

Contrat doctoral Etablissement 2024/2027**Formulaire de candidature**

Titre de la thèse	Décompositions tensorielles couplées pour le démixage de signaux multidimensionnels. Application en spectroscopie de fluorescence.
Ecole Doctorale	ED548
Laboratoire	Laboratoire d'Informatique et des Systèmes (LIS)
Discipline	Discipline : Automatique, Signal, Productique, Robotique Spécialité : Traitement du signal
Directeur(s) de Thèse & Encadrant(s)	Xavier Luciani (LIS) (Directeur de Thèse), Roland Redon (MIO) (Encadrant)

Description du sujet de recherche

(3 pages maximum - contexte scientifique, objectifs, mots clé, références)

Contexte, originalité et pertinence par rapport à l'état de l'art :

Le démixage de signaux est un domaine important du traitement du signal et répond à un problème que l'on rencontre dans de nombreuses applications de la physique et des sciences de l'ingénieur, pour lequel les signaux observés sont vus comme des mélanges de signaux sources d'intérêt que l'on cherche à estimer. Dans le cas des signaux discrets qui nous intéressent ici et parmi les nombreuses approches proposées, les décompositions tensorielles fournissent un cadre théorique naturel pour le démixage de signaux multidimensionnels (c'est-à-dire variant en fonction d'au moins trois variables, on parle également de signaux multivariés) ainsi qu'un ensemble d'outils algorithmiques [1]. Ce type d'approche a notamment été appliqué avec succès en spectroscopie, pour le démixage de signaux de fluorescence [2]. L'équipe SIIM du LIS, en collaboration avec l'équipe CEM du laboratoire MIO a une expertise reconnue dans le domaine depuis plusieurs années [3,4].

Nous appelons ici décomposition tensorielle toute représentation exacte ou approchée d'un tableau de nombres à plus de deux dimensions à l'aide d'un ensemble fini de vecteurs appelés facteurs de la décomposition. Si le modèle de décomposition est adapté (problème direct), alors le calcul effectif des facteurs permet de résoudre le problème inverse d'estimation de signaux sources à partir de leur mélange. Dans le cas multidimensionnel le plus simple, les signaux observés sont tridimensionnels et issus d'un même dispositif de mesure. Ils représentent l'évolution d'un même phénomène $F(x_1, x_2, x_3)$ en fonction de 3 variables discrètes (x_1, x_2, x_3) et sont modélisés comme des mélanges trilinéaires de signaux sources bidimensionnels. La décomposition tensorielle correspondante est la Décomposition Canonique Polyadique (DCP, aussi appelée PARAFAC) d'ordre 3. Elle admet une solution unique qui permet de donner une interprétation physique aux différents facteurs et de nombreux algorithmes permettent de la calculer [5,6]. Ce cas de figure est désormais bien maîtrisé, il est représenté par l'exemple de la figure 1 dans le cas de signaux de fluorescence bivariés.

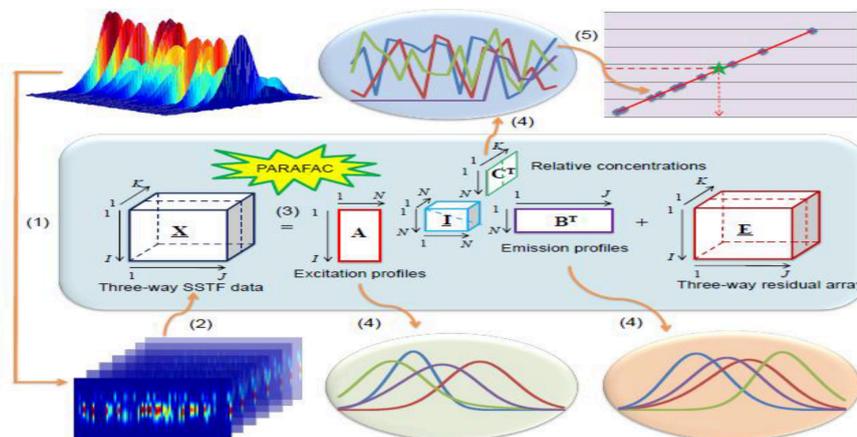


Figure 1: Décomposition CP (PARAFAC) de signaux de fluorescence 3D

Cette thèse sera donc l'occasion 1) d'envisager des scénarios de mélange plus généraux et d'en déduire les modèles de décomposition tensorielle sous-jacents 2) développer des algorithmes de calculs spécifiques à ces décompositions 3) les appliquer à des données réelles issues de la spectroscopie de fluorescence.

