

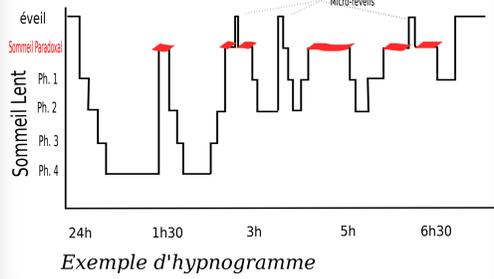
Jonas Abernot<sup>1,2</sup>, Fabien Perrin<sup>2</sup>, Florent Gobert<sup>2,3</sup>, Alexandra Corneillie<sup>2</sup>,

<sup>1</sup>Insa de Lyon, <sup>2</sup>Equipe CAP (Cognition Auditive et Psychoacoustique) CRNL (Centre de Recherche en Neurosciences de Lyon), <sup>3</sup>Hospices civils de Lyon

En 2013, Florent Gobert, déjà docteur en médecine, encadré par Fabien Perrin, entame une thèse sur la **récupération des cycles veille-sommeil** chez les personnes dans le **coma**. Son outil : L'**électroencéphalographie**, couplée à quelques autres capteurs physiologiques et environnementaux. À travers ce dispositif, il espère déceler des traces physiologiques de « **cycles** », de « **fluctuations** », qui apparaîtraient dans une période de **24 heures**, signe d'une évolution positive pour le patient.

S'appuyant sur une tradition et une expérience de la lecture visuelle du signal, il tente de construire des **hypnogrammes**, c'est à dire d'associer à chaque fenêtre de 30s un état de sommeil, résumant ainsi des milliers de valeurs en une seule.

Appliquée à des personnes dans le coma depuis peu, cette approche s'avère très difficile : les signaux sont trop différents de ceux observés chez des personnes saines. Fabien et lui décident alors d'aborder leurs données sous un **nouvel angle de vue**: celui de l'analyse automatique, mathématique, statistique, de cette abondance de signaux. Rejoins par Alexandra Corneillie et Jonas Abernot, plus versés dans les questions de traitement du signal, d'informatique et de statistiques que dans les questions de biologie humaine, d'anatomie cérébrale et d'étude de la conscience, et prêts à en découdre avec un peu de **data-mining**.



Ils travaillent donc à construire une **compréhension commune** d'un certain nombre d'**opérateurs mathématiques** destinés à décrire le signal : choisir, interpréter, comprendre ce qu'ils captent de différent ou de semblable, et lesquels vont permettre d'identifier la présence de cycles et de fluctuations.

## Réduction de dimensionnalité : Capturer les fluctuations...

La plupart des opérateurs ainsi identifiés sont liés à des mesures de **complexité**, dans le sens mathématique du terme. Cette notion renvoie à celle d'organisation, de régularité à opposer à des notions d'aléatoire, d'imprévisible.

### La géométrie fractale:

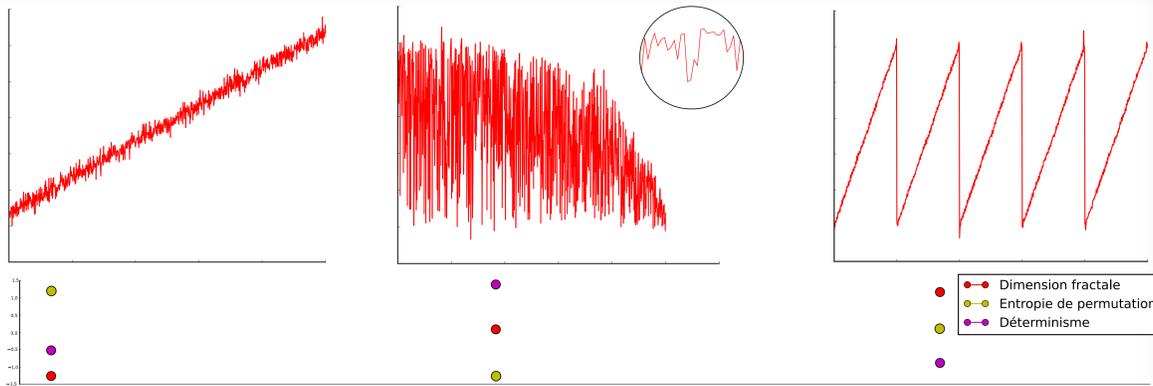
Cette géométrie s'intéresse à la description des objets rugueux. Par association d'idée, en s'interrogeant sur la 'rugosité' du signal, on s'intéresse à sa régularité.

### Les entropies:

En théorie de l'information, l'entropie est une mesure de la quantité d'information comprise dans un message. Les nombreuses variantes de cette mesure définissent chacune de manière différente la quantité d'information élémentaire.

### La dynamique du chaos:

L'étude mathématique des systèmes dynamiques a abouti à diverses mesures du niveau de déterminisme ou de stochasticité d'un signal.



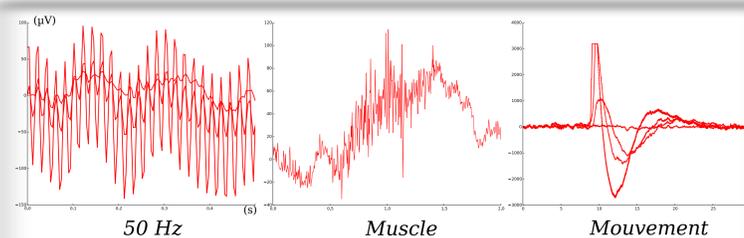
3 Exemples de descripteurs, mesurés sur 3 signaux ayant la même distribution, mais une organisation différente.

