

Classification of single-trial Local Field Potential spectrograms from reversal learning tasks using neural network

Jose Rey Lopez (s1033937) ; jose.rey.lopez@outlook.com

Supervisor: Mike X Cohen, Morgane Boillot

1 Abstract

Reversal learning paradigm is a test intended to test cognitive flexibility. In this work we present an end-to-end work on classification of single trial neural spectrograms for reversal learning tasks. Starting with the collection of data in a laboratory set about reversal learning on rats with electrophysiological recordings in the orbitofrontal cortex (OFC) and basolateral amygdala (BLA) regions. The data is preprocessed and analysed. First, we analysed the data by computing the Event Related Potentials (ERPs). It was possible to identify bursts of activity in the probed regions when the task of interest was being performed. Therefore confirming that the probes are in the regions of interest. Secondly, we performed a time-frequency analysis using Morlet wavelets in order to find differences in power for different experimental conditions. We found that the expectation of a reward generates a different power spectrum across different frequency bands in all the probed regions. Finally, we tried to classify the spectrograms for single trials for the three tasks studied produced by the time-frequency using both Convolutional Neural Networks (CNN) and XGBoost, but no attempt was successful. Among the hypothesized causes for the unsuccessful classification are the variability across single trials, or the difference between the training set for the CNN and the spectrograms to be classified.

2 Introduction

For an agent interacting with an environment, the optimal strategy is to maximize the rewards it obtains from those interactions. However, it is very common for an environment to change and for previously rewarded action to have different outcomes, forcing the agent to adapt its behavior in order to keep maximizing the rewards. To adapt to a new environment it's required to have cognitive flexibility.

Such cognitive flexibility can be measured in a research context by using reversal learning paradigms [1]. During an initial phase called discrimination learning, subjects are presented with at least two distinct stimuli, which might be visual, auditory or spatial locations. Then the agent is asked to perform a task, press a different button when the stimulus is presented, for instance. One of the stimulus is associated with a reward, each time it is presented and the agent performs the task correctly the subject receives a reward. The other stimulus is not associated with a reward even if the task is completed correctly.

After the subject has been performing the tasks until it reaches some proficiency criterion, the stimuli association is reversed and the subject must adapt its behavior to get rewards again.

During this reversal phase it is common to have probabilistic trials, that is, the stimuli are not always reversed. In these probabilistic trials, the mapping stimuli-reward association is the same as in discrimination. This is done to slow down the rate of learning and avoid the use of simple strategies.

The brain regions which show the highest correlation with these reversal learning tasks in previous electrophysiological studies [izquierdo] are the cortical and amygdalar regions. Specifically the Orbitofrontal cortex (OFC) and Basolateral amygdala (BLA) respectively.

For this report, the signal collected to study these regions is Local Potential Fields (LFP). LFPs are electrical fields generated by the synchronized firing of groups of neurons. When a neuron fires, it releases ions that generate an electric potential that excites surrounding neurons and makes them more likely to fire. Therefore it is very common for groups of neurons to fire together. Additionally, groups of neurons can communicate with each other, which creates an excitation that goes back and forth between both groups. This characteristic makes LFP signal somewhat rhythmic, and gives it an sinusoidal appearance. A property that can be exploited

by using frequency analysis to extract a more detailed representation of the signal. An example of the sinusoidal nature of neural information can be seen in Figure 1.

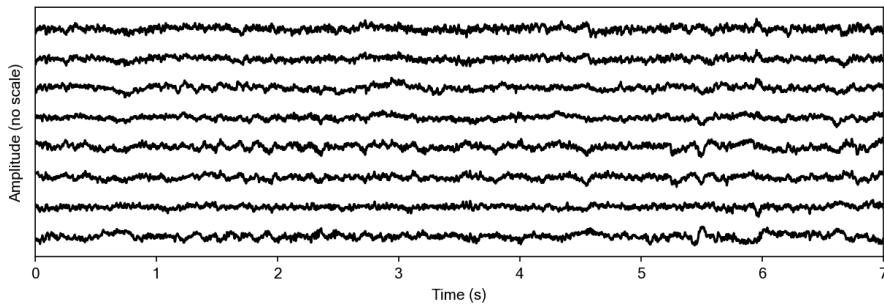


Figure 1: Example of data recorded, it presents a slightly oscillatory behavior

LFP signal properties change very fast with time, so the best practice is to take short windows of time of data from the signal and perform on it a Fourier transform, what we will call a time-frequency analysis. Different groups of neurons fire at different frequencies to communicate and do computations. The most common frequency bands found across studies are: Delta (2-4 Hz), Theta (4-8 Hz), Alpha (8-12 Hz), Beta (15-30 Hz), Low Gamma (30-80 Hz) and High Gamma (80-150 Hz).

