar15, Farias16, Chatzichristos18]. However, these methods cannot be directly applied to the problem at hand because of the particular nature of the acquisition scheme.

Another interesting question is how to include the information provided by the particle sensors into these algorithms. Indeed, some industrial processes generate both gas and particle pollutants (this is, for example, the case of cheese rubbing operation in the dairy industry). In this case, the signal delivered by the particle detector is correlated to the gas signal, and therefore, taking it into account seems like a clever thing to do.

The second phase concerns the validation of the developed method. INRS has an experimental chamber with controlled environnemental parameters (temperature, humidity, etc.). For such conditions, there exists mathematical models that allow to predict the dispersion pattern of a gas in space and time. We will use these mathematical predictions as a form of "ground truth" to validate our method. We will compare the cartography estimated by our algorithms using gas concentration measurements to the maps generated by the gas dispersion models. A strong correlation between the two will indicate that the developed methods produce reliable results.

The last phase of this project is dedicated to the analysis of pollutant exposure risk for workers. To this end, we will equip workers with mobile pollution sensors and indoor localization devices. INRS already disposes of an Ultra WideBand (UWB) localization system that could be used for this task. A 3D scanner, recently acquired by INRS, will be utilized to digitize the indoor environnement and thus improve localization accuracy. The objective is to estimate the pollutant exposure of a person during an industrial cycle, using only the pollutant maps (generated by the algorithms developed in the first phase of the project) and the person space-time trajectory. The mobile pollutant sensors will be used as "ground truth" for the validation of the exposure estimation method. The ultimate goal is to be able to estimate workers exposure risks without any intrusion, by simply using the surveillance cameras to estimate their trajectories.

This project is naturally in line with the actual digital transformation of manufacturing/production and related industries, known as the *Industry 4.0*. In the future, we expect the industrial machines to be natively equipped with various types of sensors and communicate directly with a data processing center for real-time evaluation of pollution risk exposure.

Collaborations and funding This research work will be done in collaboration with Philippe Duquenne (INRS) and his team. A Ph.D. thesis on this topic, that I will co-supervise, will start in October 2021 and will be fully funded by INRS.

Bibliographie

[Acar15] Acar E., Bro R., and Smilde A.K. Data fusion in metabolomics using coupled matrix and tensor factorizations. *Proceedings of the IEEE*, 103(9):1602–1620, 2015.

[Al-Sayouri20] Al-Sayouri S., Gujral E., Koutra D., Papalexakis E.E., and Lam S.S. t-pine: Tensor-based predictable and interpretable node embeddings. *Social Network Analysis and Mining*, 10(1):1–11, 2020.

[An18] An J., Zhang X., Zhou H., and Jiao L. Tensor-based low-rank graph with multimanifold regularization for dimensionality reduction of hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 56(8):4731–4746, 2018.

[Andersson00] Andersson C.A. and Bro R. The N-way toolbox for MATLAB. *Chemometrics and intelligent laboratory systems*, 52(1):1–4, 2000.

[Antonik06] Antonik P., Wicks M.C., Griffiths H.D., and Baker C.J. Frequency diverse array radars. In *2006 IEEE Conference on Radar*, pages 3–pp. IEEE, 2006.

[Belohlavek10] Belohlavek R. and Vychodil V. Discovery of optimal factors in binary data via a novel method of matrix decomposition. *Journal of Computer and System Sciences*, 76(1):3–20, 2010.

[Berry77] Berry H.G., Gabrielse G., and Livingston A. Measurement of the Stokes parameters of light. *Applied optics*, 16(12):3200–3205, 1977.

[Berry07] Berry M.W., Browne M., Langville A.N., Pauca V.P., and Plemmons R.J. Algorithms and applications for approximate nonnegative matrix factorization. *Computational statistics & data analysis*, 52(1):155–173, 2007.

[Bertsekas99] Bertsekas D.P. *Nonlinear programming*. Athena Scientific, second edition, 1999.

[Bienvenu79] Bienvenu G. and Kopp L. Principe de la goniométrie passive adaptative. In *Proc. 7° Colloque sur le traitement du signal et des images, GRETSI*. Nice, France, 1979.

[Bioucas-Dias12] Bioucas-Dias J.M., Plaza A., Dobigeon N., Parente M., Du Q., Gader P., and Chanussot J. Hyperspectral unmixing overview: geometrical, statistical, and sparse regression-based approaches. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(2):354–379, 2012.

[Born80] Born M. and Wolf E. *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Pergamon Press, 6th edition, 1980.

[Boyd04] Boyd S. and Vandenberghe L. *Convex optimization*. Cambridge University Press, 2004.

[Boyd11] Boyd S., Parikh N., Chu E., Peleato B., Eckstein J., et al. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning*, 3(1):1–122, 2011.

[Brenner51] Brenner J.L. Matrices of quaternions. *Pacific Journal of Mathematics*, 1(3):329–335, 1951.

[Brie11] Brie D., Miron S., Caland F., and Mustin C. An uniqueness condition for the 4-way CANDECOMP/PARAFAC model with collinear loadings in three modes. In *2011 36th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4108–4111. IEEE, Prague, Czech Republic, May 2011.

[Brie14a] Brie D., Klotz R., Miron S., Moussaoui S., Mustin C., Bécuwe P., and Grandemange S. Joint analysis of flow cytometry data and fluorescence spectra as a non-negative array factorization problem. *Chemometrics and Intelligent Laboratory Systems*, 137:21–32, 2014.

[Brie14b] Brie D., Miron S., Bécuwe P., and Grandemange S. A non-negative multilinear block tensor decomposition approach to flow cytometry data analysis. In *2014 IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*. IEEE, Reims, France, September 2014.

[Bro98a] Bro R. *Multi-way Analysis in the Food Industry : Models, Algorithms, and Applications*. Ph.D. thesis, Royal Veterinary and Agricultural University, Danemark, 1998.

[Bro98b] Bro R. and Andersson C.A. Improving the speed of multiway algorithms. *Chemometr. Intell. Lab.*, 42(1-2):105–113, August 1998.

[Bro99] Bro R., Sidiropoulos N., and Giannakis G. A fast least squares algorithm for separating trilinear mixtures. In *Proc. Int. Workshop Independent Component Analysis and Blind Signal Separation (ICA)*. Aussois, France, Jan. 1999.

