Exploration of Neuropixels Recordings of Mice during Spontaneous Behaviors

Toan Tran[1], Maurizio De Pittà[2], Gordon Erlebacher[1]

[1]Department of Scientific Computing, Florida State University, Tallahassee, United States
[2]Basque Center for Applied Mathematics, Bilbao, Spain

Introduction

Spontaneous neural activity is known to occur in the absence of external stimuli and has noisy pattern. Whether it contains information about the workings of the brain is an open question which has received increasing attention. Our preliminary work involves analyzing recordings of spontaneous activity in mouse brains in search for distinguishable structure, which may indicate functional dynamics in such activity. We employed data analysis techniques that captures the topological structure of the data and obtained complementary results.

Experimental Data

Spontaneous activity is typically recorded from specimens with resting state brains [1] which ranges from those of anaesthetized and sleeping animals to those which are awake but confined to an environment with limited exposure to stimuli. In our work, we use the publicly available dataset containing Neuropixels recordings of three mice during a period of spontaneous behavior [3]. The mice were free to run on an air-floating ball in darkness while neural activity across several brain regions was simultaneously recorded. After preprocessing, the data takes the form of multiple time series (one per neuron) with similar timescale. For all our analysis so far, each time series or activity vector is considered a data point and, thus, the activity of all neurons a high-dimensional point cloud, the topological structure of which is of interest to us.
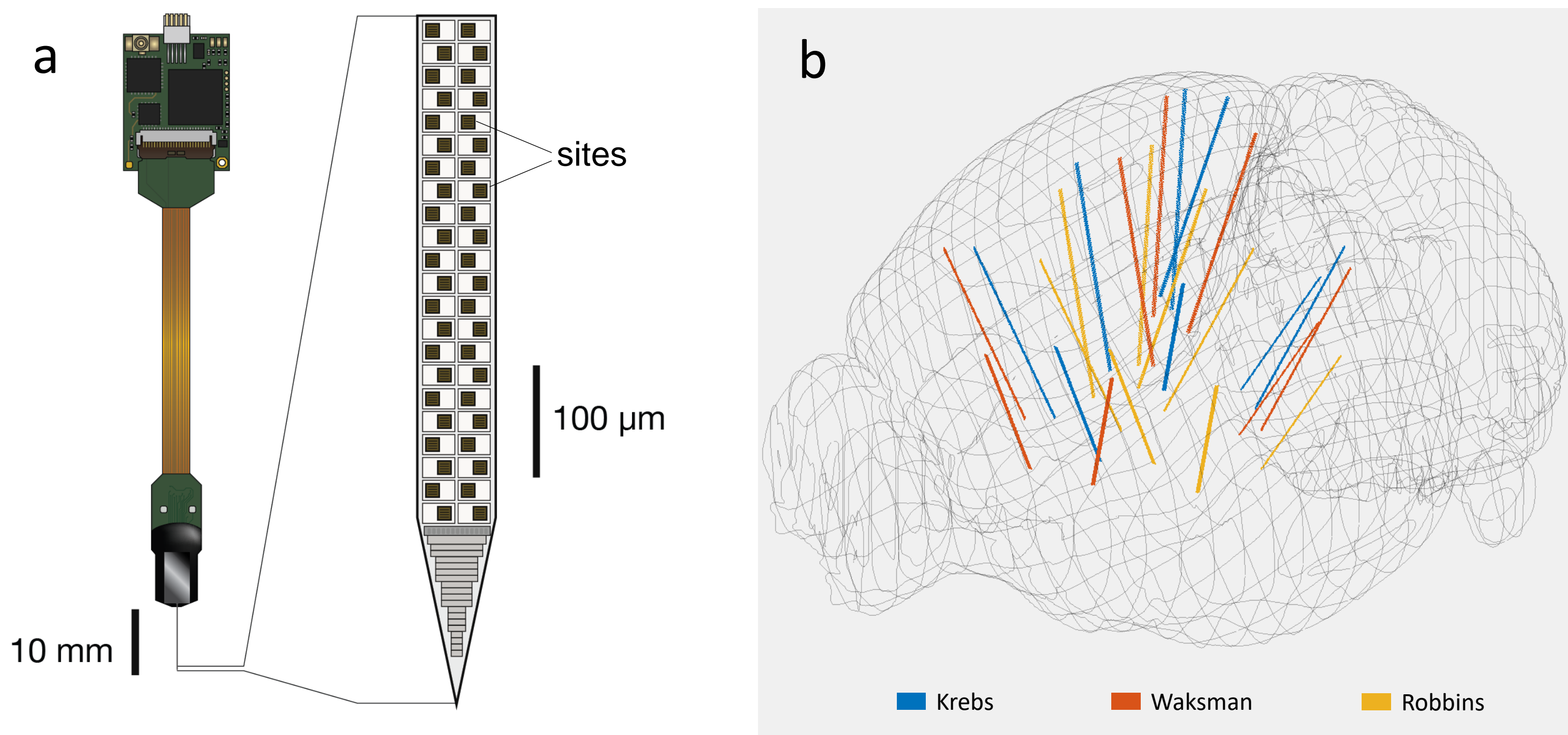

