Multivariate methods to unveil the brain regions commonly spatially active across subjects that are watching a movie

Louis-Alexandre Leger, Nicolas Francio, Mehdi Hajoub, Shiyi Huang, Pingsheng Li

1 Introduction

In our project, we examined the use of multivariate methods on fMRI data from 24 subjects viewing "The Twilight Zone" movie. Leveraging the high quality and preprocessed nature of the dataset, we aimed to identify **what are the brain spatial maps that are active while watching a movie?**

2 Methods

Independent Component Analysis (ICA)

Independent Component Analysis (ICA) is a method used to separate mixed signals into ICs. ICA operates under the assumption that natural signals often have non-Gaussian distributions, which makes non-Gaussianity a valuable criterion for separating and recovering the underlying statistical Independent Components (ICs). Let the fMRI signal be represented by the space-time data matrix of measurements $\mathbf{Y}$. In the linear mixing case, we assume that the matrix can be modeled as follows:

$$\mathbf{Y} = UA^T + E^T$$

with the columns of U representing spatial sources across voxels and the columns of A representing timecourses. In spatial ICA, we assume that the columns of the matrix U are statistically independent processes, whereas in temporal ICA, the rows of A are assumed to be independent. ICA aims to simultaneously estimate the mixing matrix A such that $Y = UW$ (with $W = A^{-1}$), and the resultant components matrix U, whose columns exhibit highest statistical independence. This is achieved through the application of algorithms (e.g infomax, FastICA) which adaptively compute weight vectors w optimizing a cost function designed to maximize non-Gaussianity (measured using kurtosis or negentropy).

Pipeline

First, we conducted ICA on an individual patient, and subsequently, on a group of 10 patients. Our ICA procedure involved converting our fMRI data into a 4D matrix, followed by preprocessing the volumes. Prior to applying ICA, we also achieved centering (subtract the mean to create a zero-mean signal), whitening, and dimensionality reduction as preprocessing steps on $\mathbf{Y}$ with Principal Component Analysis (PCA), with a number of components twice that of our ICs. Subsequently, we used FastICA algorithm to obtain the ICs. We ran this entire process with and without PCA, and the results were virtually identical (PCA achieving a cumulative explained variance of 99%). The pipeline was run with k values of 4, 8, and 20, and the results are shown below.