[Bro09] Bro R., Harshman R.A., Sidiropoulos N.D., and Lundy M.E. Modeling multi-way data with linearly dependent loadings. *J. Chemometr.*, 23(7-8):324–340, July-August 2009. Special Issue: In Honor of Professor Richard A. Harshman.

[Brosseau98] Brosseau C. *Fundamentals of polarized light: a statistical optics approach*. Wiley-Interscience, 1998.

[Bucak09] Bucak S.S. and Gunsel B. Incremental subspace learning via Non-negative Matrix Factorization. *Pattern recognition*, 42(5):788–797, 2009.

[Caland11] Caland F., Miron S., Brie D., and Mustin C. A CANDECOMP/PARAFAC approach to the estimation of environmental pollutant concentrations using biosensors. In *2011 IEEE Statistical Signal Processing Workshop (SSP)*, pages 801–804. IEEE, Nice, France, June 2011.

[Caland12] Caland F., Miron S., Brie D., and Mustin C. A blind sparse approach for estimating constraint matrices in Paralind data models. In *2012 20th European Signal Processing Conference (EUSIPCO)*, pages 839–843. IEEE, Bucharest, Romania, August 2012.

[Caland13] Caland F. *Décomposition tensorielle de signaux luminescents émis par des biosenseurs bactériens pour l'identification de Systèmes Métaux-Bactéries*. Ph.D. thesis, Universié de Lorraine, 2013.

[Cao15] Cao Y. and Xie Y. Categorical matrix completion. In *2015 IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pages 369–372. IEEE, 2015.

[Carroll70] Carroll J.D. and Chang J.J. Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckardt-Yong" decomposition. *Psychometrika*, 35(3):283–319, September 1970.

[Carteret09] Carteret C., Dandeu A., Moussaoui S., Muhr H., Humbert B., and Plasari E. Polymorphism studied by lattice phonon raman spectroscopy and statistical mixture analysis method: Application to calcium carbonate

polymorphs during batch crystallization. *Crystal Growth and Design*, 9:807–812, 2009.

[Chatzichristos18] Chatzichristos C., Davies M., Escudero J., Kofidis E., and Theodoridis S. Fusion of EEG and fMRI via soft coupled tensor decompositions. In *2018 26th European Signal Processing Conference (EUSIPCO)*, pages 56–60. IEEE, 2018.

[Chen19] Chen Y., Xiao X., and Zhou Y. Low-rank quaternion approximation for color image processing. *IEEE Transactions on Image Processing*, 29:1426–1439, 2019.

[Cichocki14] Cichocki A. Era of big data processing: A new approach via tensor networks and tensor decompositions. *CoRR*, 2014.

[Comon09] Comon P., Berge J.M.F.T., DeLathauwer L., and Castaing J. Generic and typical ranks of multi-way arrays. *Linear Algebra Appl.*, 430(11–12):2997–3007, June 2009. doi: 10.1016/j.laa.2009.01.014.

[Comon14] Comon P. Tensors: a brief introduction. *IEEE Sig. Proc. Magazine*, 31(3):44–53, May 2014. Special issue on BSS.

[Conway03] Conway J.H. and Smith D.A. *On quaternions and octonions: their geometry, arithmetic, and symmetry*. AK Peters, 2003.

[Cui18] Cui C., Xu J., Gui R., Wang W.Q., and Wu W. Search-free DOD, DOA and range estimation for bistatic FDA-MIMO radar. *IEEE Access*, 6:15431–15445, 2018.

[De Lathauwer00] De Lathauwer L., De Moor B., and Vandewalle J. A multilinear singular value decomposition. *SIAM J. Matrix Ana. Appl.*, 21(4):1253–1278, April 2000.

[De Lathauwer06a] De Lathauwer L. A link between canonical decomposition in multilinear algebra and simultaneous matrix diagonalization. *SIAM Journal on Matrix Analysis*, 28(3):642–666, 2006.

[De Lathauwer06b] De Lathauwer L. A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization. *SIAM journal on Matrix Analysis and Applications*, 28(3):642–666, 2006.

[De Lathauwer08a] De Lathauwer L. Decompositions of a higher-order tensor in block terms - Part I: Lemmas for partitioned matrices. *SIAM J. Matrix Anal. Appl.*, 30(3):1022–1032, 2008.

[De Lathauwer08b] De Lathauwer L. Decompositions of a higher-order tensor in block terms - Part II: Definitions and uniqueness. *SIAM J. Matrix Anal. Appl.*, 30(3):1033–1066, 2008.

[De Silva08] De Silva V. and Lim L.H. Tensor rank and the ill-posedness of the best low-rank approximation problem. *SIAM Journal on Matrix Analysis and Applications*, 30(3):1084–1127, 2008.

[deAlmeida08a] de Almeida A., Favier G., and Mota J. A constrained factor decomposition with application to MIMO antenna systems. *IEEE Trans. Signal Processing*, 56(6):2429–2442, 2008.

[deAlmeida08b] de Almeida A., Favier G., and Mota J. Constrained tensor modeling approach to blind multiple-antenna CDMA schemes. *IEEE Trans. Signal Processing*, 56:2417–2428, 2008.

[Diop19] Diop M., Miron S., Souloumiac A., and Brie D. Boolean CP decomposition of binary tensors: uniqueness and algorithm. In *2019 44th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, Brighton, UK, May 2019.

[Dobigeon16] Dobigeon N., Altmann Y., Brun N., and Moussaoui S. Linear and nonlinear unmixing in hyperspectral imaging. In *Data Handling in Science and Technology*, volume 30, pages 185–224. Elsevier, 2016.

[Farias16] Farias R.C., Cohen J.E., and Comon P. Exploring multimodal data fusion through joint decompositions with flexible couplings. *IEEE Transactions on Signal Processing*, 64(18):4830–4844, 2016.

[Fernandes18] Fernandes S., Fanaee-T H., and Gama J. Dynamic graph summarization: a tensor decomposition approach. *Data Mining and Knowledge Discovery*, 32(5):1397–1420, 2018.

[Ferrara83] Ferrara Jr. E.R. and Parks T.M. Direction finding with an array of antennas having diverse polarizations. *IEEE Trans. Antennas Propagat.*, AP-31(2):231–236, March 1983.

[Flamant17] Flamant J., Le Bihan N., and Chainais P. Spectral analysis of stationary random bivariate signals. *IEEE Transactions on Signal Processing*, 65(23):6135–6145, 2017. ISSN 1053587X. doi: 10.1109/TSP.2017.2736494.