Figure 1: (a) Schematic of a Neuropixels 1.0 probe with the zoom-in view of the 10-mm shank covered with switchable recording sites. Sites are small and dense enough to reliably obtain well-isolated recordings of individual neurons' firing. (b) Illustration of the color-coded shanks inside a wire mesh brain model.
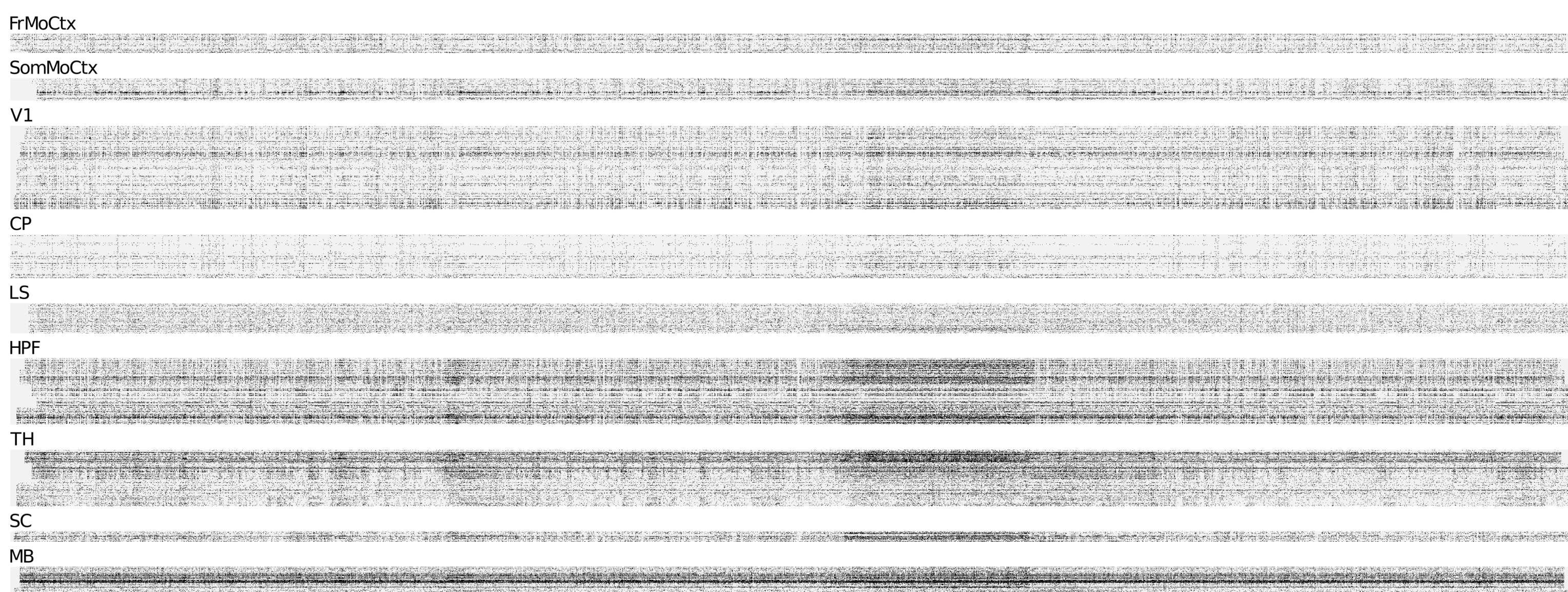

Figure 2: Raster plot of Krebs' neural activity, pooled from all 8 probes and then grouped based on brain region.