The product of a time-frequency analysis is a spectrogram, a representation on how the power distribution of a signal changes over time for the analyzed frequencies. Since the regions of interest in reversal learning process two different stimuli, we expect the computations they have to perform to be different and to see a difference in their spectrograms. We further hypothesize machine learning algorithms to differentiate such spectrograms depending on what kind of stimuli was presented.

3 Data Collection

3.1 Materials and methods

In order to motivate the rats to perform more trials during the experiment time, they are not provided water ad libitum. They can only drink during the experiment and for a period of 15 minutes after the daily experiment has finalized. After that, they are not longer provided any water until next day's experiment.

In our experiment, rats are trained in a Skinner box that contains three holes. The central hole has a green LED that marks that it is possible to start a trial if it is lit. In order to familiarize the rats with the workings of the box, there is a pre-training phase 1. If the rat has its nose in the central nose-poke for more than 500 ms and goes to any of the side holes, the rat receives a reward in form of 50 µl of water from a pump. After the rat gets used to this mechanic, there is a pre-training phase 2. In this phase the rat needs to initialize the trial by putting the nose in the central hole for more than 500 ms, a neutral tone (won't be used in further phases) is played and a side hole is selected randomly. Only if the rat goes to that side, it will receive a reward. In this phase rats get used to the possibility of not getting a reward.

Once the rats are used to the mechanics of the box, it is possible to start with the discrimination phase. When a rat puts its nose in the central hole a tone of either 2000 Hz or 8000 Hz is played. These tones are what is called a Conditioned Stimulus (CS), since their presentation marks the possibility of receiving a reward, in which case it is a CS+ stimulus, or not receiving it, a CS- stimulus. Each tone is linked with a side. If the rat goes to that side, the trial is considered to be a correct trial and the rat will receive a reward (CS+ trials), or no reward (CS- trials). If the rat does not go to the correct side, the rat does not receive any reward under any circumstance and 5 seconds of white noise is played. Moreover, when the rat starts a new trial the same stimulus is played as many times as necessary until the response is correct.

What tone corresponds to a CS+ or CS- trial changes between rats to avoid biases, along with the side that provides the rewards. A visual schematic of how a discrimination trial is done can be seen in Figure 2.

The main metric for the performance of this phase is whether the rat is able to perform correctly a trial in the first try. The criterion for deciding if a rat is proficient at performing this discrimination task is correctly performing 60% of the CS- trials in the first try during 3 consecutive days of training. After reaching proficiency, a rat is ready to undergo surgery to have 2 probes with a total of 62 electrodes implanted in the brain. The probes cover 4 regions related with reversal learning: the Orbitofrontal cortex and the Basolateral amygdala, in both hemispheres. Location of the electrodes and an image of a probe can be seen in Figure 3.

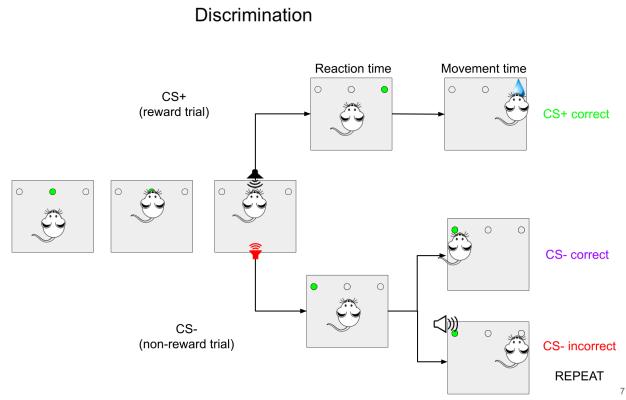


Figure 2: Schematic on how a discrimination trial can be performed for all possible responses