[Flamant18] Flamant J., Chainais P., and Le Bihan N. A complete framework for linear filtering of bivariate signals. *IEEE Transactions on Signal Processing*, 66(17):4541–4552, Sept 2018. ISSN 1053-587X. doi: 10.1109/TSP.2018.2855659.

[Flamant20] Flamant J., Miron S., and Brie D. Quaternion non-negative matrix factorization: definition, uniqueness and algorithm. *IEEE Transactions on Signal Processing*, 68(1):1870–1883, 2020. doi: 10.1109/TSP.2020.2974651. URL http://dx.doi.org/10.1109/TSP.2020.2974651.

[Flamant21] Flamant J., Miron S., and Brie D. A general framework for constrained convex quaternion optimization. *IEEE Transactions on Signal Processing [submitted]*, 2021. URL https://arxiv.org/abs/2102.02763.

[Fu15] Fu X., Ma W.K., Huang K., and Sidiropoulos N.D. Blind separation of quasi-stationary sources: Exploiting convex geometry in covariance domain. *IEEE Transactions on Signal Processing*, 63(9):2306–2320, 2015.

[Fu16] Fu X., Huang K., Yang B., Ma W.K., and Sidiropoulos N.D. Robust volume minimization-based matrix factorization for remote sensing and document clustering. *IEEE Transactions on Signal Processing*, 64(23):6254–6268, 2016.

[Fu18] Fu X., Huang K., and Sidiropoulos N.D. On identifiability of Non-negative Matrix Factorization. *IEEE Signal Processing Letters*, 25(3):328–332, 2018.

[Gil16] Gil J. and Ossikovski R. *Polarized Light and the Mueller matrix approach*. CRC Press, 2016. ISBN 9781482251562.

[Gong09] Gong X., Liu Z., Xu Y., and Ahmad M.I. Direction-of-arrival estimation via twofold mode-projection. *Signal Process.*, 89(5):831–842, May 2009.

[Grippo00] Grippo L. and Sciandrone M. On the convergence of the block nonlinear gauss–seidel method under convex constraints. *Operations research letters*, 26(3):127–136, 2000.

[Guan12] Guan N., Tao D., Luo Z., and Yuan B. On-line Non-negative Matrix Factorization with robust stochastic approximation. *IEEE Transactions on Neural Networks and Learning Systems*, 23(7):1087–1099, 2012.

[Guo08] Guo X., Miron S., and Brie D. Identifiability of the PARAFAC model for polarized source mixture on a vector sensor array. In *2008 33rd IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2401–2404. IEEE, Las Vegas, U.S.A., March 2008.

[Guo09] Guo X., Miron S., and Brie D. Identifiabilité du mélange de sources polarisées: le cas où le nombre de capteurs est inférieur au nombre de sources. In *2009 XXIIème Colloque GRETSI Traitement du Signal & des Images (GRETSI)*. Dijon, France, September 2009.

[Guo11] Guo X., Miron S., Brie D., Zhu S., and Liao X. A CANDECOMP/PARAFAC perspective on uniqueness of DOA estimation using a vector sensor array. *IEEE Transactions on Signal Processing*, 59(7):3475–3481, 2011.

[Guo12] Guo X., Miron S., Brie D., and Stegeman A. Uni-mode and partial uniqueness conditions for CANDECOMP/PARAFAC of three-way arrays with linearly dependent loadings. *SIAM Journal on Matrix Analysis and Applications*, 33(1):111–129, 2012.

[Gutch12] Gutch H.W., Gruber P., Yeredor A., and Theis F.J. ICA over finite fields—separability and algorithms. *Signal Processing*, 92(8):1796–1808, 2012.

[Haardt08] Haardt M., Roemer F., and Del Galdo G. Higher-order SVD-based subspace estimation to improve the parameter estimation accuracy in multidimensional harmonic retrieval problems. *IEEE Trans. Signal Processing*, 56(7):3198–3213, July 2008.

[Harshman70] Harshman R.A. Foundations of the PARAFAC procedure: Models and conditions for an 'explanatory' multimodal factor analysis. *UCLA Working Papers in Phonetics*, 16:1–84, December 1970.

[Harshman72] Harshman R.A. Determinant and proof of minimum uniqueness conditions for PARAFAC1. *UCLA Working Papers Phonetics*, 22:111–117, 1972.

[He17] He Z.H., Navasca C., and Wang Q.W. Tensor decompositions and tensor equations over quaternion algebra. *arXiv preprint arXiv:1710.07552*, 2017.

[Hitchcock27a] Hitchcock F.L. The expression of a tensor or a polyadic as a sum of products. *J. Math. Phys.*, 6:164–189, 1927.

[Hitchcock27b] Hitchcock F.L. Multiple invariants and generalized rank of a p-way matrix or tensor. *J. Math. Phys.*, 7:39–79, 1927.

[Ho98] Ho K.C., Tan K.C., and Tan B.T.G. Linear dependence of steering vectors associated with tripole arrays. *IEEE Trans. Signal Processing*, 46(11):1705–1711, November 1998.

[Hoffman71] Hoffman K. and Kunze R. *Linear Algebra*. Prentice-Hall, Englewood Cliffs, NJ, 2 edition, 1971.

[Huang13] Huang J., Nie F., and Huang H. Robust discrete matrix completion. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 27. 2013.

[Huang14] Huang K., Sidiropoulos N.D., and Swami A. Non-negative Matrix Factorization revisited: uniqueness and algorithm for symmetric decomposition. *IEEE Transactions on Signal Processing*, 62(1):211–224, 2014.

[Huang16] Huang K., Sidiropoulos N.D., and Liavas A.P. A flexible and efficient algorithmic framework for constrained matrix and tensor factorization. *IEEE Transactions on Signal Processing*, 64(19):5052–5065, 2016.

[Huo16] Huo Z., Liu J., and Huang H. Optimal discrete matrix completion. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 30. 2016.

[Ioannidis20] Ioannidis V.N., Marques A.G., and Giannakis G.B. Tensor graph convolutional networks for multi-relational and robust learning. *IEEE Transactions on Signal Processing*, 68:6535–6546, 2020.

[Jiang04] Jiang T. and Sidiropoulos N.D. Kruskal's permutation lemma and the identification of CANDECOMP/PARAFAC and bilinear models with constant modulus constraints. *IEEE Trans. Signal Processing*, 52(9):2625–2636, September 2004.