## Questions de méthodologie

Au-delà des **conditions d'acquisition** particulièrement difficiles : acquisition en milieu hospitalier (mauvaises impédances, soins réguliers...), sujets dans le coma (parfois lié à une lésion du cerveau), etc, des questions de méthodologies mathématiques se posent.

### Artefacts, outliers, surprises :

Les signaux EEG sont sujets à de nombreux artefacts physiologiques (signaux produits par les muscles) et électriques (mouvements des fils, interférence avec le milieu). Comment distinguer ces points érronés des points d'intérêt et que faire avec ? Quelle influence ont-ils sur nos descripteurs ?

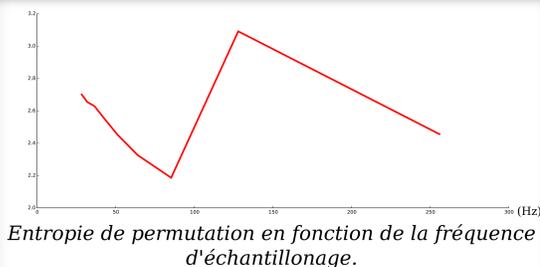


### Échantillonnage :

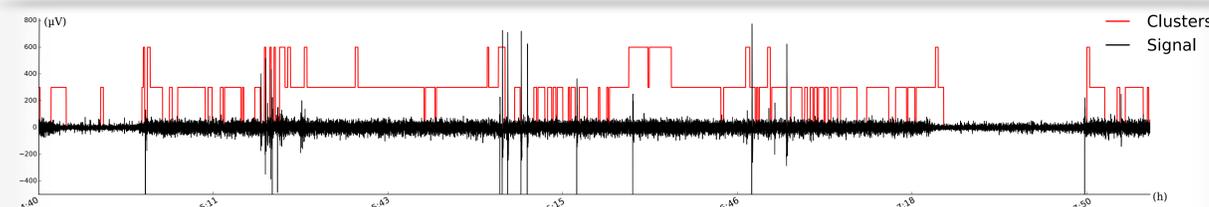
La fréquence d'échantillonnage des signaux peut avoir une forte influence sur les descripteurs. Comment identifier ces cas particuliers et les éviter ?

### Échelle :

Les descripteurs dépendent aussi souvent de la taille des signaux décrits. Quelle taille de fenêtre est pertinente ? Est-ce la même pour toutes les mesures ? Les mesures saturent-elles à partir d'une certaine taille ?



## Travaux en cours



▲ Ci-dessus : A la manière d'un hypnogramme, cette figure propose une segmentation du signal brut en différents « états ». Elle a été obtenue par un **clustering hiérarchique de fenêtres de 30 secondes**, à partir d'un ensemble de descripteurs de ces fenêtres (moyenne, variance, skewness, svd entropy, permutation entropy, detrended fluctuating analysis...), et aucune information sur la succession temporelle des fenêtres. A la différence d'un hypnogramme, cette construction algorithmique ne présuppose pas la nature des différents états, mais les construit par association de fenêtres semblables.

► Ci-contre : Nos mesures sont elles capables de distinguer les patients eux-mêmes ? Pour essayer de répondre à cette question, les mesures proposées sont réalisées sur les signaux complets. Une analyse factorielle permet d'identifier l'espace permettant le mieux de distinguer les patients.

## Python

### Étapes de l'analyse

#### Extraire

les données du formats spécifiques au matériel d'acquisition EEG.

#### Organiser

les données, les résultats intermédiaires, pour plusieurs utilisations ultérieures...

#### Stocker

ces mêmes données et résultats intermédiaires...

#### Visualiser

les données, les comparer, à plusieurs échelles, pour comprendre un résultat...

#### Transformer/Mesurer

pour relever une caractéristique du signal, une puissance, une entropie, une dimension particulière...

#### Faciliter les calculs

Paralléliser, calcul fenêtre par fenêtre...

#### Laisser une trace

de ses tentatives, de ses explorations de bibliothèques, pour pouvoir les partager, les retrouver...

#### Classifier/Segmenter

les signaux. Le coeur de l'analyse automatique.

#### Évaluer

les résultats des classifications

### Solutions Python

#### Numpy

Point de départ de l'ensemble des bibliothèques ci-dessous. Fournit une représentation en mémoire efficace des tableaux numériques homogènes.

#### Neo

Modèle objet pour données d'électrophysiologie.

#### Pandas

Tableaux annotés, outils expressifs pour les séries temporelles.

#### HDF5

À travers pandas et pytables. Stockage de données homogènes sur disque, accès partiel, enregistrement itératif...

#### Vispy

Visualisation dynamique et rapide de données lourdes, exploration.

#### Matplotlib, Seaborn

Visualisation statique, 'figures' démonstratives.

#### Scipy

Large bibliothèque comprenant notamment des fonctions de traitement du signal, de statistiques, et d'autres.

#### Pyeeeg

Petite bibliothèque proposant quelques fonctions d'analyse du signal électro-encéphalographique.

#### Des p'tits bouts maison

Si vous avez des idées...

#### Ipython notebook

Interpréteur python muni d'une interface web permettant de mêler code, explications et figures. Export en Html ou autres formats.

#### Scikits-learn

Algorithmes de classification et outils facilitant leur usage (normalisation, description...)

#### Reste à trouver...

