Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Yating Yang April 10, 2018

Low-Dimensional Dynamics Encoding in Human Brain Data

by

Yating Yang

Gordon Berman, Ph.D.
Adviser

Department of Biology

Gordon Berman, Ph.D.

Adviser

Astrid Prinz, Ph.D.

Committee Member

Shella Keilholz, Ph.D.

Committee Member

2018

Low-Dimensional Dynamics Encoding in Human Brain Data

By

Yating Yang

Gordon Berman, Ph.D.

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Sciences with Honors

Department of Biology

2018

Abstract

Low-Dimensional Dynamics Encoding in Human Brain Data

By Yating Yang

The brain is a complex, multiscale structure where different levels work together to produce behaviors and cognition. People have placed much emphasis on the functioning of individual parts of the brain, but in order to achieve one of the ultimate goals of neuroscience, we need to consider the brain dynamics as a whole, and focus not only on each part of the brain but the functional connectivity within the brain dynamics. Here, we find lower dimensional states in ECoG data, looking for long-time scale patterns.

One of the biggest challenges we face is the complexity of the data; the brain data is high dimensional, and it contains structures across different length scales and different time scales. To analyze the data, we performed different dimensionality reduction methods to visualize brain dynamics in a lower dimensional state. Then, by looking at the embedding space from t-SNE, we found clustering features of embedding space are contiguous and also discrete. By looking at the amplitudes of all neural channels, we could group clusters on the embedding space together into different brain regions, with right hippocampus dentate gyrus (DG) being most dominant across the map. Then, we compare the transitions of the dataset to that of Markovian model generated from the data; we observe the dataset contains a much longer time scale, far beyond what Markov model can predict, indicating the presence of a complex dynamical structure that bridges scales.

Low-Dimensional Dynamics Decoding in Human Brain Data

By

Yating Yang

Gordon Berman, Ph.D.

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Sciences with Honors

Department of Biology

2018

Acknowledgements

First, I would like to express my gratitude to my advisor Prof. Gordon Berman who has given me generous help and support in the past three years. When I first joined the lab, I had very little knowledge in coding and modeling, and as my mentor, he taught me everything, and offered me the opportunity to develop my interests in computational neuroscience. He is extremely knowledgeable and patient, and his humor makes our lab such a fun place to do research in.

Secondly, I would like to thank my committee members, Prof. Astrid Prinz and Prof. Shella Keilholz for their genuine advice. I would also like to thank our collaborators, Dr. Robert Gross, Dr. Jon Willie, and Dr. Cory Inman for providing us with the ECoG data from Emory Hospital.

Last but not least, I would like to thank all members in Berman's Lab, Katherine Overman, Josuan Calderon, Itai Pinkovezky, Sena Agezo, Liz O'Gorman, Jeff Gao and Hanyao Sun. Special thanks go to Josuan who has helped me a lot in analyzing data in the lab and answered my endless questions and emails.

Table of Contents

Introduction

The human brain is a complex multiscale structure in terms of temporal and spatial domains; molecular, cellular and neuronal phenomena work together to form the basis of cognition and behaviors (Bassett et al, 2011). Human behaviors are determined by more than the pure sum of each parts of the brain, but also by the underlying connectivity across different regions. To understand the human brain's ability to interact with one's behavior at certain time point, it is crucial to understand not only certain parts of the brain structure but the brain as a whole. One of the ultimate goals in neuroscience is to understand the brain-mind mechanisms; we are here to look at the brain as a whole and work towards the goal of decoding brain dynamics.

Complicating this endeavor, however, is the complexity of the human brain. In terms of spatial scaling, it demonstrates a hierarchical structure which facilitates global communication (Bassett 2014). In terms of temporal dynamics, it shows a broad range. For example, the highest frequency gamma band (>30 Hz) is shown to be important in cognition of binding of information from sensory, whereas the beta (12-30 Hz), alpha (8-12 Hz), theta (2-4 Hz), and delta (1-2 Hz) bands each respond to different but complementary functions (Uhlhaas PJ, et al. 2008). To analyze brain data, we face the challenge of its inherently high-dimension and its varying structures across different length scales and different time scales.

In recent years, interest in examining interactions between brain dynamics and human behaviors is mounting. Methods have been invented to measure behaviors and analyze them. For example, freely-moving behaviors can be mapped into a space in which an animal's action is

classified and organized upon its underlying structures of postural movement data (Berman et al, 2014). Unsupervised approaches to decode neural states from brain recording data have been introduced to help analyze interactions between brain dynamics and human behaviors without tedious and time-consuming manual labeling of video and audio (Brunton et al, 2016). Various computational methods have been created to infer the connectivity from brain data (Magrans de Abril, 2017). Moreover, the advent of invasive and noninvasive neuroimaging techniques allows scientists to observe a wide range of activities within human brains. Diffusion tensor imaging (DTI) is used to characterize changes in microstructures within the neuropathways (Alexander, 2017). Electrocorticography (ECoG) is a type of electrophysiological monitoring technique that places electrodes directly onto the surface of the brain to record electrical activity; it can provide brain signals with high signal-to-noise ratio, and high spatial and temporal resolution (Hill 2017).

Here in our lab, we have the access to ECoG data of patients with epilepsy; local field potential (LFP) signals were measured during the ECoG recording (Figure 1). LFP is the electric potential in the extracellular space of the neurons in the brain (Destexhe 2013). In the experiment, our goal is to find lower-dimensional structures and long time-scale patterns within the brain data and decodes human behaviors from brain dynamics. To accomplish this, we have to overcome the challenge of inherent high dimension and complexity of human brain by performing different data analysis methods.