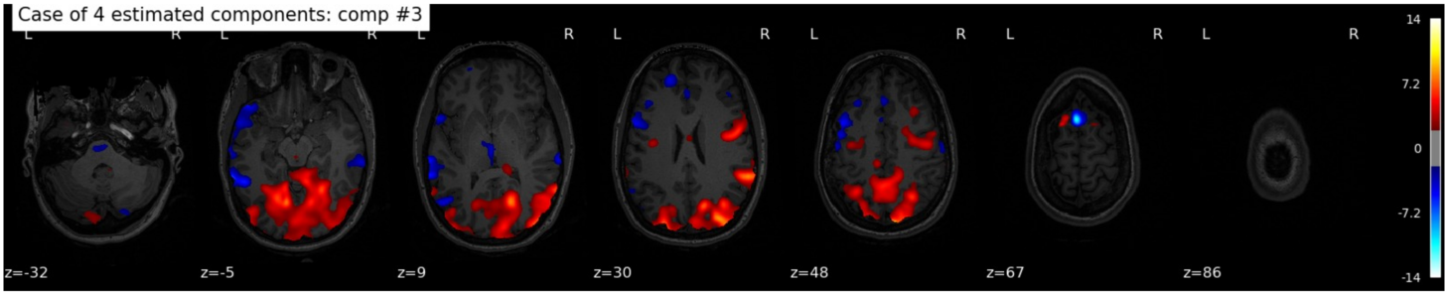
3 Results

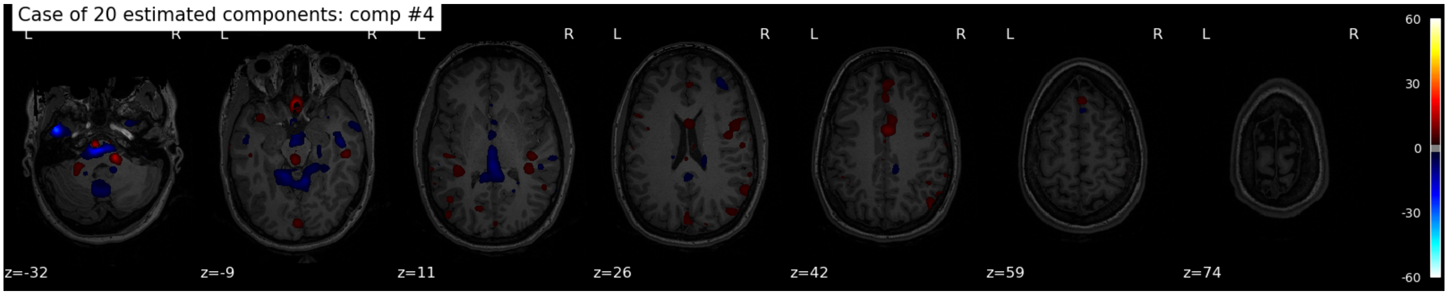
ICA and Atlas Association

As presumed, we can see the spatial maps involved in movie watching and these seem to be coherent as they obviously involve many different brain areas corresponding to the visual cortex (see figures above Table 1). To gain a deeper understanding of these results and move beyond visually descriptive analysis, we matched our ICs with a well-known atlas (Harvard-Oxford Atlas), resampling the volume for each IC to align with the atlas dimensions and determining the percentage of overlap with labeled brain regions. Table 1 shows that the most represented brain region for 3^{rd} IC is the Lateral Occipital Cortex, which is primarily involved in object recognition (see Grill-Spector et al.). Other components showed high activity in the Precuneous Cortex, associated with the default network of resting consciousness, and the Frontal Pole, linked to higher cognitive functions.

Temporal ICA

Additionally, we conducted temporal ICA to analyze functional data, identifying three distinct and mutually exclusive brain states that can be visualized in Fig. 1. Interestingly, these states show periodic signals, and notably, IC3 exhibits a significant deviation from the mean towards the movie's end, suggesting a possible reaction to an auditory stimulus or anticipatory brain activity.





Brain Regions	Lat. Occipital Cortex (SD)	Lat. Occipital Cortex (ID)	Precuneous Cortex	Occipital Pole
$\%IC3$	10.0	4.4	3.5	3.3

Table 1: **A)** 3^{rd} IC from $k = 4$ ICA components. **B)** 4^{th} IC from $k = 20$ ICA components. The background representation of the anatomical brain was obtained after brain extraction on one subject. **C)** Table representing top 4 brain parts 3^{rd} IC for $k = 4$

Functional Connectivity derived from ICA

Once we had the percentages of our ICs mapped onto the different annotated brain areas, we also quickly explored the possibility of illustrating the brain's functional connectivity network from our ICA. We did this by first selecting the top 5 brain parts by IC composition (%) and then creating a bipartite graph between the different brain parts and IC's. After creating this graph, we projected the graph onto the brain areas to reveal the functional connectivity network of our ICs. Fig 2 allows to illustrate Donald Hebb's dogma that "N*eurons that fire together, wire together*" and transition into seed voxel functional connectivity.

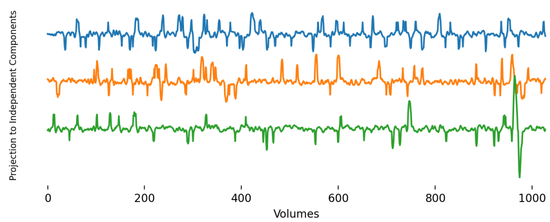


Figure 1: Temporal ICA.
Blue represents IC1, Orange IC2, and Green IC3.

Seed voxel functional connectivity

We shifted from using a single, highly variable voxel to a specific region as our seed for consistency, specifically seed 10, the posterior division of the superior temporal gyrus (Wernicke's Area), which is key to listening and language comprehension. Mean correlation coefficients between this area and others were calculated over 24 subjects, with detailed results displayed in Fig. 3 and additional data in the code. Strong functional connectivity observed between Wernicke's Area and associated brain regions during movie viewing indicates a multimodal brain response to complex auditory, linguistic, memory, emotional, and social stimuli presented in the film.