Correlation-Based Clustering

We examined the connectivity of the data, which is characterized by the persistence of 0-dimensional homology, using hierarchical clustering. The algorithm constructs a tree-like diagram called a dendrogram which is a multi-scale representation of the point cloud. Imagine standing close to a wall with a two-dimensional point cloud drawn on it and then slowly moving away until all the points becomes one from your perspective. This is akin to how the algorithm captures the "hierarchy" of the data.

Clustering algorithms take a distance matrix as input and we constructed ours using the following definition of correlation distance

$$d(X,Y) = 1 - \rho_{X,Y}$$

where $\rho_{X,Y}$ is the Pearson's correlation

$$\rho_{X,Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}$$

A sorting scheme was derived from the dendrogram which we applied to three "relational" matrices in which each element is

- Pearson's correlation between two activity vectors
- Log of mutual information between two activity vectors with mutual information defined as

$$I(X;Y) = \sum_{x,y} P_{XY(x,y)} \log \frac{P_{XY(x,y)}}{P_X(x)P_Y(y)}$$

- Physical distance between the neurons in microns computed from the recording site locations in Allen Common Coordinate Framework (CCF)

We computed the matrices, then identically permutate their columns and rows to reflect the clustering result. That the clustering result observed from all three matrices are complementary, if not consistent, shows promise and provides potential support for future analysis.
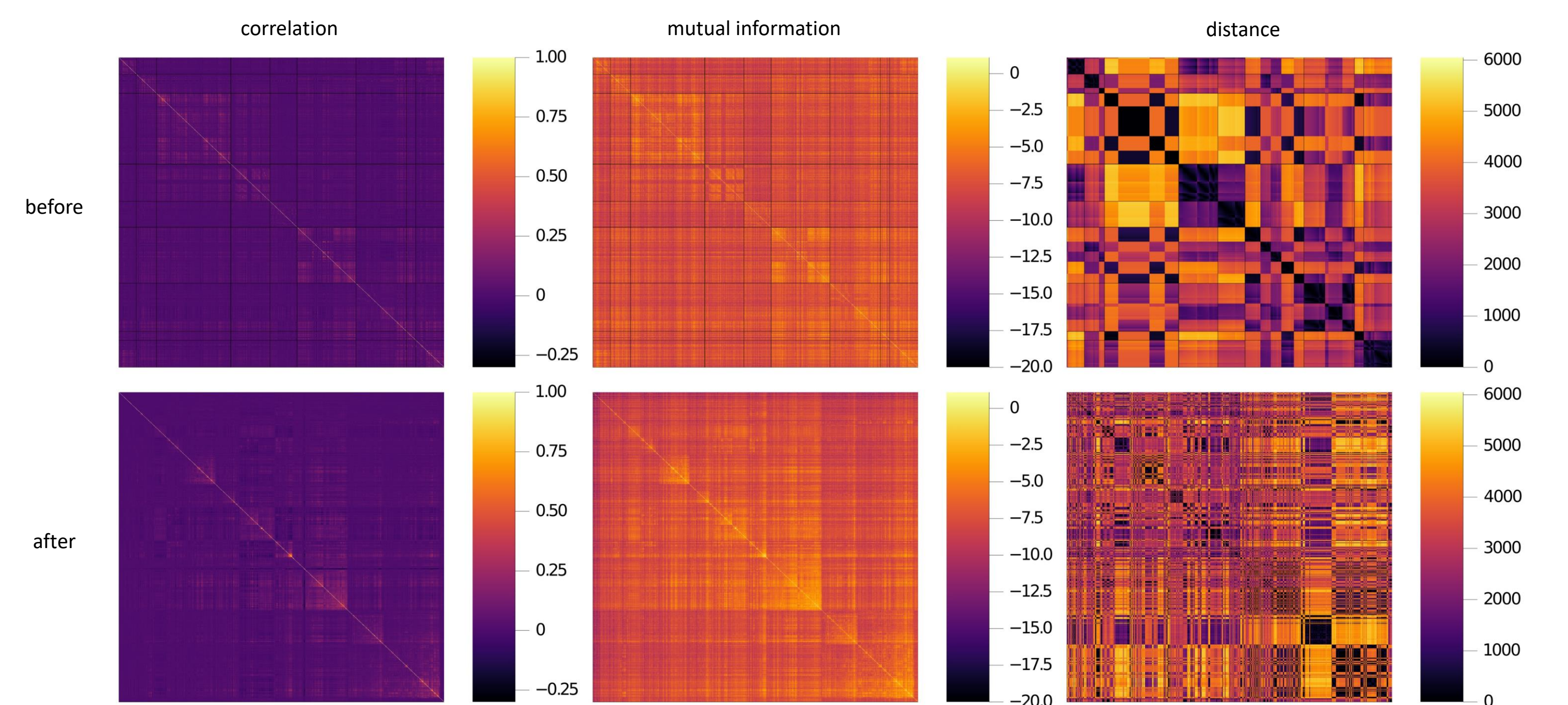

