

Proposition de Sujet de Thèse pour Contrat Doctoral UCA

Adresse e-mail à utiliser pour toute correspondance :

klounici@math.unice.fr

Titre de la thèse

Tri de potentiels d'action sur des données neurophysiologiques massives

Thesis Title

Spike sorting for massive neurophysiological data sets

Directeur de Thèse (HDR ou assimilé)

Nom : Lounici

Prénom : Karim

Téléphone : +33(0)4 92 07 69 96

Courriel : klounici@math.unice.fr

Laboratoire d'accueil

LJAD

Co-directeur

Nom : Flamary

Prénom : Remi

HDR : Non

Unité de recherche : Lagrange

Téléphone :

Courriel : remi.flamary@unice.fr

Domaine Scientifique

DS1 - Mathématiques et leurs Interactions

Description du sujet

La communauté mathématique poursuit, à l'heure actuelle, de nombreuses recherches à l'interface avec les Neurosciences, dans le but de comprendre les mécanismes cognitifs et comportements. Toute avancée dans ce domaine peut potentiellement ouvrir la voie vers de nombreuses applications dans le domaine biomédical. Ce projet de thèse porte spécifiquement sur un problème de tri des potentiels d'action (spike sorting) en neurophysiologie sur des données massives.

L'analyse des réseaux de neurones au niveau du système nerveux central a récemment franchi certaines frontières grâce à l'utilisation de capteurs multi-électrodes (MEA Biocam de 4096 électrodes) capables d'enregistrer les signaux bio-électriques générés par le système nerveux. Ce type de capteur est capable de mesurer l'activité de milliers de neurones simultanément. Une fois que l'activité neuronale de chacun des neurones est extraite des enregistrements des électrodes (spike sorting), on est capable de capturer les structures de corrélation au sein du réseau neuronal, et donc de reconstruire la connectivité fonctionnelle à l'intérieur du réseau. Cette reconstruction permettrait à terme de mieux comprendre les processus neurophysiologiques mis en jeu dans le cadre des mécanismes cognitifs et comportements.

Néanmoins, la complexité des signaux mesurés rend leur traitement particulièrement difficile. En effet, les milliers d'électrodes contenues sur le capteur de petite taille mesurent l'activité de plusieurs milliers de neurones simultanément avec des effets d'atténuation dus à la distance relative des neurones par rapport aux électrodes du capteur. Le traitement des données enregistrées impose donc une étape de séparation de sources, appelée tri des potentiels d'action (spike sorting). Cela consiste en la détection et l'attribution de chaque potentiel d'action ou spike enregistré au neurone qui l'a généré. Avec les capteurs classiques qui contiennent au plus quelques électrodes, cette étape de spike sorting peut être réalisée manuellement (avec éventuellement l'utilisation de méthodes semi-automatiques de pré-traitement telles que le clustering). Malheureusement, pour les capteurs de plusieurs milliers d'électrodes, cette étape de spike sorting se complexifie considérablement si bien que les méthodes existantes ne sont plus applicables.

L'objectif principal de ce doctorat est de développer une nouvelle méthodologie automatique de spike sorting implémentable sur des données massives. Ce projet comporte une composante statistique et computationnelle. Du point de vue statistique, il s'agira de mettre en évidence les paramètres intrinsèques du problème qui déterminent la complexité statistique et les vitesses de résolution de ce problème. Les résultats prendront la forme d'inégalités d'oracle parcimonieuses et leur optimalité sera aussi étudiée dans le cadre minimax. Du point de vue computationnel, ce projet développera des algorithmes de spike sorting efficaces en grande dimension basés principalement sur la théorie de l'optimisation. Ces algorithmes seront testés sur des données réelles en collaboration avec l'Institut de Biologie de Valrose autour d'une problématique de traitement de la douleur. Un objectif pratique de ce projet sera de fournir un code facilement implémentable sur des données biomédicales.

Ce doctorat fera l'objet d'un co-encadrement entre le Laboratoire J.A. Dieudonné et le laboratoire Lagrange. Il sera attendu du doctorant une bonne connaissance en probabilité

et statistiques et un intérêt marqué pour les interactions des mathématiques. Notamment, le doctorant sera amené à collaborer fréquemment avec des biologistes et des neurologues et à manipuler des données biomédicales durant ce projet.

