NX-421: Neural signals and signal processing

# MINI PROJECT 2

For this mini-project, we propose 3 variants and you will do only the one your group was assigned to. Each of the 3 variants has a practical part that corresponds to your methods and a theoretical part that you should address and integrate in your report and presentations.

**Groups AB, D, F, J, L, N, P will present on Thursday 09/11** during the exercise session. Presentations should last **10 min + 5 min of QAs**.

All the other groups should provide us with **a report of 3 pages max** (including images), single column, font-size 11px in Arial. Make sure to shape the structure in a paper-like format: Introduction, Methods, Results and Discussion sections + **the code** you used for solving the mini-project. Both should be uploaded on Moodle.

Report+Code are due on **Thursday 9th November at 16:00**.
**Note that this is a sharp deadline, no extensions are allowed.**

Everyone is encouraged to come to the presentations.

EPFL | Neuro X Institute

NX-421: Neural signals and signal processing

# Variant 1

## Groups: AB, F, AA, E, C, H, K

***TAs of reference:***
***Fabrice Guibert (fabrice.guibert@epfl.ch) & Zhiwei Huang (zhiwei.huang@epfl.ch)***

***Description***
Spatial maps derived from group fMRI data using multivariate decomposition methods like Independent Component Analysis (ICA) reveals the additive subcomponents in a multivariate signal like BOLD. In this project you will work on extracting these spatial maps using ICA and explore independent sources present in the BOLD signal.

**Dataset**:
Rather than importing data through nilearn, we provide you with an already preprocessed dataset. We conducted for you all steps of anatomical preprocessing as well as all steps of functional preprocessing. For the sake of archive size, we do not include intermediate quantities nor initial volumes. To include this dataset, you can download it here.

In this task, participants alternated between rest sessions and watching a checkerboard on screen (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8520012/ Fig 1 for the full experimental design). In the interest of time, we provide you only with the *full checkerboard* sessions.

***Practical part***
1. Start by visualizing your dataset. Are the subjects aligned to one another functionally? *Note: You do not need to provide us with a visualization, this is just for you!*
2. Apply group ICA on the dataset by concatenating sessions together (make sure to standardize each session before concatenation!).
3. Show each ICA component separately overlaid on the anatomy, in axial view only. *Note-You should assign to each ICA a number or letter to help identify it.*
4. Plot the time course of each ICA component, follow the same naming convention as you followed in the last part.
   Based on your observation, comment on some of the components? *Think about what happens to components that are appearing very rarely?*
5. Use Principal Component Analysis (PCA) to reduce the **spatial** dimensionality of your dataset for three explained variance thresholds: 25%, 50% and 95%. For each variance threshold, recompute the ICA on the lower-dimension data. For the 95% variance PCA, plot the resulting ICA components by reprojecting them to the original dimension.

6. [Bonus] What would be a relevant number of ICA components for this dataset to delineate between meaningful additive components? (Hint: it is sufficient to identify the relevant ICA components without rerunning the ICA)

***Theoretical part***

1. Given the provided files, would you be able to run a GLM? If not, what file type might you be missing to do so, and why is it not included?
2. You found several ICA components. Do some of them match relevant brain networks? Name two of them.
3. Can you see any difference between applying PCA and not applying PCA in the quality of the recovered ICA components? (Answer this for explained variance: 25%, 50% and 95%).
4. Can you explain why some components might have different contributions over time in ICA? Based on this, the spatial components and your inspection of the fMRI data, do you think ICA can be used to detect some abnormalities in the data?

NX-421: Neural signals and signal processing

# Variant 2

## Groups: D, L, P, AC, B, G, W

***TAs of reference:***
***Ekansh Sareen (ekansh.sareen@epfl.ch) & Michael Chan (chunheimichael.chan@epfl.ch)***

***Description***

In this mini-project variant, you will play with the Principal Component Analysis and its potential use as a noise reduction tool. It has already been shown that using PCA to reduce the dimensionality of data entails information loss. Simply performing PCA presents the risk that we will lose the true signal especially since the variance explained by the signal is very small relative to the noise in fMRI data.

To investigate PCA as a clean up tool, we will make use of the Pixar dataset -
Dataset description: https://openneuro.org/datasets/ds000228/versions/1.0.0.
Get the dataset using nilearn module (check notes for more details):
https://nilearn.github.io/dev/modules/generated/nilearn.datasets.fetch_development_fmri.html

Afterwards, you will also analyse the data using Partial Least Squares pipeline provided in Week 6 exercise.