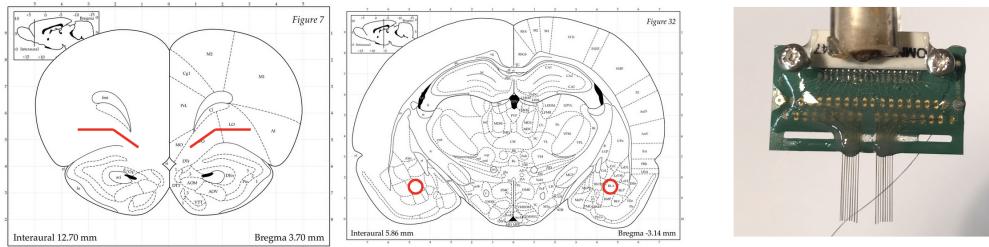


Figure 3: (1,2) Location of the implantation of the electrodes. (3) Image of one of the probes used.

After a period of around 2 weeks, the rat recovers from the surgery and is ready to return to the experiment. It continues to perform the discrimination phase for a period of around 30 days, which usually produces 1000-1500 trials of recorded data. Only discrimination phase data is used in this report. However, for a matter of completeness, the rest of the experimental procedure is described.

After finishing the discrimination phase, the reversal phase starts. In the reversal phase the tone that is linked with a reward and the tone that is linked with no reward are switched, but only in 80% of the trials. The remaining 20% of the trials still have the same tone-reward association as in the discrimination phase. They are the so-called 'probabilistic trials', their function is to slow down the learning rate of the rat. This first reversal phase lasts for around 500 trials. After that, the tones are switched again, 80% of the trials present the tones as done in discrimination phase. There are in total around 4 or 5 reversal phases, depending on the time the rat takes to finish each phase, and the quality of the recorded signal, which decreases over time. A schematic for this phase can be seen in Figure 4

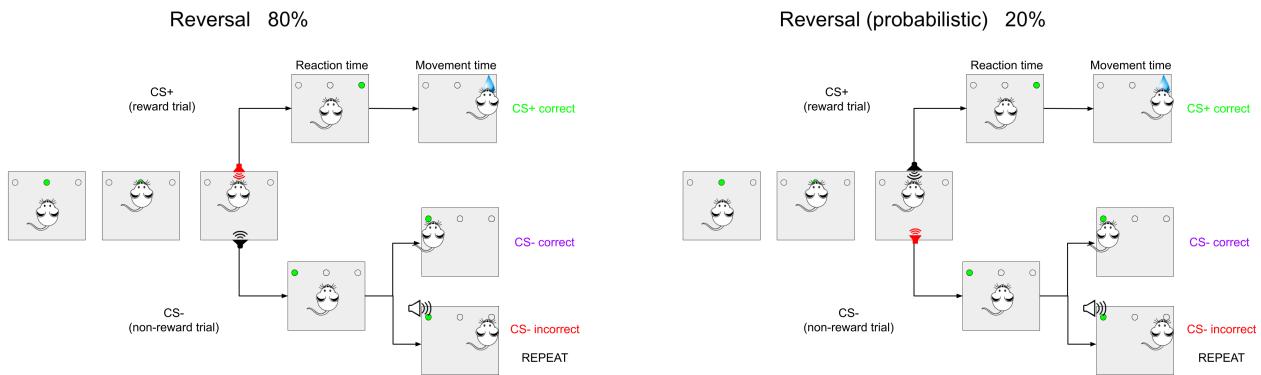


Figure 4: Schematic for the possible responses for a reversal trial, either probabilistic or not probabilistic

An example of how tones and types of trials change for all stages of the experiment can be seen in Table 1

	Discrimination (Hz)	Reversal 1 (Hz)	Reversal 2 (Hz)	Reversal 3 (Hz)
CS+	8000	2000	8000	2000
CS-	2000	8000	2000	8000
CS+*		8000	2000	8000
CS-*		2000	8000	2000
Reward Side	Left			

Table 1: Table showing how the tone changes for different phases and types of trials

3.2 Results

We first analyze the overall performance in the task for a rat in particular, and how the change from discrimination to reversal phase and between different rounds of reversal phases affects the performance, to visualize it we use Figure 5.