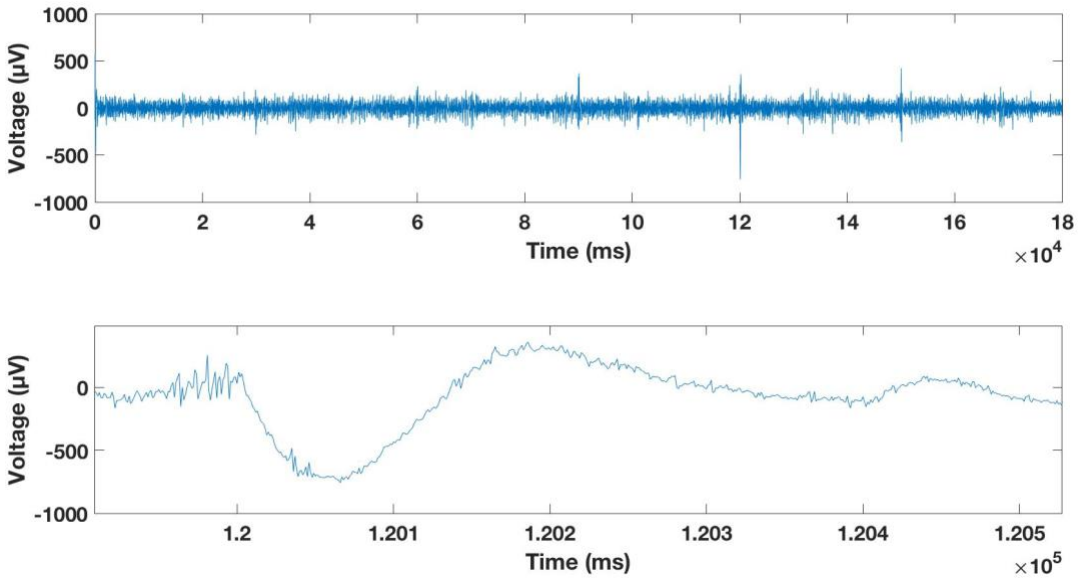


Figure 1 (A) LPF signals in the first 6-minute data. (B) Zoomed in one of the sections of the abrupt potential changes.

Materials and Methods

Data collection and pre-processing

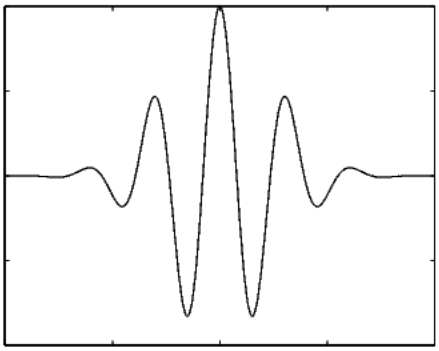
Thanks to our collaborators, Dr. Robert Gross, Dr. Jon Willie, and Dr. Cory Inman, we have the access to ECoG data from epilepsy patients in Emory Hospital. We focused on one of the patients' 24-hour ECoG recording along with the video tape of her movement at the same time. Note that during the time of recording, there was no seizure activity. The patient had electrodes implanted on the exposed cortex of her brain, and on one electrode, there were multiple recording sites (channels). There were 128 neural channels in the recording, and the data was taken at 500Hz. However, not every channel is the actual recording of neural activities; some channels are taking signals from both sites on the same electrode, some channels are set to record the activity of previous channels, and some channels are for the purpose of stimulation.

Therefore, we excluded channels with these types of artifacts, and only preserved 104 channels in the analysis to follow.

Among these 104 channels, the data needed to be filtered first; the harmonic frequency and everything below 1Hz were filtered out in this step to decrease the noise, like movement artifacts caused by appliance, in the data. We also filter out 60 Hz noise from the electronics and its harmonic frequencies (Sankaran, 1999). The filtering window around each harmonic is 1 Hz.

Wavelet transforms

We chose to use wavelet transform in our analysis, because for most natural signals, there are slowly varying components interspersed by abrupt changes, but it is hard to separate major events from noise by only looking at the time domain. Wavelet transforms allow us to look at time and frequency domain at the same time. To reduce the dimensionality of neural data, we calculated the amplitudes of the continuous wavelet transform for each neural channel. We used the Morlet wavelet in our continuous wavelet transform (Figure 2).



The continuous wavelet transform is able to build a time-frequency representation that is good at describing brain dynamics occur at multiple time scales. Moreover, this particular

choice of continuous wavelet transform is able to isolate epochs of periodic dynamics (or "chirps"). In our research, we observe 25 frequency channels per neural channel, spaced between 1 Hz and 250 Hz, where the maximum frequency is determined by Nyquist frequency of the system.

In details, we calculate this Morlet continuous wavelet transform, $W_{S,\tau}[y(t)]$, by

$$W_{S,\tau}[y(t)] = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} y(t)\psi * (\frac{t-\tau}{s})dt \quad \text{(Eq. 1)}$$

where

$$\psi(\eta) = \pi^{-\frac{1}{4}} e^{i\omega_0\eta} e^{-\frac{1}{2}\eta^2}. \quad \text{(Eq. 2)}$$

In the equations above, y(t) is the neural time series, s is the time scale, and ω_0 is a non-dimensional parameter which we set to 5 here.

Time scale, s, is related to the Fourier frequency, f, by

$$s(f) = \frac{\omega_0 + \sqrt{2+\omega_0^2}}{4\pi f}. \quad \text{(Eq. 4)}$$

The scaler function C(s) we use to normalize the Morlet wavelet across frequencies (Berman, 2014) is

$$C(s) = \frac{\pi^{-\frac{1}{4}}}{\sqrt{2s}} e^{\frac{1}{4}\left(\omega_0 - \sqrt{\omega_0^2+2}\right)^2}.$$

Spatial Embedding

The power spectrum includes 25 frequency channels per each of the 104 neural channels, making each data point a 2600-dimensional feature vector in time space. One of the common dimensionality reduction method is t-distribution stochastic neighbor embedding (t-SNE), which aims to take data points from a high dimension space and embed it into a much lower dimensional space (van der Maaten, 2008). t-SNE models a random walk on the data points in the high dimensional space, attempting to preserve that random walk as best as possible in a low dimensional embedding.

The first stage of the t-SNE algorithm is to compute the probability distribution in which objects pairs in the high dimension that are similar have higher probability of being picked first. We define the transition probability from time point $\mathrm{t_i}$ to time point $\mathrm{t_j}$ as $\mathrm{p}_{j|i}$ which is proportional to the Gaussian kernel of the distance between these two time points

$$\mathrm{p}_{j|i} = \frac{\exp(-d(t_i,t_j)^2/2\sigma_i^2)}{\Sigma_{k\neq i}\exp(-d(t_i,t_j)/2\sigma_i^2)}. \quad \text{(Eq. 5)}$$

All self-transitions (i.e. $\mathrm{p}_{i|i}$) are set to zero.