***Practical part***

**[Loading and Visualizing]**

1. Load fMRI data for 25 subjects and visualize the fMRI volume.
2. Convert NIFTI images into matrices, using a consistent TR and slice for all subjects.(*Hint: Middle frames give better spatial coverage of the brain*)
3. Introduce additive salt noise to the voxels of the chosen slice and TR.

**[Denoising with SVD]**

4. Generate covariance matrix across subjects (Hint: use np.cov())
5. Apply SVD and plot singular values.
6. Pick the top 2 components from SVD and project the matrix onto them.
7. Restore denoised matrix, visualize it, and compare with the original denoised matrix. Note your observations in terms of denoising.

**[Behaviour Analysis with PLSC]**

8. Load the original data matrix from (1) and load the behavioral labels from participants.tsv (check nilearn resource: 'phenotypic' attribute), specifically the 3rd column (Child_Adult).
9. Create distinct columns for child and adult labes.
10. Perform temporal averaging preserving the original spatial dimensions.
11. Apply the provided visual mask on every subject (link).
12. Compute behavior-spatial covariance matrix.
13. Implement the PLS pipeline from Week 6 exercise.
14. Visualize the spatial maps for both behavioral scores.

***Theoretical part***

15. **What are the advantages and limitations of the PCA and PLS method on the proposed dataset?**
16. **What is the effect of changing the number of components in PCA?**
17. **How do you interpret the differences in behavioural components after applying PLS?**
18. **What do different spatial components represent after applying PLS?**

***References***

**Richardson, Hilary and Lisandrelli, Grace and Riobueno-Naylor, Alexa and Saxe, Rebecca (2018). MRI data of 3-12 year old children and adults during viewing of a short animated film. OpenNeuro. [Dataset] doi: 10.18112/openneuro.ds000228.v1.0.0**

NX-421: Neural signals and signal processing

# Variant 3

## Groups: J, N, I, M, O, X, Z

***TAs of reference:***
***Ilaria Ricchi (ilaria.ricchi@epfl.ch) & Sélima Zahar (Selima.Zahar@epfl.ch)***

***Description***

In this variant, we will give you the option to choose the dataset you like the most: you can take a dataset from the weekly exercises or from the previous mini project or download another dataset from openneuro.

You are working in a Neuroimaging lab and you will be a *real* researcher: from the chosen dataset, you are asked to look for a specific scientific question that is related to the dataset but you will have to think of addressing the question with a multivariate method *(e.g. What are the brain spatial maps that are commonly present in multiple subjects while they are listening to music? What brain regions are commonly spatially active across subjects that are watching a movie?)*.

Shape the dataset in a matricial form, investigate the multiple components you are able to find and discuss the spatial maps. Afterwards compute seed functional connectivity.

***Practical part***

0. *If you have chosen a not preprocessed dataset, remember to run at least minimal steps of preprocessing.*
   *We recommend taking a preprocessed dataset so that you can focus on the rest!*

1. Briefly describe the dataset and the scientific question you want to answer.
2. List the basic steps of the multivariate method you chose and explicit the maths behind it.
3. Vary the number of components you are considering and visualise the spatial maps. Check if you are able to find meaningful maps according to the question you asked or the type of the dataset. *You can also look at the atlas to have hints on the brain regions of interest.*
4. Select a seed (voxel) and perform seed functional connectivity and perform this analysis on voxels that could be relevant for your question. Interpret the results.
5. BONUS: average the time series of multiple voxels belonging to 10 different brain regions to obtain 10 signals. Then, generate and obtain a functional connectivity matrix from these 10 regions. Use your favorite graph metric on the graph.

# NX-421: Neural signals and signal processing

***Theoretical part***

1) What are the advantages and limitations of the chosen method?
2) Why did you choose this method for the scientific question you want to address?
3) What do you observe changing the number of components? Could some components be only noise? If so, can ICA be useful for preprocessing or complementing GLM?
4) What could be the limitations of seed voxel functional connectivity? Can you think of other ways to perform functional connectivity without taking one voxel at the time?
5) BONUS: Why could it be useful to have a graph? Describe how graph theory can help investigating specific questions.