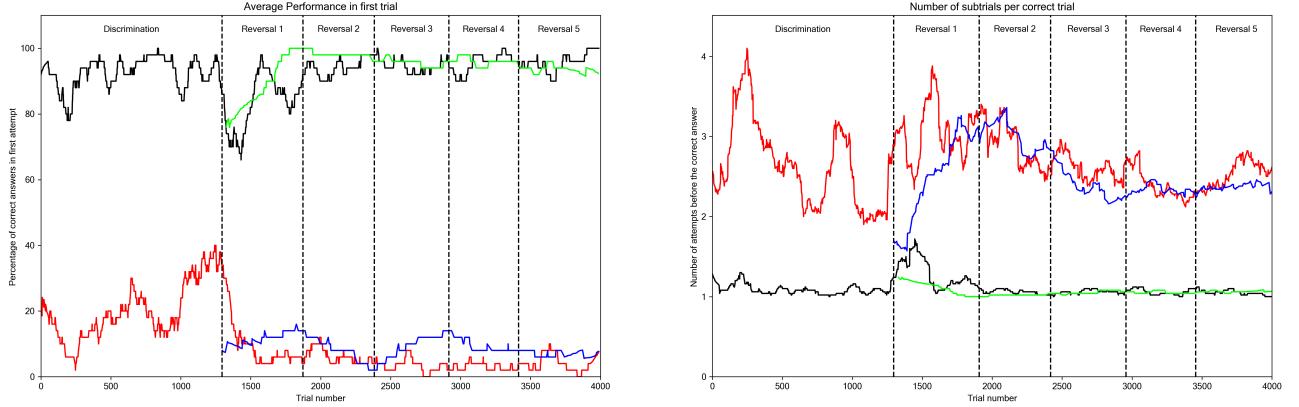


Figure 5: (Left) Percentage of correct responses in the fist attempt for different conditions and phases. The black line indicates the performance of CS+ trials, red line for CS- trials, lime for CS+* and blue for CS-*. (Right) Average number of attempts required for a correct trial for each condition, same color as in the previous figure. All values are product of a moving average of 50 elements.

In Figure 5, we show:

1. The average performance

For each of the possible trial conditions, the rat can perform it correctly in the first attempt or not. This type of data is binary and a moving average of 50 elements was used to smooth it out. The value was then scaled up to 100 to express the resulting value as a percentage of correct first trials within that window of trials.

- **CS+ trials** : Marked as a black line. By the time the recordings starts, the average performance is very high, close to 100%. This is due to the rat already knowing how to perform the task. When the reversal phase starts, the initial performance falls abruptly but soon recovers to around 80% - 90% and stays there for the rest of reversal rounds. This percentage corresponds to the trials in the reversal round that for the new rewarded tone there is actually a reward, the rat learns it and adapts to it. The 10-20% left corresponds to probabilistic trials of the reversal round that the rat can't predict so it is impossible for it to adapt and has to fail if done correctly.
- **CS- trials** : Marked as a red line. In the discrimination phase the performance has an overall positive trend reaching up to 40%. For this rat, the criterion to change from discrimination to reversal by performance was not met, and the change was due to time (30 days of discrimination). After the change to reversal, performance falls to around a 5% - 10% in the first attempt and did not recover. The rat is not putting any effort into doing CS- trials correctly at first. It always goes to the reward side.
- **'Probabilistic' CS+ trials** : By having the same tone as CS- but providing a reward, the performance for probabilistic CS+ trials sharply increases and remains very high fo all reversal round.
- **'Probabilistic' CS- trials** : Marked as a blue line. Same tone as CS+. The performance is low for all reversal rounds.