Then, t-SNE embeds the data points into the lower dimensional space with new transition probability, $\mathrm{q}_{j|i}$, similar to the $\mathrm{p}_{j|i}$. However, the $\mathrm{q}_{j|i}$ here is proportional to a Cauchy distribution.

We define the distance function, $d(t_i, t_j)$, as a Kullback-Leibler (KL) divergence between two featured vectors.

$$d(t_1, t_2) = D_{KL}(t_1||t_2)$$

The algorithm then aims to minimize the cost function (van der Maaten, 2008)

$$C = D_{KL}(P||Q) = \sum_{ij} p_{ij} \log(\frac{p_{ij}}{q_{ij}}) \quad \text{(Eq. 6)}$$

Where $p_{ij} = \frac{1}{2(p_{j|i}+q_{j|i})}$

$$q_{ij} = \frac{\left(1+\Delta_{ij}^2\right)^{-1}}{\sum_k \sum_{l \neq k} \left(1+\Delta_{k,l}^2\right)^{-1}}$$

And Δ_{ij} is the Euclidean distance between *i* and *j* in the embedding space.

Training Set Generation and Re-embedding

For each channel, we have 43,200,000 (500 Hz * 24 hour *60 min * 60s) data points, and we have 104 neural channels to work with. Due to the limitation of computer memory, we cannot compute the t-SNE for the whole data set at once. Therefore, we use sampling techniques to generate a training set comprised of 36,000 data points and re-embed the remaining data point into the space created by the training set. We re-assembled the 24-hour data into 240 data files, and each file contains a 6-minute recording. To make sure that the training set can represent the whole data set, we did not randomly select 36,000 data points. Here, we used subsampling method in which we separated the whole dataset into 240 files and performed t-SNE on each file. Then, we plot out the embedding space of that file, and partition the space into various regions using watershed transformation (Meyer, 1994). Data points were taken from every region of the embedding space proportional to the size of the region to build the training set.

Then, to add more points into the space, we re-embed the remaining data set into the space by running t-SNE for each file. We define z as a new featured vector to be embedded into

the space according to the mapping between vectors in the training set (X) and embedding by t-SNE (X'), and ζ be the embedding of z that need to be computed. The cost function is calculated by

$$p_{i|j} = \frac{\exp(-d(z,j)^2/2\sigma_z^2)}{\sum_{x\in X}\exp(-d(z,k)^2/2\sigma_z^2)}$$

Where $d(z,j)$ is the KL divergence between z and $x \in X$. Similarly, the transition probability is calculated by

$$q_{j|\zeta} = \frac{\left(1+\Delta_{\zeta,j}^2\right)^{-1}}{\sum_{x\prime\in X\prime}\left(1+\Delta_{\zeta,x\prime}^2\right)^{-1}}$$

Then, we minimize the KL divergence between the transition probability in two spaces.

Markovian Models Generation

The Markovian model data is generated by using the similar method in Shannon's work on information theory (Shannon 1948). The simulated data is generated by randomly selecting a state from the watershed regions on the embedding space which is the list of regions that visited during embedding process, and the states come next are being randomly chosen next. This process continues until the generated sequence is the same size of the original data set.

Transition Matrices and Non-Markovian Time Scale

To calculate the transition matrix over different time scales (τ), we compute

$$[\mathbf{T}(\tau)]_{i,j} \equiv p(S(n+\tau) = i | S(n) = j) \qquad \text{(Eq. 7)}$$

which describes the probability an object transitions from i to j after τ time steps. This matrix can also be analyzed in terms of their eigenvalues and eigenvectors

$$\left[\mathbf{T}(\tau)_{\mathrm{i,j}}\right] = \sum_{\mu} \lambda_{\mu}(\tau) u_i^{\mu}(\tau) v_j^{\mu}(\tau) \qquad \text{(Eq. 8)}$$

Here, $\mathbf{u}^{\mu}$ and $\mathbf{v}^{\mu}$ are the eigenvectors, and $\lambda_{\mu}(\tau)$ is the eigenvalue. Since probability is conserved in the transitions, the largest eigenvalue $\lambda_1(\tau) = 1$, and all other eigenvalues are less than 1, describing the loss of predictability.

Results

Embedding space dynamics

We generated a probability density map by convolving data points in the embedded space under a Gaussian of small width (σ = 2, Figure 2). The embedding space is not uniformly distributed, but contains multiple local maxima, illustrated in red; these regions have the potential to represent significant brain activity or even distinctive human behaviors.

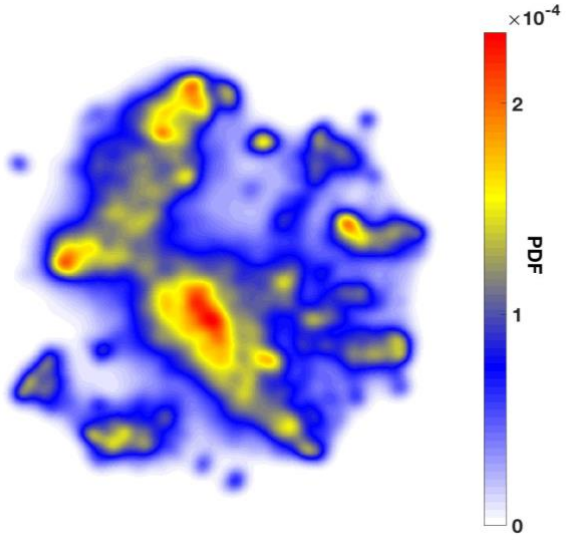


Figure 2. Embedding space probability density function (PDF) of the whole dataset.