Figure 3: "Relational" matrices of Krebs before and after clustering is applied. Note the black vertical and horizontal lines in matrices in the first row separating each of them into regions along the diagonal. Neuron pairs in these regions are in the same brain area which explains the existence of clusters even before clustering.
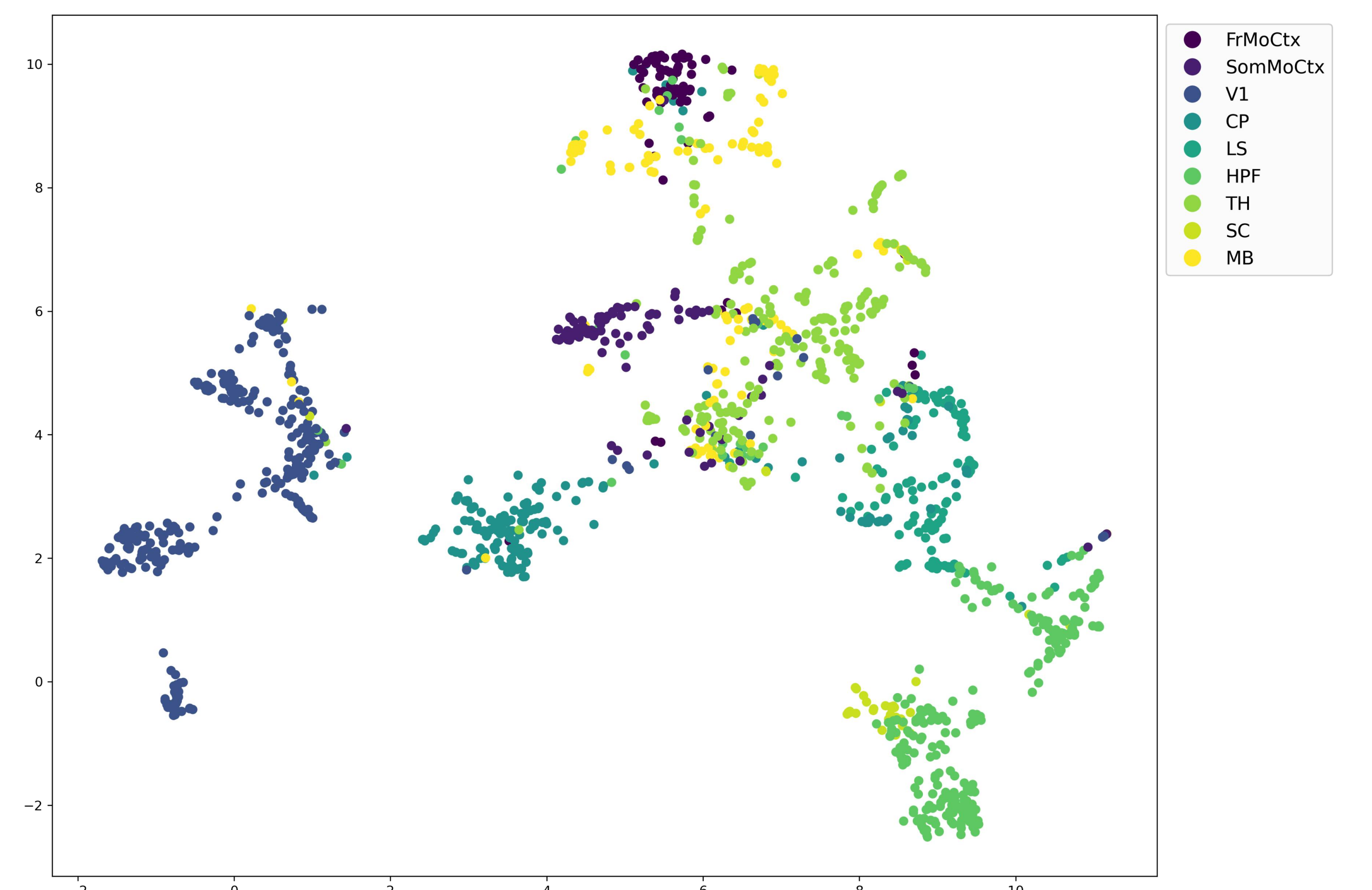
UMAP