2. The number of correction trials per correction trial.

The previous metric only takes into account the result of the first attempt. In this experiment if a trial is performed incorrectly, the same stimulus is played repeatedly until the required response is performed right. Successive trials are considered correction trials of the first one. When these additional trials are taken into account:

- **CS+** : The average number of attempts required is very low, close to 1 all the time. The only remarkable feature is the small bump seen in the transition from discrimination to reversal.
- **CS-** : The average number of correction trials required is highly variable across the whole discrimination phase. We cannot pinpoint any specific event that could explain the spikes it shows. During reversals it stabilizes around 2.5, which means that in spite of always going for the reward side on the first attempt, the rat learns to perform well in the second or third attempt, even when the reversals are changed. This means that the rat learns and is able to adapt to its environment. The rat just decides not to do it at the first attempt.
- **'Probabilistic' CS+** : For the reasons presented in the performance metric, this type of trial follows the same pattern as CS+
- **'Probabilistic' CS-** : For the previous reasons, the type of trial follows the same pattern as CS-

Altogether, we observe that very soon in the training the rat identifies the tone predictive of the reward. Its strategy seems to be selecting the reward side in the first try most of the time, regardless of the type of trial. Such strategy gets reinforced with the change from discrimination into reversal phase, when the uncertainty about the functioning of the box is at its peak.

4 LFP Analysis

4.1 ERPs

4.1.1 Materials and methods

Before applying any analysis on the data, the signal underwent several steps of preprocessing, including cleaning and local averaging (the sum of all signals from one of the 4 regions is zero for a trial).

ERPs (Event Related Potentials) is the most basic analysis that can be performed on LFP data. ERPs are bursts of activity seen when the LFP data centered around a specific task-related event from many trials of the same task are averaged. Neural data is composed by signal and noise. If the data from many trials of the same task is averaged, there will still be signal left. Noise on the other hand follows a symmetric probabilistic distribution and when averaged it tends to zero.

ERPs is the most basic analysis that can be performed on LFP data. Initially the signal needs some preprocessing, it has to be cleaned and locally averaged (the sum of all signals from one of the 4 regions will be zero for a trial).

There are three task-related events in our experiment that can trigger ERPs:

1. Stimulus presentation, after 0.5 seconds of nose poking in the central hole, a tone is played.
2. Reaction, the rat releases its nose from the central hole
3. Response, the rat answers the trial by poking one of the side holes

A method to visualize the possible existence of ERPs is plotting all the trial voltages as an image. To create an ERP image (Figure 6), we plotted the voltage for each trial (y axis) over a period of time of 7 seconds (x axis), 2 s before stimulus and 5 s after stimulus. The vertical black line marks the stimulus presentation. The curved

line marks the reaction time (nose removed from the central hole), which is also used to sort the order of the trials presented. The scattered points mark the movement time (nose introduced in one of the side holes).

4.1.2 Results

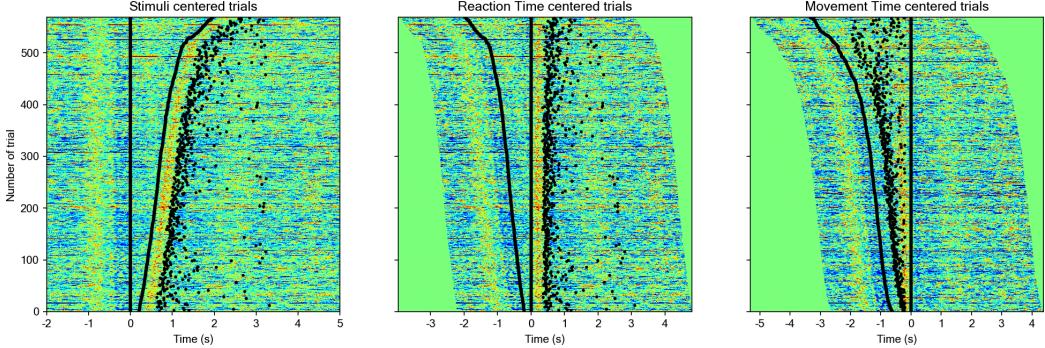


Figure 6: ERP images for different centerings, the yellow section indicates a burst of activity.