By partition the space using a watershed transformation (Meyer, 1994), we divided the space into 60 regions with each region has their own peak points with each region has their own peak points (Figure 3A). Then, we can extract data points from each region, and study the properties of the brain dynamics in each region. The mean wavelets are calculated and plotted for 60 regions (Figure 4). Since we used 25 frequency channels per neural channel, every 25 points on the x-axis belongs to one neural channel; peaks occur at the similar position on continuous regions indicate there might be a pattern of active channels through the embedding space.

After calculating the mean wavelet for each of the 104 neural channels in each region (Figure 4), the embedding space can be further clustered based on the channels with the highest mean wavelet amplitude (Figure 5). According to the labels of the channels, we can label the clusters of the embedding space. The embedding space shows clustering features of channels in different brain regions that are both continuous and discrete, with right hippocampus DG contain the continuous trend cross the embedding space. Along the continuous dynamical structures, we observe that the overall amplitudes are about the same (Figure 6), if we look at the mean amplitude of each neural channels in this continuous trend, despite right hippocampus DG being the most active, the left hippocampus is the next most active brain structure.

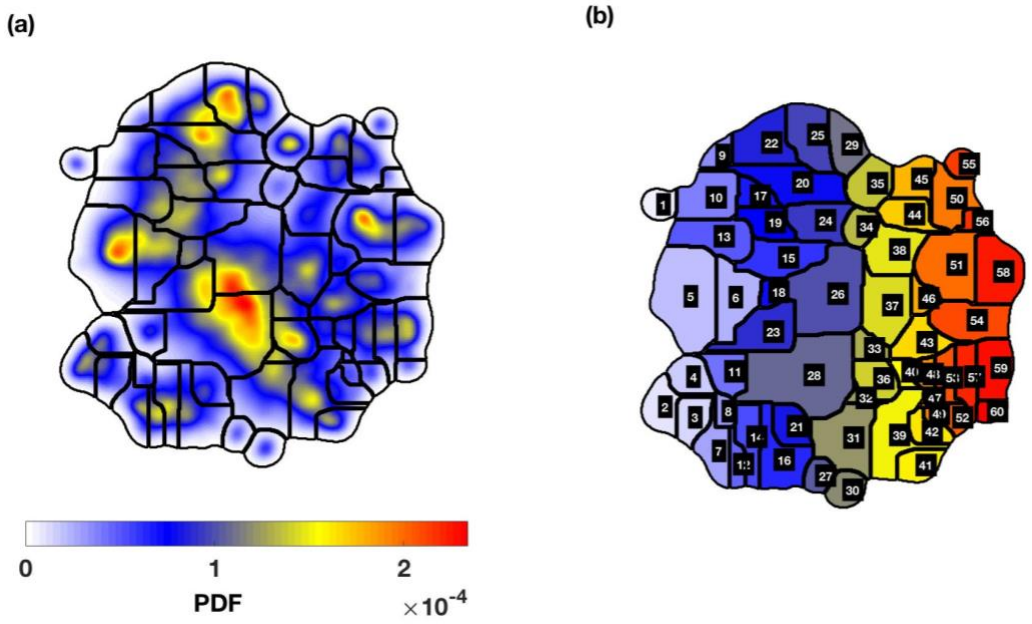


Figure 3. Segmentation into 60 regions. (a) Boundary obtained from performing a watershed transformation on the PDF from figure 1. (b) Number index of the 60 regions. The sequence of the index is determined by the x-axis from left to right.

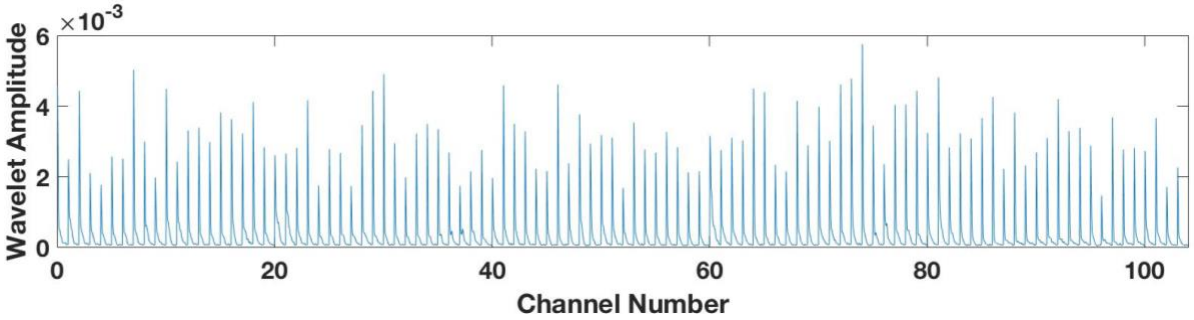


Figure 4. Mean wavelet amplitude of region 1.