[Kanatsoulis18] Kanatsoulis C.I., Fu X., Sidiropoulos N.D., and Ma W.K. Hyperspectral super-resolution: A coupled tensor factorization approach. *IEEE Transactions on Signal Processing*, 66(24):6503–6517, 2018.

[Kanatsoulis19] Kanatsoulis C.I., Fu X., Sidiropoulos N.D., and Akçakaya M. Tensor completion from regular sub-Nyquist samples. *IEEE Transactions on Signal Processing*, 68:1–16, 2019.

[Kaur18] Kaur D., Aujla G.S., Kumar N., Zomaya A.Y., Perera C., and Ranjan R. Tensor-based big data management scheme for dimensionality reduction problem in smart grid systems: SDNperspective. *IEEE Transactions on Knowledge and Data Engineering*, 30(10):1985–1998, 2018.

[Keshava02] Keshava N. and Mustard J.F. Spectral unmixing. *IEEE Signal Processing Magazine*, 19(1):44–57, 2002.

[Kiers99] Kiers H.A., Ten Berge J.M., and Bro R. PARAFAC2—part i. a direct fitting algorithm for the PARAFAC2 model. *Journal of Chemometrics: A Journal of the Chemometrics Society*, 13(3-4):275–294, 1999.

[Kim82] Kim K.H. *Boolean matrix theory and applications*, volume 70. Dekker, 1982.

[Kim14] Kim J., He Y., and Park H. Algorithms for nonnegative matrix and tensor factorizations: A unified view based on block coordinate descent framework. *Journal of Global Optimization*, 58(2):285–319, 2014.

[Kolda09] Kolda T.G. and Bader B.W. Tensor decompositions and applications. *SIAM review*, 51(3):455–500, 2009.

[Kruskal77] Kruskal J.B. Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra Applicat.*, 18(2):95–138, 1977.

[Kuang16a] Kuang L., Yang L.T., and Qiu K. Tensor-based software-defined internet of things. *IEEE Wireless Communications*, 23(5):84–89, 2016.

[Kuang16b] Kuang L., Yang L.T., Wang X., Wang P., and Zhao Y. A tensor-based big data model for QoS improvement in software defined networks. *IEEE Network*, 30(1):30–35, 2016.

[Laurberg08] Laurberg H., Christensen M.G., Plumbley M.D., Hansen K.L., and Jensen S.H. Theorems on positive data: On the uniqueness of NMF. *Computational Intelligence and Neuroscience*, 2008(764206), 2008.

[Le Bihan04a] Le Bihan N. and Mars J. Singular value decomposition of quaternion matrices: a new tool for vector-sensor signal processing. *Signal Processing*, 84(7):1177–1199, 2004.

[Le Bihan04b] Le Bihan N. and Mars J. Singular value decomposition of quaternion matrices: a new tool for vector-sensor signal processing. *Signal Processing*, 84(7):1177–1199, 2004.

[Le Bihan17] Le Bihan N. The geometry of proper quaternion random variables. *Signal Processing*, 138:106–116, 2017.

[Lebedev14] Lebedev V., Ganin Y., Rakhuba M., Oseledets I., and Lempitsky V. Speeding-up convolutional neural networks using fine-tuned CP-decomposition. *arXiv preprint arXiv:1412.6553*, 2014.

[Lee48] Lee H. Eigenvalues and canonical forms of matrices with quaternion coefficients. In *Proceedings of the Royal Irish Academy. Section A: Mathematical and Physical Sciences*, volume 52, pages 253–260. JSTOR, 1948.

[Lee99] Lee D.D. and Seung H.S. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788–791, 1999.

[Lefèvre11] Lefèvre A., Bach F., and Févotte C. On-line algorithms for Non-negative Matrix Factorization with the Itakura-Saito divergence. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 313–316. 2011.

[Leveau02] Leveau J.H. and Lindow S.E. Bioreporters in microbial ecology. *Current opinion in microbiology*, 5(3):259–265, 2002.

[Li05] Li T. A general model for clustering binary data. In *The 11th ACM SIGKDD International Conference on Knowledge Discovery in Data Mining*, pages 188–197. ACM, 2005.

[Li18a] Li B., Bai W., and Zheng G. Successive ESPRIT algorithm for joint DOA-range-polarization estimation with polarization sensitive FDA-MIMO radar. *IEEE Access*, 6:36376–36382, 2018.

[Li18b] Li X., Wang D., Wang W.Q., Liu W., and Ma X. Range-angle localization of targets with planar frequency diverse subaperturing MIMO radar. *IEEE Access*, 6:12505–12517, 2018.

[Lim09] Lim L.H. and Comon P. Nonnegative approximations of nonnegative tensors. *J. Chemometr.*, 23(7-8):432–441, 2009.

[Lin15] Lin C.H., Ma W.K., Li W.C., Chi C.Y., and Ambikapathi A. Identifiability of the simplex volume minimization criterion for blind hyperspectral unmixing: The no-pure-pixel case. *IEEE Transactions on Geoscience and Remote Sensing*, 53(10):5530–5546, 2015.

[Liu01] Liu X. and Sidiropoulos N.D. Cramér-Rao lower bounds for low-rank decompositions of multidimensional arrays. *IEEE Trans. Signal Processing*, 49(9):2074–2089, September 2001.

[Liu16] Liu Y., Ruan H., Wang L., and Nehorai A. The random frequency diverse array: A new antenna structure for uncoupled direction-range indication in active sensing. *IEEE Journal of Selected Topics in Signal Processing*, 11(2):295–308, 2016.

[Long02] Long D.A. *The Raman Effect: A Unified Treatment of the Theory of Raman Scattering by Molecules*. John Wiley & Sons. Ltd, 2002.

[Lowd05] Lowd D. and Domingos P. Naive Bayes models for probability estimation. In *Proceedings of the 22nd International Conference on Machine Learning*, pages 529–536. ACM, 2005.

[Lu12] Lu H., Vaidya J., Atluri V., and Hong Y. Constraint-aware role mining via extended boolean matrix decomposition. *IEEE Transactions on Dependable and Secure Computing*, 9(5):655–669, 2012.

[Lu14] Lu G. and Fei B. Medical hyperspectral imaging: a review. *Journal of Biomedical Optics*, 19(1):010901, 2014.