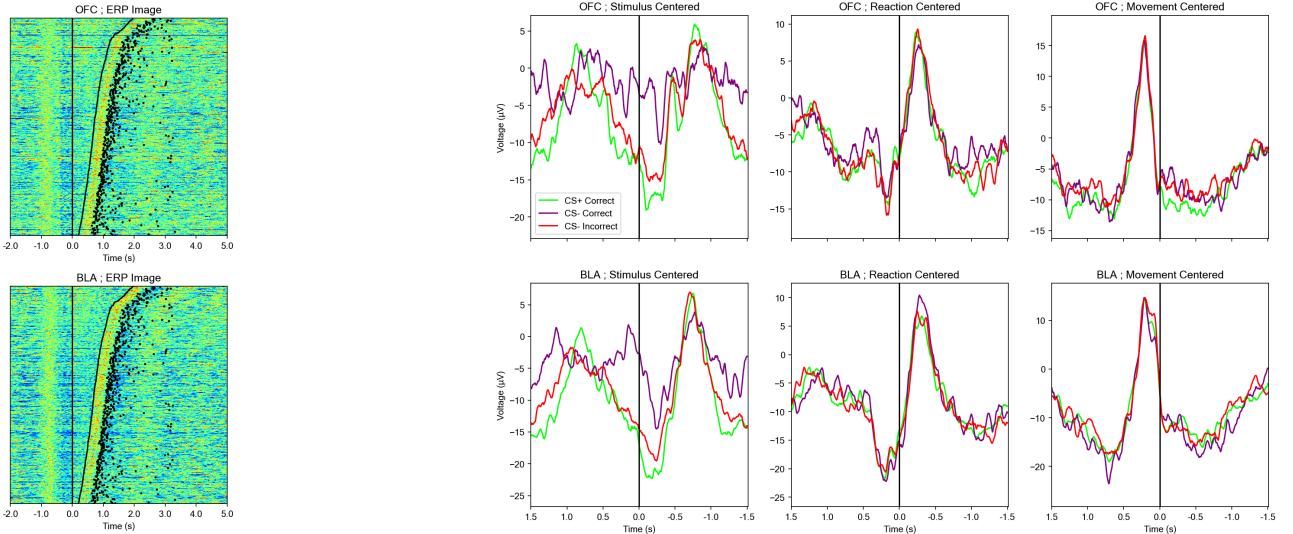


Figure 7: For two channels, one in the OFC and one in the BLA, the subplots on the left show the ERP images. The subplots on the right show the ERPs when the data is centered in each of the three events

The ERP images from Figure 6 show that around -1 to -0.5 seconds before the stimulus there is a burst of activity. Another burst of activity happens right after reaction time. This is a positive result because it means the brain activity in the regions (upper panel: OFC, lower panel: BLA) being probed is related with the task the rat is performing. Apart from this, it is not possible draw further conclusions.

In the right subplot in figure , we plotted the averaged voltage of one electrode in the OFC and one in the BLA centered around our 3 centerings for each type of trial. For stimulus centered ERPs, the CS- correct trials have a different value than CS+ correct and CS- incorrect trials. This centering

There is a peak after 1 second, but the height of the peak is not as big as for the other two centerings. This is due to the peak of activity being correlated with the reaction time and it's spread along a wide time window after the stimulus.

For the centerings for reaction time and movement time there's a very strong peak because all bursts of activity are very close with respect to the events, in case of the movement time the peak is already behind the centering, additionally, the difference between conditions that could be seen in the first centering is lost. The most

reasonable explanation is that the rat already heard the tone, processed it and took a decision. By the time it takes its nose out the central hole there's no difference between conditions because the decision making process has already finished.

4.2 Time-Frequency analysis

4.2.1 Materials and methods

LFP signal presents a rhythmic activity which arises from oscillations. The best tool to study the composition of an oscillatory signal is a frequency analysis, which can be done with a Fourier Transform for instance. Fourier transforms decomposes the electrical signal in time into the amplitudes of sinusoidal waves of different frequencies. However, Fourier transform only provides good results for steady state signals, signals whose characteristics, amplitude or frequency for instance, do not vary in time too much. This is not the case for LFP data, so we must perform a time-frequency analysis.

Time-frequency analysis is a technique that allows for the study of the frequency spectrum of non-transient signals (frequencies and/or amplitude of the signal varies in time). It is done by taking small sections of the signal which correspond to periods of time where the signal does not vary greatly and perform a Fourier-like analysis on it. This analysis doesn't have a temporal precision as high as ERPs do, but in turn is much more informative.