Objectifs et Méthodes :

Nous considérerons ainsi l'analyse conjointe d'ensembles de signaux multivariés d'ordres quelconques $F(x_1, \dots, x_p)$, $F(y_1, \dots, y_q)$, $F(z_1, \dots, z_r)$... regroupés dans différents tenseurs et pour lesquels l'une des variétés au moins est commune (justifiant ainsi le traitement conjoint). Le problème sera donc réécrit sous la forme de décompositions tensorielles couplées, permettant de prendre en compte cette information commune. Ce type d'approche s'est développé durant la dernière décennie, en particulier pour de la fusion de données. Nous pourrions ainsi nous appuyer sur un état de l'art conséquent dont un exemple est donné dans [7]. Cependant, il nous sera probablement nécessaire d'introduire de nouveaux modèles de décompositions couplées répondant aux contraintes suivantes.

Tout d'abord les valeurs prises par les variables communes pourront être différentes d'un ensemble de signaux à un autre, par exemple du fait de l'utilisation de capteurs ayant des résolutions très différentes. A cette occasion, nous nous poserons également la question de savoir jusqu'à quel point un tenseur peut être sous-échantillonné suivant certaines de ses dimensions sans dégrader l'estimation des facteurs liés aux autres dimensions. Deuxièmement, deux cas de figures seront envisagés lors du calcul de la décomposition : avec ou sans connaissance a priori sur les facteurs. Dans le premier cas nous étudierons la possibilité d'injecter pour certains facteurs des modèles paramétriques pertinents dans un cadre applicatif donné. Par ailleurs, jusqu'ici, nous avons considéré implicitement que les signaux sources présents dans le modèle de mélange sont identiques en forme pour tous les mélanges. Cette hypothèse, très classique, n'est pas toujours réalisée en pratique. Nous serons donc amenés à proposer des modèles de décomposition tenant compte de déformations possibles des signaux sources d'un mélange à l'autre sans que ceux-ci soient identifiés comme résultants de sources différentes lors de la résolution du problème inverse. Ces signaux sources sont supposés unimodaux ou bien bimodaux. Une première transformation des signaux sources consiste à leur appliquer une translation sans modification de forme. Cette approche modélisée par Harshman, dont la résolution numérique (PARAFAC2) a été proposée par Kiers [8] constituera une des méthodes de référence dans ce contexte. La deuxième transformation possible des signaux sources sera la dilatation autour du maximum de leur profil, dilatation qui sera fonction de chaque signal source et de chaque échantillon. Enfin, il est envisagé une troisième transformation qui consisterait à déformer chaque signal source pour un échantillon donné en lui appliquant un profil multiplicatif proche de un, profil qu'il serait donc nécessaire d'optimiser en plus des amplitudes de translation et des facteurs de dilatation liés aux deux transformations précédentes. Enfin, au-delà des modèles pré-existants ou non, nous proposerons des solutions algorithmiques originales pour le calcul effectif des décompositions. Nous privilégierons ainsi les approches de résolution semi-algébriques, basées sur la diagonalisation conjointe de matrices [6,9].

L'application considérée est la spectroscopie de fluorescence. Cette technique est employée depuis plusieurs années au laboratoire MIO pour caractériser la matière organique dissoute dans les eaux de mer et de rivière. Celle-ci consiste à mesurer l'intensité de fluorescence d'un échantillon liquide en fonction des longueurs d'onde d'excitation et d'émission. Dans le cas particulier des molécules fluorescentes en milieu liquide, le spectre d'émission de fluorescence est décalé vers des longueurs d'onde plus élevées que les longueurs d'onde du spectre d'excitation et la forme du spectre d'émission ne dépend pas de la longueur d'onde en excitation. De plus, si l'absorbance de l'échantillon est suffisamment faible, le modèle physique de l'intensité de fluorescence devient trilineaire : il dépend alors uniquement des spectres d'excitation, des spectres d'émission et des concentrations liées aux molécules fluorescentes (fluorophores) ainsi que d'autres paramètres physiques et instrumentaux. Le laboratoire MIO dispose d'un spectrofluorimètre Fluoromax (Horiba) permettant de mesurer l'intensité de fluorescence en mode stationnaire (source d'excitation continue) ou bien résolue en temps (source d'excitation impulsionnelle). A l'aide de ce système de mesure, il sera donc possible d'obtenir des intensités de fluorescence F d'un échantillon k en fonction de plusieurs variables : les longueurs d'onde d'excitation et d'émission λ_{ex} et λ_{em} et le temps t de décroissance de la fluorescence. A ces premières variables liées au système de mesure, on pourra aussi ajouter les variables physico-chimiques suivantes : la concentration en métal ajouté à l'échantillon (notée C_M) qui provoque une atténuation de l'intensité de fluorescence (appelé quenching), et le pH (potentiel hydrogène) qui peut engendrer un décalage en longueurs d'onde des spectres d'excitation et d'émission. Les différentes combinaisons de ces variables permettront de constituer des tenseurs d'ordres 3, 4 ou 5 ayant des variétés communes comme par exemple :