[Ma14] Ma W.K. et al. A signal processing perspective on hyperspectral unmixing: Insights from remote sensing. *IEEE Sig. Proc. Mag.*, 31(1):67–81, 2014.

[Mairal10] Mairal J., Bach F., Ponce J., and Sapiro G. On-line learning for matrix factorization and sparse coding. *Journal of Machine Learning Research*, 11(Jan):19–60, 2010.

[Mandic11] Mandic D.P., Jahanchahi C., and Took C.C. A quaternion gradient operator and its applications. *IEEE Signal Processing Letters*, 18(1):47–50, 2011.

[McCauley19] McCauley P.I., Cairns I.H., White S.M., Mondal S., Lenc E., Morgan J., and Oberoi D. The low-frequency solar corona in circular polarization. *Solar Physics*, 294(8):106, 2019.

[Meeds06] Meeds E., Ghahramani Z., Neal R.M., and Roweis S.T. Modeling dyadic data with binary latent factors. In *Advances in neural information processing systems*, pages 977–984. 2006.

[Miao20] Miao J., Kou K.I., and Liu W. Low-rank quaternion tensor completion for recovering color videos and images. *Pattern Recognition*, 107:107505, 2020.

[Miettinen08] Miettinen P., Mielikainen T., Gionis A., Das G., and Mannila H. The discrete basis problem. *IEEE Transactions on Knowledge and Data Engineering*, 20(10):1348–1362, 2008.

[Miron05] Miron S., Le Bihan N., and Mars J.I. Vector-sensor MUSIC for polarized seismic sources localization. *EURASIP Journal on Applied Signal Processing*, 2005:74–84, 2005.

[Miron06] Miron S., Le Bihan N., and Mars J.I. Quaternion-MUSIC for vector-sensor array processing. *IEEE Transactions on Signal Processing*, 54(4):1218–1229, 2006.

[Miron08] Miron S., Guo X., and Brie D. DOA estimation for polarized sources on a vector-sensor array by PARAFAC decomposition of the fourth-order

covariance tensor. In *2008 16th European Signal Processing Conference (EUSIPCO)*. Lausanne, Switzerland, August 2008.

[Miron10] Miron S., Guo X., and Brie D. Vector sensor array processing for polarized sources using a quadrilinear representation of the data covariance. In *Signal Processing*, pages 19–36. InTech, 2010.

[Miron11] Miron S., Dossot M., Carteret C., Margueron S., and Brie D. Joint processing of the parallel and crossed polarized Raman spectra and uniqueness in blind nonnegative source separation. *Chemometrics and Intelligent Laboratory Systems*, 105(1):7–18, 2011.

[Miron13] Miron S., Song Y., Brie D., and Wong K.T. CANDECOMP/PARAFAC (CP) direction finding with multi-scale array. In *2013 5th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pages 224–227. Saint Martin, France, December 2013.

[Miron15a] Miron S. and Brie D. Some rank conditions for the identifiability of the sparse Paralind model. In *2015 International Conference on Latent Variable Analysis and Signal Separation (LVA-ICA)*, pages 41–48. Springer, Liberec, Czech Republic, August 2015. ISBN 978-3-319-22481-7.

[Miron15b] Miron S., Song Y., Brie D., and Wong K.T. Multilinear direction finding for sensor-array with multiple scales of invariance. *IEEE Transactions on Aerospace and Electronic Systems*, 51(3):2057–2070, 2015.

[Miron21] Miron S., Diop M., Larue A., Robin E., and Brie D. Boolean decomposition of binary matrices using a post-nonlinear mixture approach. *Signal Processing*, 178:107809, 2021. doi: 10.1016/j.sigpro.2020.107809. URL http://dx.doi.org/10.1016/j.sigpro.2020.107809.

[Moussaoui05] Moussaoui S., Brie D., and Idier J. Non-negative source separation: Range of admissible solutions and conditions for the uniqueness of the solution. In *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, volume 5, pages 289–292. Philadelphia, PA, USA, March 2005.

[Moussaoui06] Moussaoui S., Brie D., Mohammad-Djafari A., and Carteret C. Separation of non-negative mixture of non-negative sources using a bayesian approach and MCMC sampling. *IEEE Trans. Signal Processing*, 54(11):4133–4145, November 2006.

[Mu12] Mu T. et al. Static hyperspectral imaging polarimeter for full linear Stokes parameters. *Opt. Exp.*, 20(16):18194–18201, 2012.

[Nehorai94] Nehorai A. and Paldi E. Vector-sensor array processing for electromagnetic source localization. *IEEE Trans. Signal Processing*, 42(2):376–397, February 1994.

[Nenova13] Nenova E., Ignatov D.I., and Konstantinov A.V. An FCA-based boolean matrix factorisation for collaborative filtering. *arXiv preprint arXiv:1310.4366*, 2013.

[Nguyen18] Nguyen D.M., Tsiligianni E., and Deligiannis N. Learning discrete matrix factorization models. *IEEE Signal Processing Letters*, 25(5):720–724, 2018.

[Nie19] Nie X., Yang L.T., Feng J., and Zhang S. Differentially private tensor train decomposition in edge-cloud computing for SDN-based internet of things. *IEEE Internet of Things Journal*, 7(7):5695–5705, 2019.

[Nus18] Nus L., Miron S., and Brie D. On-line blind unmixing for hyperspectral pushbroom imaging systems. In *2018 IEEE Statistical Signal Processing Workshop (SSP)*, pages 418–422. IEEE, Freiburg, Germany, June 2018.

[Nus20] Nus L., Miron S., and Brie D. An ADMM-based algorithm with minimum dispersion regularization for on-line blind unmixing of hyperspectral images. *Chemometrics and Intelligent Laboratory Systems*, 204:104090, 2020.

[Orlin77] Orlin J. Contentment in graph theory: covering graphs with cliques. In *Indagationes Mathematicae (Proceedings)*, volume 80, pages 406–424. Elsevier, 1977.

[Orozco Suárez07] Orozco Suárez D., Bellot Rubio L., del Toro Iniesta J., Tsuneta S., Lites B., Ichimoto K., Katsukawa Y., Nagata S., Shimizu T., Shine R., et al. Quiet-Sun internetwork magnetic fields from the inversion of Hinode measurements. *The Astrophysical Journal*, 670:L61–L64, 2007.