The main tool we use to perform time-frequency analysis is the Morlet wavelet. A Morlet wavelet is combination of a complex sinusoid and a gaussian signal. The complex sinusoid accounts for the frequency and the phase of the signal and the gaussian centres the frequency analysis in a window of time within the whole signal.

The parameters that can be tuned for a Morlet Wavelet are the frequency and the number of cycles of the that are not attenuated by the gaussian.

$$\text{Complex Morlet Wavelet} = \underbrace{A}_{\text{Factor}} \cdot \underbrace{e^{-\frac{t^2}{2s^2}}}_{\text{Gaussian}} \cdot \underbrace{e^{i2\pi f t}}_{\text{Sine}} \quad \text{Where } A = \frac{1}{(s\sqrt{\pi})^{\frac{1}{2}}}$$

To illustrate how these two parameters influence the outcome of the analysis, we now use an example comprised from 697 trials which contains an electrical leak at 100 Hz. Used to show how the parameters affect the spread of the frequency power but not used in further analysis due to the leak.

The frequency of the sinusoid tunes which frequency of the signal will be studied. Close frequencies will overlap to some extent. This effect can be very easily visualized in the upper plots of Figure 8. The part of the signal highlighted in green has a big and slow oscillation. When convolved with a Morlet wavelet of a low frequency (6 Hz), the spectrogram in the upper right part shows a peak of activity at the time of the center of the signal at that frequency. The part of the signal highlighted in blue shows some very fast oscillations due to electrical line leaks (100 Hz), when convolved with a Morlet wavelet at that frequency the spectrogram again presents a peak of activity, but this leak does not filter far away from its central frequency of the leak.

The number of cycles controls how long the analyzed time window is. For lower frequencies, each cycle is longer than for cycles of higher frequencies. This means that to avoid taking too long time windows, we use a lower number of cycles for lower frequencies. Furthermore, in longer time windows the condition of steadiness of the signal's parameters is not fulfilled. Higher frequencies, on the contrary, require a higher amount of cycles so the time window is not too small and the activity of interest is properly captured.

The effect of the number of cycles can be seen in the lower plot of Figure 8. On the left plot, with a constant number of cycles of 6, it produces good results (balance between frequency and temporal precision) in the lower frequencies but it will distort the 100Hz electrical noise used to illustrate this effect. The signal is distorted into a wider range of frequencies centered in the original value. The contrary happens on the right plot, a very high value for number of cycles will not provide good results in low frequencies, the results have high frequency precision but bad temporal precision. The higher frequencies' data is not distorted.

Using Morlet wavelets, a more in-depth analysis can be performed by looking at the activity by frequency of each one of the channels and comparing the power for each region for each frequency band.

We use a frequency range of 2 to 120 Hz, with 20 frequencies logarithmically spaced in the 2-25 Hz range, and 30 linearly spaced frequencies in the 25-120 Hz range. This combination allows for a good trade-off to visualize