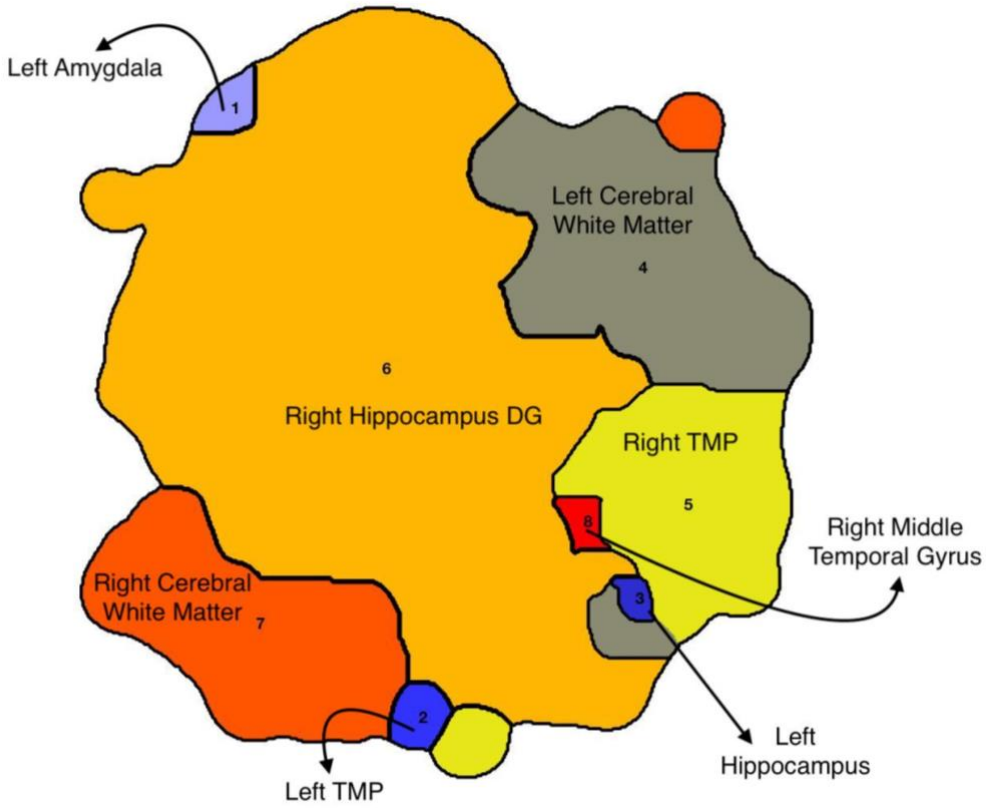


Figure 5. Clustering based on the mean wavelets of each neural channels in each region. We labeled the neural channel with the highest mean wavelet in each region, and cluster the space based on the label of the neural channels.

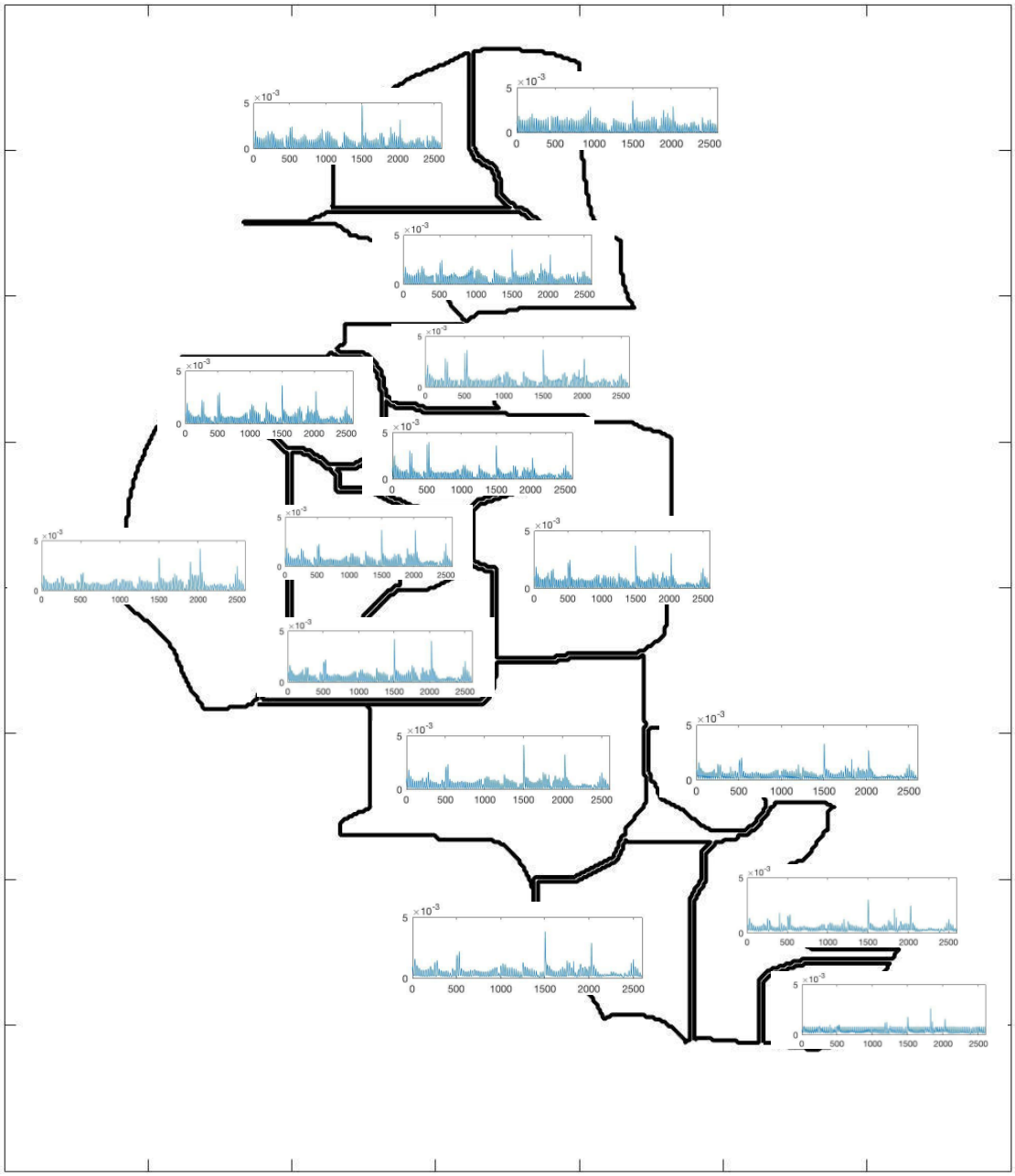


Figure 6. Frequency spectrum in right hippocampus regions, along the continuous trend.

By calculating the velocity within embedding space and plot the velocity on a histogram (Figure 7), we observe a two-state pattern: resting states vs. moving states, with a higher distribution in resting states.

