

Recalage de spectres RMN

Contexte

La spectroscopie par résonance magnétique nucléaire (RMN) est une technique d'acquisition de signal qui exploite les propriétés magnétiques des noyaux atomiques. On s'en sert par exemple pour analyser des échantillons bio-organiques. Les échantillons des mélanges complexes de molécules et l'on souhaite savoir par exemple quelles molécules sont présentes et en quelle proportion. Les signaux mesurés, que l'on appelle des spectres RMN, sont des superpositions des spectres de chaque molécule que l'on appelle les spectres purs. Les spectres mesurés sont des mélanges complexes de composés purs qu'il faut démêler. A ce problème de démêlage (appelé aussi séparation de sources en traitement de signal) s'ajoute un problème d'alignement des spectres [3]. En effet, un changement l'environnement chimique d'une molécule induit un déplacement du spectre de cette molécule. Ces déplacements doivent être corrigés avant de procéder au démêlage. Cette procédure est actuellement faite manuellement par la personne en charge de l'acquisition des données. Le but de ce stage est de proposer une procédure automatique de démêlage qui intègre un réaligement des spectres RMN.

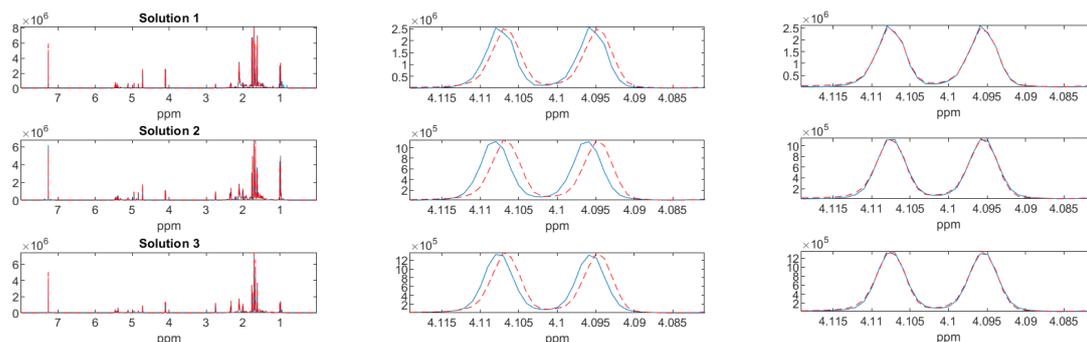


FIGURE 1 – Spectres de mélanges (gauche), zoom sur une partie du spectre avant recalage (centre), zoom sur une partie du spectre après manuel recalage (droite).

Objectifs

L'objectif de ce stage est de développer un algorithme de recalage de spectres RMN automatique. On se basera pour cela sur une formulation du problème sous une forme d'un problème d'optimisation de type moindre carrées non-linéaires séparables, qui se résout par exemple en utilisant l'algorithme *Variable Projection* [4].

Le stage comportera des développements théoriques en optimisation et des phases d'implémentation en Python, qui seront assorties de tests qui permettront petit à petit d'affiner le problème d'optimisation et l'algorithme associé. Après la phase de prise en main, qui inclura l'application brute de l'algorithme *Variable Projection* au problème de recalage, le premier raffinement consistera à introduire un terme de régularisation quadratique, de manière à stabiliser les solutions (technique classique en problèmes inverses). L'algorithme *Variable Projection* sera facilement modifiable pour traiter ce cas mais il faudra alors étudier l'influence de l'hyperparamètre attaché à la régularisation de manière empirique. Dans un second temps, on cherchera comment écrire l'algorithme sous forme déroulée [5], cette technique d'intelligence artificielle permettant alors d'apprendre l'hyperparamètre de régularisation, au prix d'une phase d'entraînement sur une base de signaux à construire. Les différentes méthodes seront comparées sur les données simulées. On procédera à des tests sur données réelles issues de [3].

Le stage se déroulera donc de la façon suivante

Phase 1 : prise en main sujet Prise en main de la problématique du recalage des signaux RMN, incluant une phase de simulation des signaux. Écriture du problème de recalage en un problème d'optimisation. Étude de l'algorithme *Variable Projection* [4].

Phase 2 : tests de référence Implémentation de l'algorithme original *Variable Projection* sur données simulées en Python.

Phase 3 : ajout d'une régularisation Ajout d'une régularisation quadratique, étude du problème d'optimisation associé et modification de l'algorithme.

Phase 4 : tests de la méthode régularisée Implémentation de l'algorithme régularisé et comparaison avec les résultats de référence.

Phase 5 : déroulement Développement de l'algorithme déroulé associé à l'algorithme régularisé.

Phase 6 : tests de la méthode déroulée Implémentation de l'algorithme déroulé avec PyTorch, création d'une base d'apprentissage, apprentissage du réseau et comparaison avec les résultats de référence et régularisés.

Phase 7 : application Application aux données réelles.

Pré-requis

Nous recherchons un.e stagiaire de niveau M2 en mathématiques appliquées ou data science avec une appétence pour ce sujet à la croisée du traitement du signal, de l'optimisation et de l'intelligence artificielle. Des bases solides en optimisation et en programmation Python sont nécessaires. Une connaissance des signaux RMN, des techniques d'apprentissage artificiel, et une maîtrise de PyTorch seraient un plus.

Informations pratiques

Durée du stage

Stage de 4 à 5 mois, avec un début dès avril 2023.

Encadrement

L'équipe encadrante sur place sera [Sandrine Anthoine](#), chargée de recherche CNRS et [Frédéric Richard](#), professeur AMU, et pourra bénéficier de l'expertise de [Caroline Chaux](#), directrice de recherche CNRS (IPAL, Singapour).

Le/la stagiaire sera accueilli.e. au sein de l'équipe [Signal-Image](#) de l'institut de Mathématique de Marseille ([I2M](#)), qui est un laboratoire commun entre Aix-Marseille Université et le CNRS. Les locaux se situent sur le technopôle de Château-Gombert (accessible en transport en commun).

Candidature

Envoyer CV et lettre de motivation à sandrine.anthoine@univ-amu.fr et frederic.richard@univ-amu.fr dès que possible.

Références

- [1] Claridge, T.D.W., High-Resolution NMR Techniques in Organic Chemistry, Third Edition, Elsevier Science, 2016.
- [2] Source Separation and Applications, IEEE Signal Processing Magazine, Vol. 31, No. 3, May 2014.
- [3] Cherni, A. et. al., Challenges in the decomposition of 2D NMR spectra of mixtures of small molecules, Royal Society of Chemistry, 2019.
- [4] Golub, G. and Pereyra, V., Separable nonlinear least squares : the variable projection method and its applications, Inverse Problems, Vol. 19, No. 2, 2003.
- [5] V. Monga and Y. Li and Y. C. Eldar, Algorithm Unrolling : Interpretable, Efficient Deep Learning for Signal and Image Processing, IEEE Signal Processing Magazine, vol. 38, no. 2, March 2021.