Tenseurs d'ordre 3	Tenseurs d'ordre 4	Tenseur d'ordre 5
$F(\lambda_{ex}, \lambda_{em}, k),$ $F(\lambda_{ex}, \lambda_{em}, t),$ $F(\lambda_{ex}, \lambda_{em}, C_M)$	$F(\lambda_{ex}, \lambda_{em}, t, k), F(\lambda_{ex}, \lambda_{em}, t, C_M),$ $F(\lambda_{ex}, \lambda_{em}, C_M, k), F(\lambda_{ex}, \lambda_{em}, t, pH),$ $F(\lambda_{ex}, \lambda_{em}, pH, k)$	$F(\lambda_{ex}, \lambda_{em}, C_M, pH, k)$

Ces tenseurs nous serviront ainsi à valider les différents modèles et algorithmes de décomposition proposés dans la partie méthodologique de la thèse sur des signaux réels.

Retombées attendues :

2 articles méthodologiques en traitement du signal, 1 article plus applicatif en spectroscopie ou sciences de l'environnement.

Mots clés : *Démélange de signaux, Décompositions tensorielles, Algèbre multilinéaire, Optimisation numérique, Modèles physiques, Chimiométrie, Spectroscopie de fluorescence*

Références :

- [1] P. Comon, "Tensors: a brief introduction", IEEE Sig. Proc. Magazine, 31(3):44--53, May 2014. special issue on BSS.
- [2] R. Bro, PARAFAC. Tutorial and applications, Chemometrics and Intelligent Laboratory Systems, Volume 38, Issue 2, 1997, Pages 149-171,
- [3] X. Luciani, R. Redon, S. Mounier et A. Bois, "A simple correction method of inner filter effects affecting FEEM and its application to the PARAFAC decomposition" Chemometrics and Intelligent Laboratory Systems 96.2 (2009)
- [4] I. W. Sanou, R. Redon, X. Luciani, S. Mounier, "Online Nonnegative and Sparse Canonical Polyadic Decomposition of Fluorescence Tensors", Chemometrics and Intelligent Laboratory Systems 225 (2022)
- [5] P. Comon, X. Luciani et A. L. F. De Almeida. "Tensor Decompositions, Alternating, Least Squares and other Tales", Journal of Chemometrics 23, 2009.
- [6] X. Luciani et L. Albera. "Canonical polyadic decomposition based on joint eigenvalue decomposition", Chemometrics and Intelligent Laboratory Systems 132, 2014,
- [7] M. Sørensen, I. Domanov, L. De Lathauwer, Coupled canonical polyadic decompositions and (coupled) decompositions in multilinear rank- $(l_r, n, l_r, n, 1)$ terms – part II: Algorithms, SIAM J. Matrix Anal. Appl., vol. 36, no. 3, 2015
- [8] H. Kiers, J.. ten Berge, R. Bro, PARAFAC2 Part I. A direct fitting algorithm for the PARAFAC2 model, Special Issue:Proceedings of the International Chemometrics Research Meeting, 1998
- [9] R. André, X. Luciani et E. Moreau. "Joint EigenValue Decomposition Algorithms Based on First-Order Taylor Expansion". In : IEEE Transactions on Signal Processing, 68 (2020)

Encadrement et conditions matérielles pour le doctorant

La direction de la thèse sera assurée par Xavier Luciani (LIS) et le co-encadrement par Roland Redon (MIO).

La thèse se déroulera au sein de l'équipe Signal Image (SIIM) du LIS sur le site de Toulon où le doctorant disposera d'un bureau en open-space, d'un ordinateur et de l'accès à toutes les ressources que le laboratoire et l'équipe mettent à disposition de leurs doctorants (accès aux salles de vie et de réunion, papeterie, reprographie, catalogue de revues en ligne, accès à différents clusters de calcul, parking, etc.) Le Laboratoire MIO assurera l'acquisition des signaux de fluorescence nécessaire à la validation des modèles et algorithmes développés par le doctorant. Celui-ci aura accès aux équipements de mesure et d'instrumentation.

Compétences attendues et personnes à contacter

Compétences attendues :

Le candidat devra avoir de solides connaissances et une appétence pour une majorité des domaines suivants: Traitement du signal, méthodes numériques, programmation (Matlab de préférence), méthodes d'optimisation, séparation de sources, algèbre linéaire. Des connaissances en optique et/ou spectroscopie seraient un plus.

Personne(s) à contacter :

Xavier Luciani (luciani@univ-tln.fr), Roland Redon (roland.redon@mio.osupytheas.fr)