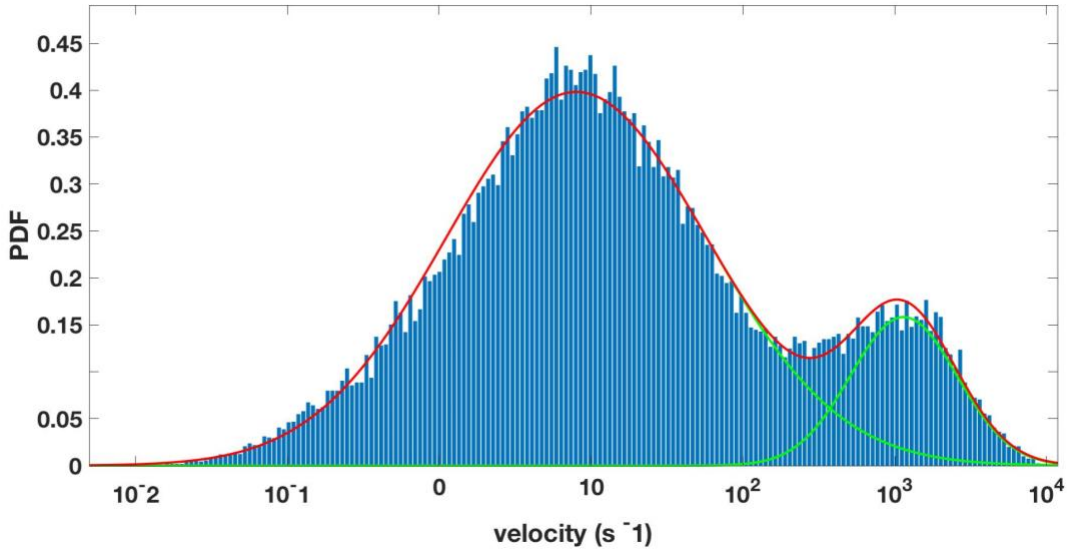


Figure 7. Histogram of velocity in the embedding space fitted into a two-component Gaussian mixture model. The blue bars represent the measured probability distribution, the red line is the fitted model, and the green lines are the mixture components.

Transition matrices and non-Markovian time scales

Plotting the transition matrix on the embedding space (Figure 8A), we observe the transitions are mostly localized with large probability transitions travel between nearby regions, and the clusters are contiguous in the embedding space. Each red dot represents a local peak probability density function, and the black lines represent the transition probability cross different regions. The thickness of line is proportional to the transition matrix $T(\tau = 1)_{ij}$.

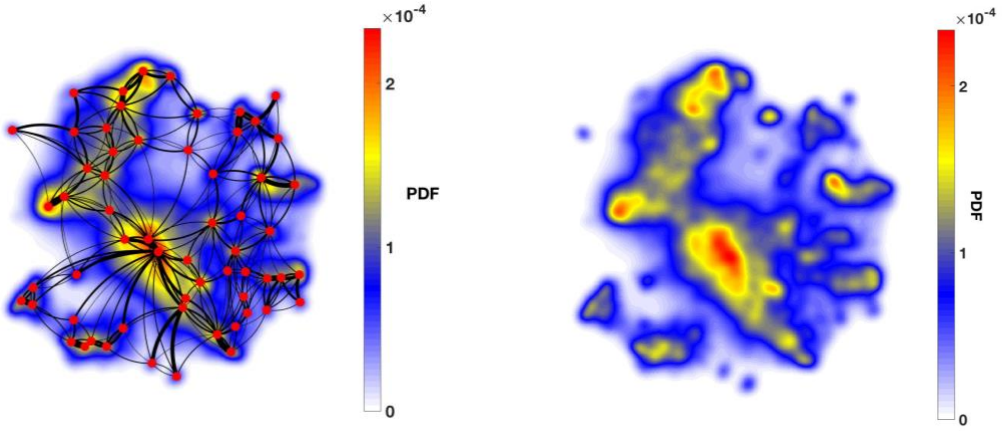


Figure 8. (A) Transition rates are plotted on the embedding space. Each red dot represents a maximum of the local PDF, and black lines represent the transition probability between two regions. The right-handed curvatures imply the direction of the transition. (B) The original embedding space without any annotation of transitions or local peak points.

If we plot out transition matrices of 1 step (τ = 1), 100 step (τ = 100), 10000 step (τ = 10000), and 100000 steps (τ = 100000) (Figure 9), we can tell that when τ = 1, there are diagonal block-structures, which implies one state only goes to a small amount of other states, while when τ gets larger, it becomes increasingly difficult to predict the future state of the brain data. Although the transition matrices did not change much after τ = 100, and the block structures

from τ = 1 figure begin to disappear, it still contains a great deal of non-randomness, indicating the existence of long time scale pattern in the dataset.

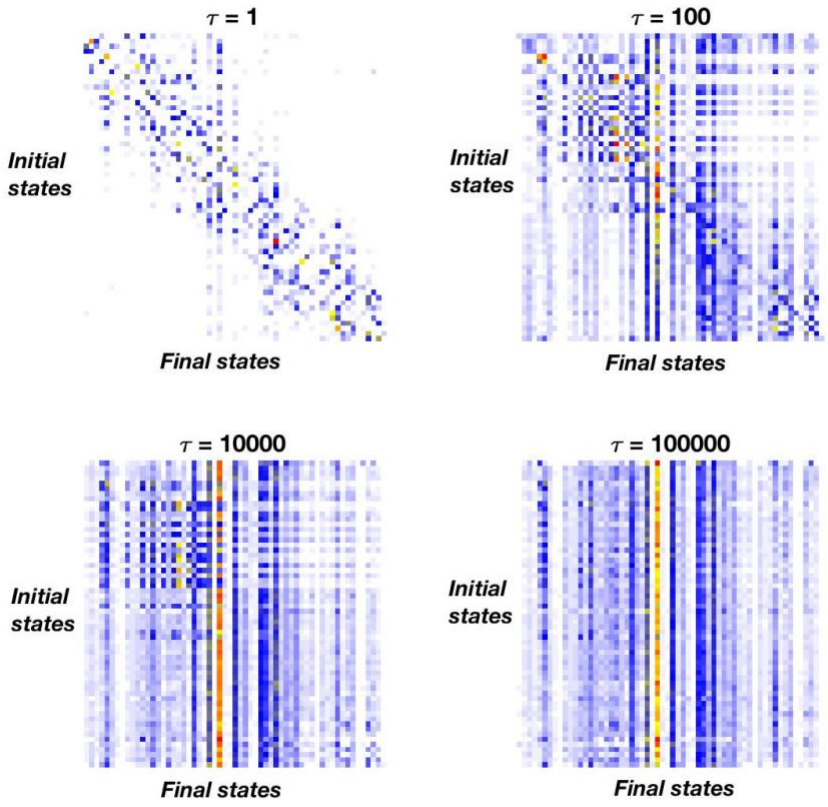


Figure 9. Transition matrices for τ = 1, 100, 10000, and 100000.