[Oseledets11] Oseledets I.V. Tensor-train decomposition. *SIAM J. Scientific Computing*, 33(5):2295–2317, 2011.

[Paatero94] Paatero P. and Tapper U. Positive matrix factorization: Anon-negative factor model with optimal utilization of error estimates of data values. *Environmetrics*, 5(2):111–126, 1994.

[Parcollet20] Parcollet T., Morchid M., and Linares G. A survey of quaternion neural networks. *Artificial Intelligence Review*, 53(4):2957–2982, 2020.

[Parrello15] Parrello D., Mustin C., Brie D., Miron S., and Billard P. Multicolor whole-cell bacterial sensing using a synchronous fluorescence spectroscopy-based approach. *PloS one*, 10(3), 2015.

[Pellat-Finet84] Pellat-Finet P. Représentation des états et des Opérateurs de Polarisation de la Lumière Par des Quaternions. *Optica Acta : International Journal of Optics*, 31:415–434, 1984. ISSN 0030-3909. doi: 10.1080/713821517.

[Prévost20] Prévost C., Usevich K., Comon P., and Brie D. Hyperspectral super-resolution with coupled Tucker approximation: Recoverability and SVD-based algorithms. *IEEE Transactions on Signal Processing*, 68:931–946, 2020.

[Qi20] Qi Y., Michałek M., and Lim L.H. Complex best r-term approximations almost always exist in finite dimensions. *Applied and Computational Harmonic Analysis*, 49(1):180–207, 2020.

[Qin16] Qin S., Zhang Y.D., Amin M.G., and Gini F. Frequency diverse coprime arrays with coprime frequency offsets for multitarget localization. *IEEE Journal of Selected Topics in Signal Processing*, 11(2):321–335, 2016.

[Rahamim04] Rahamim D., Tabrikian J., and Shavit R. Source localization using vector sensor array in a multipath environment. *IEEE Trans. Signal Processing*, 52(11):3096–3103, November 2004.

[Rasmussen03] Rasmussen C.E. Gaussian processes in machine learning. In *Summer school on machine learning*, pages 63–71. Springer, 2003.

[Razon91] Razon A. and Horwitz L.P. Tensor product of quaternion Hilbert modules. *Acta Applicandae Mathematica*, 24(2):141–178, 1991.

[Richartz49] Richartz M. and Hsü H.Y. Analysis of Elliptical Polarization. *Journal of the Optical Society of America*, 39(2):136, 1949. ISSN 0030-3941. doi: 10.1364/JOSA.39.000136.

[Rodman14a] Rodman L. *Topics in quaternion linear algebra*, volume 45. Princeton University Press, 2014.

[Rodman14b] Rodman L. *Topics in quaternion linear algebra*, volume 45. Princeton University Press, 2014.

[Roy89] Roy R. and Kailath T. Esprit-estimation of signal parameters via rotational invariance techniques. *IEEE Trans. Acoust., Speech, Signal Processing*, 37(7):984 – 995, July 1989.

[S. Ramanathan97] S. Ramanathan M.E. and Daunert S. Bacterial biosensors for monitoring toxic metals. *Trends Biotechnol.*, 15:500–506, 1997.

[Salmi09] Salmi J., Richter A., and Koivunen V. Sequential unfolding svd for tensors with applications in array signal processing. *IEEE Transactions on Signal Processing*, 57(12):4719–4733, 2009.

[Sangwine98] Sangwine S.J. Colour image edge detector based on quaternion convolution. *Electronics Letters*, 34(10):969–971, 1998.

[Schaefer07] Schaefer B., Collett E., Smyth R., Barrett D., and Fraher B. Measuring the Stokes polarization parameters. *American Journal of Physics*, 75(2):163–168, 2007.

[Schmidt81] Schmidt R.O. *A signal subspace approach to multiple emitter location and spectral estimation*. Ph.D. dissertation, Stanford University, Stanford, CA, May 1981.

[Shan85] Shan T.J., Wax M., and Kailath T. On spatial smoothing for direction-of-arrival estimation of coherent signals. *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-33(4):806–811, August 1985.

[Shen09] Shen B.H., Ji S., and Ye J. Mining discrete patterns via binary matrix factorization. In *The 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 757–766. ACM, 2009.

[Shen17] Shen Y., Baingana B., and Giannakis G.B. Tensor decompositions for identifying directed graph topologies and tracking dynamic networks. *IEEE Transactions on Signal Processing*, 65(14):3675–3687, 2017.

[Sidiropoulos00a] Sidiropoulos N.D. and Bro R. On the uniqueness of multilinear decomposition of N-way arrays. *J. Chemometr.*, 14(3):229–239, 2000.

[Sidiropoulos00b] Sidiropoulos N.D., Bro R., and Giannakis G.B. Parallel factor analysis in sensor array processing. *IEEE Trans. Signal Processing*, 48(8):2377–2388, August 2000.

[Sidiropoulos01] Sidiropoulos N. and Dimić G. Blind multiuser detection in W-CDMA systems with large delay spread. *IEEE Signal Processing Lett.*, 8:87–89, 2001.

[Silva13] Silva D.G., Nadalin E.Z., Montalvão J., and Attux R. The modified MEXICO for ICA over finite fields. *Signal processing*, 93(9):2525–2528, 2013.

[Singh19] Singh A., Aujla G.S., Garg S., Kaddoum G., and Singh G. Deep-learning-based SDN model for internet of things: An incremental tensor train approach. *IEEE Internet of Things Journal*, 7(7):6302–6311, 2019.

[Song19] Song Q., Ge H., Caverlee J., and Hu X. Tensor completion algorithms in big data analytics. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 13(1):1–48, 2019.

[Sorber13] Sorber L., Van Barel M., and De Lathauwer L. Tensorlab v1.0. http://esat.kuleuven.be/sista/tensorlab/, 2013. [Available online, February 2013].

[Stegeman09a] Stegeman A. On uniqeness conditions for CANDECOMP/PARAFAC and INDSCAL with full column rank in one mode. *Linear Algebra Appl.*, 431(1-2):211–227, July 2009.

[Stegeman09b] Stegeman A. and de Almeida A.L.F. Uniqueness conditions for constrained three-way factor decompositions with linearly dependent loadings. *SIAM J. Matrix Anal. Appl.*, 31(3):1469–1499, 2009.