Figure 4: 2D projection of activity vectors of Krebs produced by the UMAP package with the following parameter values: n_neighbors=5, min_dist=0.1, n_components=2, metric=correlation. Embedded points are color-coded based on brain regions.

UMAP is a versatile dimension reduction algorithm with many applications in data analysis. Its theoretical root in Riemannian geometry and algebraic topology enables it to find a low dimensional embedding that preserves the essential topological structure of data. This makes it a powerful tool for visualizing and exploring data.

The algorithm has primary parameters which can considerably affect the output of the projection. The construction of a neighborhood graph is crucial step in creating the embedding and so is the number of nearest neighbors of the data points as a parameter. Specifically, it controls the balance between local and global structure. We set the parameter to a low value which priorities local structure and results in better clustering in the embedding. The minimum distance between dimensionally reduced points was also kept low for more distinct clusters.

The brain-region-based coloring of the projection shows evident correlation between the activity of a neuron and its location in the brain. If the recorded activity is truly spontaneous, meaning it is not confounded with its induced counterpart, this suggests that there are multiple distinguishable regimes in spontaneous activity based on brain region.

Conclusion & Future Work

The neural activity under investigation may not be intrinsically spontaneous and likely corresponds to the motor movements such running and sniffing. Whether or not this is true and to what extent are important questions to answer for the legitimacy of future findings. There are videos recordings of mice behaviors included in the dataset which can be examined in conjunction with the neural activity for this purpose.

Our analysis so far treats the full recordings of the neurons as data points and, thus, do not consider time as a factor, at least not explicitly. These recordings are time series and may exhibit interesting properties such as autoregression; therefore, finding a way to incorporate a time component into our analysis is the next step for this project. Potential candidates include autocorrelation analysis and time-delay embedding.

References

[1] Xiao Liu, Xiao-Hong Zhu, Peihua Qiu, and Wei Chen. A correlation-matrix-based hierarchical clustering method for functional connectivity analysis. Journal of neuroscience methods, 211(1):94–102, 2012.

[2] Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction, 2020.

[3] Nick Steinmetz, Marius Pachitariu, Carsen Stringer, Matteo Carandini, and Kenneth Harris. Eightprobe Neuropixels recordings during spontaneous behaviors. 3 2019.

[4] Carsen Stringer, Marius Pachitariu, Nicholas Steinmetz, Charu Bai Reddy, Matteo Carandini, and Kenneth D Harris. Spontaneous behaviors drive multidimensional, brainwide activity. Science, 364(6437), 2019.