Comparison with the Markovian model

A Markovian model is a stochastic model used in randomly changing system (Gagniuc 2017). If the dynamics of the brain data are purely Markovian, then the transition of the state at any time point is not determined by any previous information but just the current state. By comparing the structures in $\mathrm{T}_M(100)$ and $\mathrm{T}(\tau)$ for τ = 100 (Figure 10), 100, 10000, and even

100000 (Figure 9), we can see that, even only by 100 steps, all rows in the Markov model are similar, and all transitions are purely random, which indicates that Markov model has already lost almost all information. This makes sense if we look at the calculation of transition matrix after τ steps. If it is a Markov model that carries no memory into transitions, the first transition matrix $\mathrm{T}(\tau = 1)$ presents a complete characterization of the dynamical system. So, the calculation of the matrix after τ transitions will be:

$$\mathrm{T}_M(\tau) \equiv [\mathrm{T}(1)]^{\tau} = \sum_{\mu} \left[\lambda_{\mu}(1)\right]^{\tau} u^{\mu}(1) v^{\mu}(1).$$

Because all eigenvalues are less than one, except for the first eigenvalue, the contribution from the $\mu > 1$ terms will drop exponentially to zero as τ goes to infinity. Therefore, for long time period, $\mathbf{T}_M(\tau)$ starts to lose all information about current state and transitions. Thus, in a Markovian model, the slowest time scale is determined by the absolute value of the second eigenvalue in the one-step transition ($|\lambda_2(1)|$), and the characteristic decaying time $\mathrm{t}_2 = \frac{-1}{\log|\lambda_2(1)|}$. Calculating the decay time for this system, we have t_2 = 15.7 transitions. In other words, there are hidden states in the system that carry the memory over longer time, beyond 2 transitions.

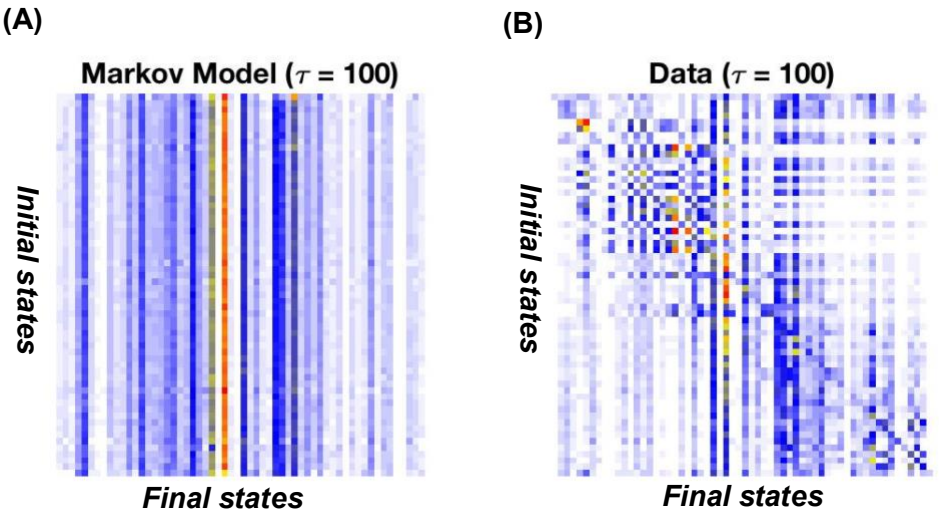


Figure 10 Long time scale transition matrices and non-Markovian dynamics. (A) Markov model transition matrix for τ =10. (B) Transition matrix of the data for τ =10.

The observations are more evident if we look at the eigenvalue spectra of the transition matrix over time. The transition matrices $\mathrm{T}(\tau)$ as a function of τ for 2-6 eigenvalues (solid line) are shown below, in addition to the prediction from Markovian model (dashed lines) based on one time-step, $\mathrm{T}(1)$. In both cases, where self-transitions are present (figure 11A) and self-transitions are excluded (figure 11B), we observe a much longer time scale in data dynamics than in the Markov model we generated from the data; without self-transitions, Markov model can only predict up to less than 100 transitions; with self-transitions, Markov model can only work up to around 1 second. The time scales in our data are way beyond what Markov model can predict.