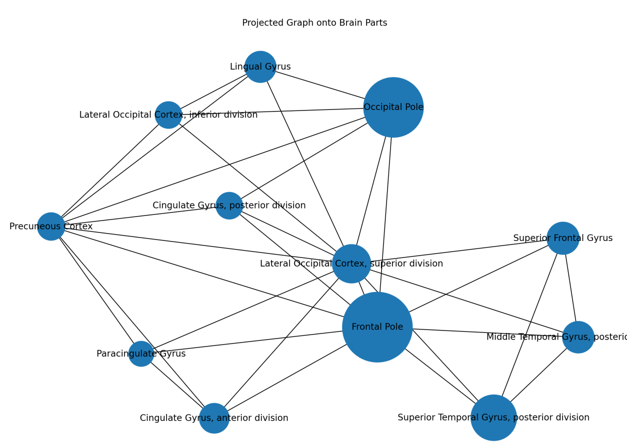


Figure 2: IC connectivity network

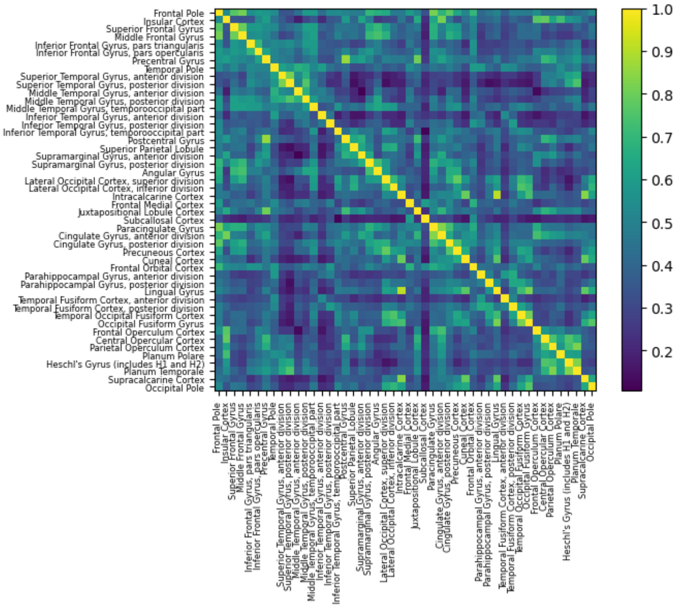


Figure 3: Functional connectivity map within 48 regions

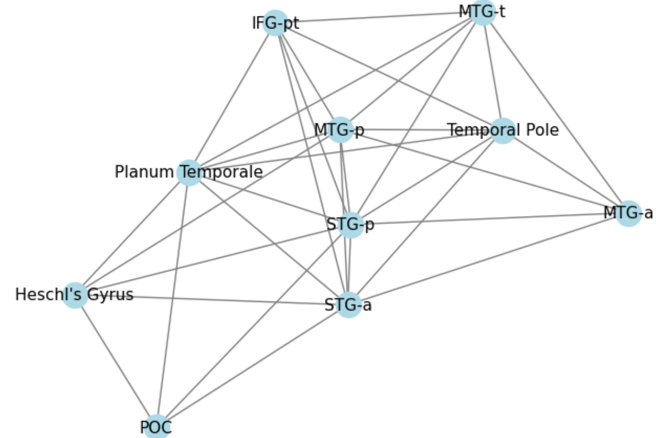


Figure 4: Graph based on functional connectivity within 10 regions

BONUS: 10-node subgraph related to Wernicke's Area

We identified the top 9 regions most correlated with the Superior Temporal Gyrus, posterior division, creating a set of 10 regions to construct a Functional Connectivity (FC) graph. This FC matrix was then converted into a binary adjacency matrix (edge density=0.3) for graph analysis. The visualization in Fig 4 shows high connectivity within auditory structures (e.g., STG, MTG). Notably, the degree distribution in this 10-node subgraph is more skewed compared to the full 48-node graph, indicating distinct connectivity patterns in local versus global brain networks (details in the code).

4 Discussion

In fMRI, ICA is a valuable tool for addressing our initial question since it aims to distinguish spatial brain patterns even when multiple regions are engaged simultaneously and especially without any prior hypothesis about brain regions involved in watching a movie, making it suitable for discovering patterns that naturally emerge from the data. However, interpreting these sources can be challenging. One notable limitation of ICA is the selection of the number of components to study. This parameter requires careful consideration because an insufficient number of components might not yield clear, distinguishable spatial maps representing meaningful neural activity. Conversely, an excessive number of components may capture signals associated with non-neural factors such as head motion, scanner noise, or physiological artifacts. We may vary the number of components as we did in our study or employ more advanced methods, developed to specifically address this challenge (see Hui et al.).
We observed that increasing the number of components didn't help reveal new spatial brain patterns. Instead, we noticed that beyond a certain threshold, additional ICs started to capture noise. Identifying these noise-related ICs allows effective noise separation and removal, leaving a cleaner signal. This makes ICA a valuable tool for noise reduction during data preprocessing or as a complement to the General Linear Model (GLM). As forecast in the code report we tried performing a GLM to compare its performances with and without pre-ICA. Unfortunately the design of experiment of our dataset didn't allow for relevant GLM results as the subjects were only watching a movie throughout almost the whole acquisition time.
Also, single voxel seed is not robust across different subjects due to vast variability. We can take brain region or cluster based analysis instead. Or we can use ICA based analysis which can work without predefined seeds. Using graph theory in neuroscience allows for the mapping and analysis of brain networks, offering insights into how brain regions interconnect and work together. Also, it provides a framework for understanding the organizational structure of brain functions, assessing the impact of diseases on connectivity, and exploring the relationship between neural networks and cognitive tasks. This approach can reveal the dynamic nature of brain connectivity and its adaptation during various activities and conditions.

References

Kalanit Grill-Spector, Zoe Kourtzi, and Nancy Kanwisher. The lateral occipital complex and its role in object recognition. *Vision Research*, 41(10-11):1409–1422. doi: 10.1016/s0042-6989(01)00073-6.

Mingqi Hui, Rui Li, Kewei Chen, Zhen Jin, Li Yao, and Zhiying Long. Improved estimation of the number of independent components for functional magnetic resonance data by a whitening filter. *IEEE Journal of Biomedical and Health Informatics*, 17(3):629–641. doi: 10.1109/JBHI.2013.2253560.