[Stein12] Stein M.L. *Interpolation of spatial data: some theory for kriging*. Springer Science & Business Media, 2012.

[Swindlehurst97] Swindlehurst A., Goris M., and Ottersten B. Some experiments with array data collected in actual urban and suburban environments. In *IEEE Workshop on Signal Proc. Adv. in Wireless Comm.*, pages 301–304. Paris, France, 1997.

[Talwar96] Talwar S., Viberg M., and Paulraj A. Blind separation of synchronous co-channel digital signals using an antenna array. Part I. Algorithms. *IEEE Transactions on Signal Processing*, 44(5):1184–1197, 1996.

[Tan96a] Tan K.C., Ho K.C., and Nehorai A. Linear independence of steering vectors of an electromagnetic vector sensor. *IEEE Trans. Signal Processing*, 44(12):3099–3107, December 1996.

[Tan96b] Tan K.C., Ho K.C., and Nehorai A. Uniqueness study of measurements obtainable with arrays of electromagnetic vector sensors. *IEEE Trans. Signal Processing*, 44(4):1036–1039, April 1996.

[Tecon06] Tecon R. and Van der Meer J.R. Information from single-cell bacterial biosensors: what is it good for? *Current Opinion in Microbiology*, 17(1):4–10, February 2006.

[Tecon08] Tecon R. and van der Meer J.R. Bacterial biosensors for measuring availability of environmental pollutants. *Sensors*, 8(7):4062–4080, 2008.

[tenBerge02] ten Berge J.M.F. and Sidiropoulos N.D. On uniqueness in CANDECOMP/PARAFAC. *Psychometrika*, 67(3):399–409, September 2002.

[tenBerge04] ten Berge J.M.F. Partial uniqueness in CANDECOMP/PARAFAC. *J. Chemometr.*, 18(1):12–16, May 2004.

[Tu11] Tu S., Chen R., and Xu L. A binary matrix factorization algorithm for protein complex prediction. *Proteome Science*, 9(1):1, 2011.

[Tudor10a] Tudor T. Vectorial Pauli algebraic approach in polarization optics . II . Interaction of light with the canonical polarization devices. *Optik - International Journal for Light and Electron Optics*, 121(23):2149–2158, 2010. ISSN 0030-4026. doi: 10.1016/j.ijleo.2009.08.001.

[Tudor10b] Tudor T. Vectorial Pauli algebraic approach in polarization optics. I. Device and state operators. *Optik - International Journal for Light and Electron Optics*, 121(13):1226–1235, 2010. ISSN 0030-4026. doi: 10.1016/j.ijleo.2009.01.004.

[Turrell72] Turrell G. *Infrared and raman spectra of crystals*. Academic Press, 1972.

[VanderVeen97] Van der Veen A.J. Analytical method for blind binary signal separation. *IEEE Transactions on Signal Processing*, 45(4):1078–1082, 1997.

[Wang11] Wang F., Li P., and König A.C. Efficient document clustering via Online Non-negative Matrix Factorizations. In *Society for Industrial and Applied Mathematics (SIAM)*, volume 11, pages 908–919. 2011.

[Wang13a] Wang D. and Lu H. On-line learning parts-based representation via incremental orthogonal projective Non-negative Matrix Factorization. *Signal Processing*, 93(6):1608–1623, 2013.

[Wang13b] Wang W.Q. and Shao H. Range-angle localization of targets by a double-pulse frequency diverse array radar. *IEEE Journal of Selected Topics in Signal Processing*, 8(1):106–114, 2013.

[Wang14] Wang W.Q. Subarray-based frequency diverse array radar for target range-angle estimation. *IEEE Transactions on Aerospace and Electronic Systems*, 50(4):3057–3067, 2014.

[Wang16a] Wang W.Q. Moving-target tracking by cognitive RF stealth radar using frequency diverse array antenna. *IEEE Transactions on Geoscience and Remote Sensing*, 54(7):3764–3773, 2016.

[Wang16b] Wang W.Q. Overview of frequency diverse array in radar and navigation applications. *IET Radar, Sonar & Navigation*, 10(6):1001–1012, 2016.

[Watson16] Watson T. Nonnegative rank vs. binary rank. *Chicago Journal of Theoretical Computer Science*, 2:1–13, 2016.

[Watts01] Watts V.L. Boolean rank of Kronecker products. *Linear Algebra and its Applications*, 336(1-3):261–264, 2001.

[Wax89] Wax M. and Ziskind I. On unique localization of multiple sources by passive sensor arrays. *IEEE Trans. Acoust., Speech, Signal Processing*, 37(7):996–1000, July 1989.

[Weiss93] Weiss A.J. and Friedlander B. Direction finding for diversely polarized signals using polynomial rooting. *IEEE Trans. Signal Processing*, 41(5):1893–1905, May 1993.

[Whitney71] Whitney C. Pauli-Algebraic Operators in Polarization Optics. *Journal of the Optical Society of America*, 61(9):1207–1213, 1971. ISSN 0030-3941. doi: 10.1364/JOSA.61.001207.

[Wong98] Wong K.T. and Zoltowski M.D. Direction-finding with sparse rectangular dual-size spatial invariance arrays. *IEEE Trans. Aerosp. Electron. Syst.*, 34(4):1320–1336, October 1998.

[Wu14] Wu Y., Shen B., and Ling H. Visual tracking via On-line Non-negative Matrix Factorization. *IEEE Transactions on Circuits and Systems for Video Technology*, 24(3):374–383, 2014.

[Xu15a] Xu D., Jahanchahi C., Took C.C., and Mandic D.P. Enabling quaternion derivatives: the generalized HR calculus. *Royal Society open science*, 2(8):150255, 2015.

[Xu15b] Xu D. and Mandic D.P. The theory of quaternion matrix derivatives. *IEEE Trans. Signal Processing*, 63(6):1543–1556, 2015.

[Xu16] Xu D., Xia Y., and Mandic D.P. Optimization in quaternion dynamic systems: gradient, Hessian, and learning algorithms. *IEEE Transactions on Neural Networks and Learning Systems*, 27(2):249–261, 2016.

[Xu20] Xu T., Yang Y., Huang M., Wang H., Wu D., and Yi Q. Tensor-based angle and range estimation method in monostatic FDA-MIMO radar. *Mathematical Problems in Engineering*, 2020, 2020.

[Zhang97] Zhang F. Quaternions and matrices of quaternions. *Linear algebra and its applications*, 251:21–57, 1997.