Bibliographie :

- [1] V. Koltchinskii, K. Lounici, A.B. Tsybakov. Nuclear-norm penalization and optimal rates for noisy low-rank matrix completion. *Annals of Statistics*, 39(5):2302-2329, (2011).
- [2] Combettes, P. L., Pesquet, J. C. Proximal splitting methods in signal processing. *Fixed-point algorithms for inverse problems in science and engineering*, 28 (1),785-272, (2011)
- [3] A. Boisbunon, R. Flamary, A. Rakotomamonjy. Active set strategy for high-dimensional non-convex sparse optimization problems, *International Conference on Acoustic, Speech and Signal Processing (ICASSP)*, (2014).
- [4] Albert, M., Bouret, Y., Fromont, M., Reynaud-Bouret, P. Surrogate data method, s based on a shuffling of the trials for synchrony detection: the centering issue. *Neural Computation*, 28 (1), 2352-2352, (2016).

Description of the thesis

The mathematical community is actively working on research projects at the interface with neuroscience, with the goal to understand brain cognitive and behavioral mechanisms. Any breakthrough in this domain can potentially unlock the path to numerous biomedical applications. More specifically, this research is about a particular spike sorting problem for brain neurophysiological recordings.

The analysis of neural in the central nervous system has recently crossed a threshold with the development of new multi electrodes sensors (MEA Biocam of 4096 electrodes) This type of sensors can record the activity of thousands of neurons simultaneously within the central nervous system. If we can extract the activity of each individual neurons from these recordings (the spike sorting step), then we should eventually be able to retrieve the correlation structures within the neural network and thus understand the functional connectivities inside this network. This reconstruction can be exploited to understand the neurophysiological processes at stake in cognitive and behavioral mechanisms.

However, this is a particularly difficult task given the complexity and the size of the measured signals. Indeed, the thousand of electrodes contained on a small size sensor are measuring the activity of thousands of neurons simultaneously with potential attenuation effects due to the relative distances of the neurons to the electrodes. The treatment of the data require a source separation step called spike sorting that consists in correctly matching each observed spike to the neuron that generated it. With classical sensors that contain at most a few electrodes, this step can be carried out manually (with maybe the assistance of semi-automatic pre-processing methods such as clustering). Unfortunately, these existing methods cannot handle the new generation of sensors containing thousands of electrodes.

The main objective of this doctorate project is to develop a new automatic spike sorting method that can handle massive data sets such as those generated by MEA Biocam sensors. This project contains both statistical and computational aspects. In statistics, the objective will be to determine the intrinsic parameters of the problem that determine its statistical complexity. The new results may take the form of oracle inequalities. Their optimality will be investigated as well using the minimax framework. The computational aspect will consist in developing new optimisation algorithms to perform efficient spike sorting on high-dimensional data. These algorithms will be tested on neurophysiological data in collaboration with the institute of Biology of Valrose. A practical objective of this project will be to develop a user-friendly code readily implementable on biomedical data.

This doctorate project will be a joint-supervision between the J.A. Dieudonne and Lagrange Laboratories at Valrose. It will be expected for prospective candidates a solid knowledge in probability and statistic and a certain interest for interactions of mathematics. In particular, the Ph.D candidate will frequently cooperate with biologists and neuroscientists and manipulate biomedical data during the duration of this project.

Bibliography :

- [1] V. Koltchinskii, K. Lounici, A.B. Tsybakov. Nuclear-norm penalization and optimal rates for noisy low-rank matrix completion. *Annals of Statistics*, 39(5):2302-2329, (2011).
- [2] Combettes, P. L., Pesquet, J. C. Proximal splitting methods in signal processing. *Fixed-point algorithms for inverse problems in science and engineering*, 28 (1),785-272, (2011)
- [3] A. Boisbunon, R. Flamary, A. Rakotomamonjy. Active set strategy for high-dimensional non-convex sparse optimization problems, *International Conference on Acoustic, Speech and Signal Processing (ICASSP)*, (2014).
- [4] Albert, M., Bouret, Y., Fromont, M., Reynaud-Bouret, P. Surrogate data methods based on a shuffling of the trials for synchrony detection: the centering issue. *Neural Computation*, 28 (1), 2352-2352, (2016).

Informations complémentaires