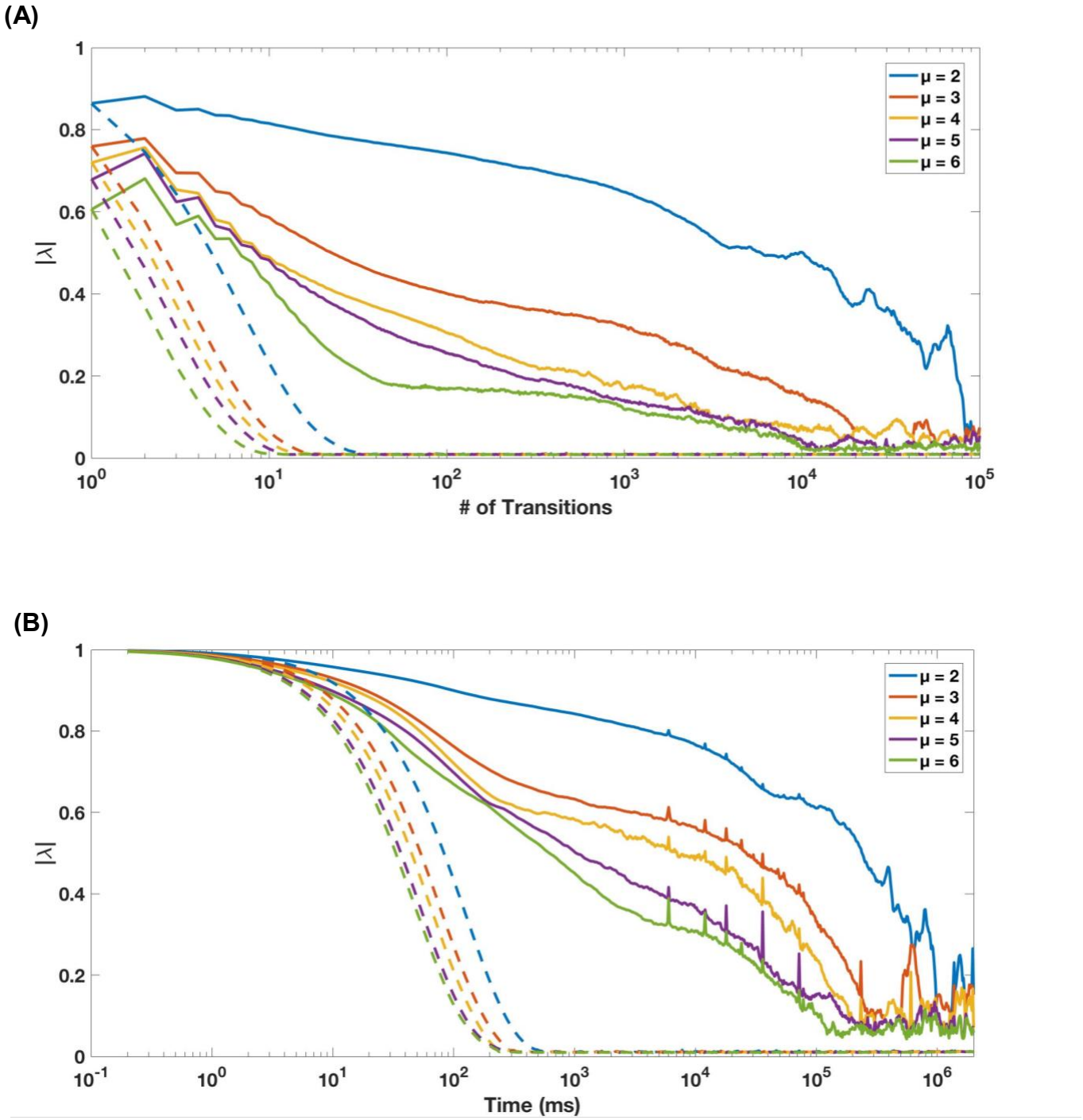


Figure 11 Eigenvalues spectrums: absolute values of the leading eigenvalues of the transition matrix **T** as a function of τ for self-transitions are excluded (A) and self-transitions are present (B).

Discussion

By looking at 24-hour ECoG recording in human brain, and consider the brain as a whole, we reduced the dimension of the dataset and looked the dynamics in its embedding space. Based on the mean wavelet of each neural channel in each region, we observed a continuous trend within the regions labeled as right hippocampus; along the trend, left hippocampus is next most active brain region. However, the structures within the embedding space are hard to describe quantitatively, we need to find a way to quantitatively describe the dynamics within the embedding space.

By looking at the movement of the patient in each region on the embedding map, we can describe the similarity and difference of human behaviors cross different watershed regions. Due to technical difficulties, we haven't completed the process of extracting video tape from each region on the map, but this will be done in the near future. However, even if we have finished the process of looking at human behaviors in each region, in order to further decode human behaviors, we should utilize methods like behavioral map (Berman et al, 2016), in which we need t0 measure the behavior repertoires of individuals and study the behavioral transitions, and long-time scale patterns. However, human behaviors are far more complicated than those of animals, making this process way more difficult.

Table of Figures

References

Hill, N. Jeremy et al. “Recording Human Electrocorticographic (ECoG) Signals for Neuroscientific Research and Real-Time Functional Cortical Mapping.” Journal of Visualized Experiments: JoVE 64 (2012): 3993. PMC. Web. 17 Oct. 2017.

van der Maaten, L.J.P.; Hinton, G.E. (Nov 2008). "Visualizing High-Dimensional Data Using t-SNE". Journal of Machine Learning Research. **9**: 2579–2605.

Berman GJ, Choi DM, Bialek W, Shaevitz JW. 2014 “Mapping the stereotyped behaviour of freely moving fruit flies”. J. R. Soc. Interface 11: 20140672

Berman GJ, Choi DM, Bialek W, Shaevitz JW. Shaevitz. 2o16 “Predictability and hierarchy in Drosophila behavior” PNAS 2016 113(42) 11943-11948

Danielle S. Bassett and Michael S. Gazzaniga. 2011 “Understanding complexity in the human brain”. Trends Cogn Sci. 2011 May; 15 (5): 200-209.

Uhlhaas PJ, et al. “The role of oscillations and synchrony in cortical networkds and their putative relevance for the pathophysiology of schizophrenia”. Schizophr Bull. 2008; 34: 927-943

Wang Nancy X. R., Olson Jared D., Ojemann Jeffrey G., Rao Rajesh P. N., Brunton Bingni. 2016 “Unsupervised Decoding of Long-Term, Naturalistic Human Neural Recordings with Automated Video and Audio Annotations”. Frontiers in Human Neuroscience: vol 10 (2016): 165.

Gagniuc, Paul A. (2017). *Markov Chains: From Theory to Implementation and Experimentation*. USA, NJ: John Wiley & Sons. pp. 1–256. ISBN 978-1-119-38755-8.

Alexander, Andrew L. et al. “Diffusion Tensor Imaging of the Brain.” *Neurotherapeutics : the journal of the American Society for Experimental NeuroTherapeutics* 4.3 (2007): 316–329. *PMC*. Web. 20 Mar. 2018.

Ildefons Magrans de Abril, et al. “Connectivity inference from neural recording data: Challenges, mathematical bases and research directions” DOI: https://doi.org/10.1016/j.neunet.2018.02.016

Meyer, Fernand. (1994). Meyer, F.: Topographic distance and watershed lines. Signal Process. 38, 113-125. Signal Processing. 38. 113-125. 10.1016/0165-1684(94)90060-4.

Sankaran, C. "Effects of harmonics on power systems"*. Electrical Construction and Maintenance Magazine. Penton Media, Inc*.

Alain Destexhe and Claude Bedard (2013) Local field potential. Scholarpedia, 8(8):10713.