[Zhang07] Zhang Z., Ding C., Li T., and Zhang X. Binary matrix factorization with applications. In *IEEE Seventh International Conference on Data Mining (ICDM 2007)*, pages 391–400. IEEE, 2007.

[Zhang10] Zhang Z.Y., Li T., Ding C., Ren X.W., and Zhang X.S. Binary matrix factorization for analyzing gene expression data. *Data Mining and Knowledge Discovery*, 20(1):28–52, 2010.

[Zhang13] Zhang L., Huang T.Z., Zhu Q.F., and Feng L. Uni-mode uniqueness conditions for CANDECOMP/PARAFAC decomposition of n-way arrays with linearly dependent loadings. *Linear Algebra and its Applications*, 439(7):1918–1928, 2013.

[Zhang20] Zhang G., Fu X., Wang J., Zhao X.L., and Hong M. Spectrum cartography via coupled block-term tensor decomposition. *IEEE Transactions on Signal Processing*, 68:3660–3675, 2020.

[Zhao09] Zhao Y., Zhang L., and Pan Q. Spectropolarimetric imaging for pathological analysis of skin. *Appl. Opt.*, 48(10):236–246, 2009.

[Zhao16] Zhao R. and Tan V.Y. On-line Non-negative Matrix Factorization with outliers. *IEEE Transactions on Signal Processing*, 65(3):555–570, 2016.

[Zhou11] Zhou G., Yang Z., Xie S., and Yang J.M. On-line blind source separation using Incremental Non-negative Matrix Factorization with volume constraint. *IEEE Transactions on Neural Networks*, 22(4):550–560, 2011.

[Zniyed18] Zniyed Y., Boyer R., de Almeida A.L., and Favier G. High-order CPD estimation with dimensionality reduction using a tensor train model. In *Proc. of the 26th European Signal Processing Conference (EUSICO 2018)*. Rome, Italy, September 2018.

[Zniyed19a] Zniyed Y., Boyer R., de Almeida A.L.F., and Favier G. Multidimensional harmonic retrieval based on vandermonde tensor train. *Elsevier Signal Processing*, 163:75–86, 2019.

[Zniyed19b] Zniyed Y., Miron S., Boyer R., and Brie D. Uniqueness of tensor train decomposition with linear dependencies. In *2019 8th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pages 460–464. Guadeloupe, France, December 2019.

[Zoltowski00a] Zoltowski M.D. and Wong K.T. Closed-form eigenstructure-based direction finding using arbitrary but identical subarrays on a sparse uniform cartesian array grid. *IEEE Trans. Signal Processing*, 48(8):2205–2210, August 2000.

[Zoltowski00b] Zoltowski M.D. and Wong K.T. ESPRIT-based 2-D direction finding with a sparse uniform array of electromagnetic vector sensors. *IEEE Trans. Signal Processing*, 48(8):2195–2204, August 2000.

Résumé

Ce mémoire synthétise mes activités de recherche au Centre de Recherche en Automatique de Nancy (CRAN) entre 2006 et 2021. Les projets menés s'articulent autour du traitement des signaux multidimensionnels et de l'analyse des données N-D, avec des applications en traitement d'antenne, spectroscopie, imagerie hyperspectrale, spectro-polarimétrie et apprentissage automatique. La première partie du manuscrit fournit un bilan synthétique de mes activités d'enseignement et de recherche menées respectivement, au département Réseaux et Télécommunication de l'IUT Nancy-Brabois, et au CRAN. Dans la seconde partie, je présente mes contributions méthodologiques, algorithmiques et appliquées sur quatre volets. Les travaux décrits dans le premier chapitre sont relatifs à l'étude de l'unicité des décompositions tensorielles. L'objectif est de trouver des conditions permettant de garantir l'unicité des décompositions en présence des dépendances linéaires dans les matrices des différents modes. Les deuxième et troisième chapitres sont dédiés respectivement, à l'application des modèles et algorithmes tensoriels en traitement d'antenne et en spectroscopie. Le quatrième chapitre illustre mes contributions au démélange de sources multidimensionnelles sous contraintes; les résultats évoqués dans ce chapitre concernent le démélange boolean de sources binaires, la séparation *en ligne* d'images hyperspectrales et la factorisation no-négative de matrices à valeurs quaternioniques (QNMF). Dans la quatrième partie, je dresse quelques perspectives à court et moyen terme dans le domaine de la factorisation des tableaux à valeurs catégorielles, des représentations de rang faible pour les tenseurs à valeurs quaterniniques, des approches tensorielles pour l'analyse des graphs dynamiques, du traitement d'antenne à diversité de fréquence et de la cartographie spatio-temporelle des polluants.

Mots-clés : signaux multidimensionnels, tenseurs, quaternions, unicité, traitement d'antenne, polarisation, imagerie hyperspectrale, spectroscopie, tableaux à valeurs binaires

Abstract

This thesis summarizes my research activities at the Centre de Recherche en Automatique de Nancy (CRAN) between 2006 and 2021. The projects carried out are focused on multidimensional signal processing and N-D data analysis, with applications in sensor array processing, spectroscopy, hyperspectral imaging, spectro-polarimetry and machine learning. The first part of the manuscript presents a summary of my teaching and research activities, carried out at the Networks and Telecommunications Department of the IUT Nancy-Brabois and at CRAN, respectively. In the second part, I present my methodological, algorithmic and applied contributions on four topics. The work described in the first chapter concerns to the study of tensor decompositions uniqueness. The objective is to find conditions that guarantee the uniqueness of the decompositions in the presence of linear dependencies between the columns of the factor matrices. The second and third chapters are dedicated to the application of tensor-based models and algorithms in sensor array processing and spectroscopy, respectively. The fourth chapter illustrates my contributions in the field of constrained unmixing of multidimensional sources; the results given in this chapter concern the Boolean unmixing of binary sources, the online separation of hyperspectral images and the nonnegative factorization of quaternion valued matrices (QNMF). In the fourth part, I outline some short-term and mid-term perspectives in the fields of categorical arrays factorization, low-rank representations of quaternion valued tensors, tensor approaches for dynamic graph analysis, frequency diverse array processing and spatio-temporal cartography of pollutants.

Keywords : multidimensional signals, tensors, quaternions, uniqueness, sensor array processing, polarization, hyperspectral imaging, spectroscopy, binary arrays