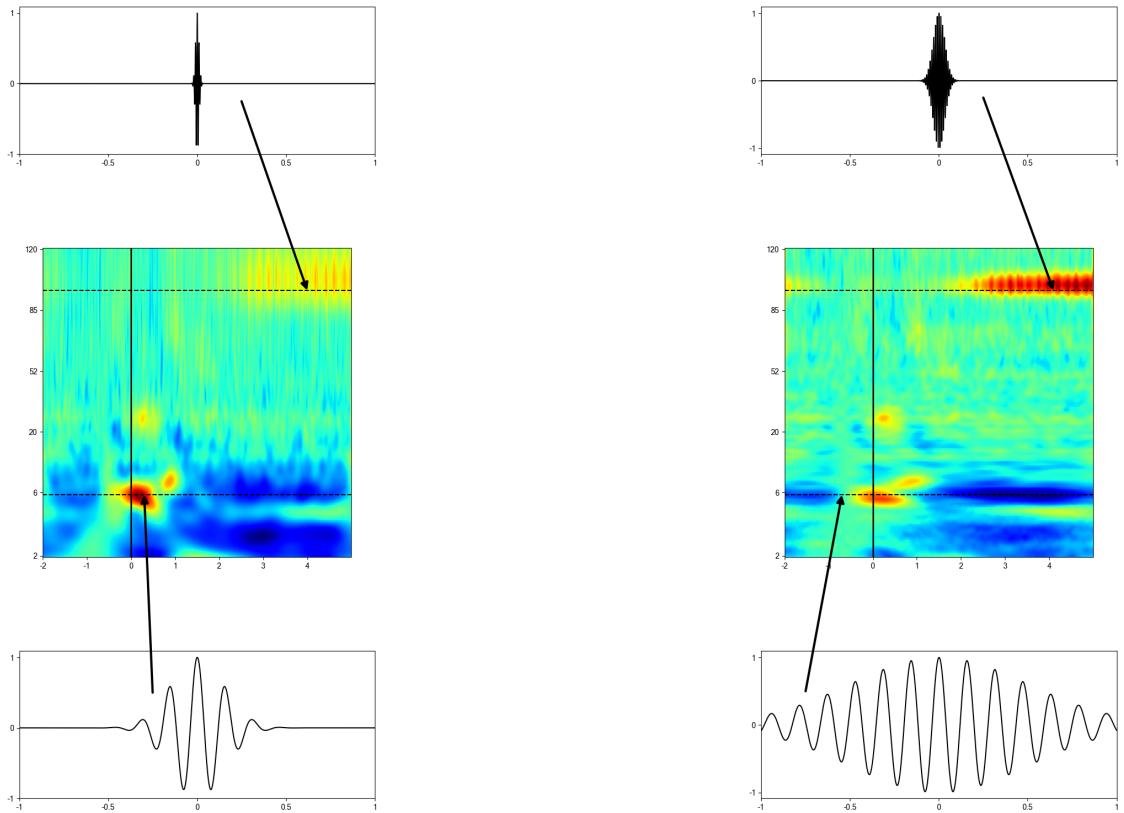
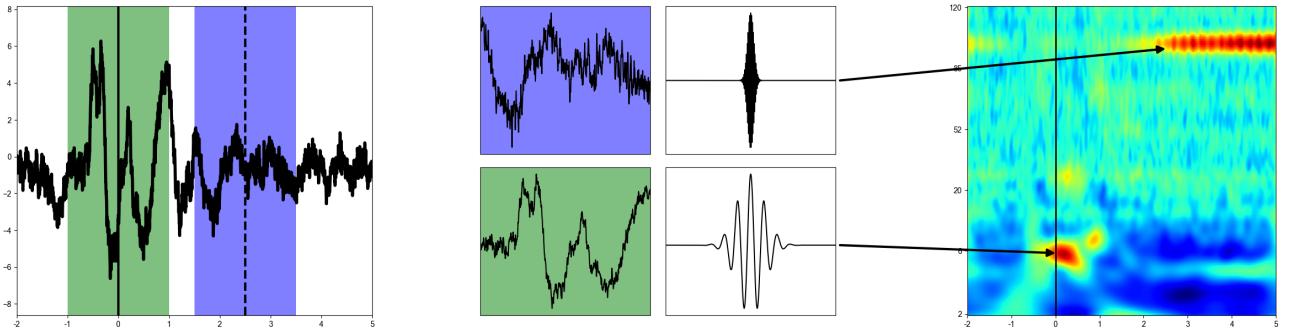


Figure 8: (Upper plots) Signal to spectrogram schematic pipeline; Effect of different values on the frequency of the sinusoid. (Lower plots) Effect on the spectrogram of different values for the number of cycles.

changes in power in the entire frequency range with a limited number of frequencies sampled, allowing for a faster execution.

The number of cycles selected changed from 6 to 20 in a logarithmic way as the frequencies increased. The selection of such parameters was done by visual inspection of time-frequency plots generated with different parameter values. We select parameter values that did not dissolve the strong theta component the signal has while also preserving gamma activity.

After performing the time-frequency analysis we applied baseline normalization using the data from the time window -0.8 to -0.5 seconds before the stimulus. The time window for the baseline was selected for not being contaminated by any activity from the trial or other trials' activity. We then transformed it into decibels.

The frequency spectrum of the electrophysiological activity follows a 1/f pattern. Lower frequencies have higher energies but by applying baseline normalization we both eliminate the 1/f pattern, that makes difficult to appreciate changes in activity. It also eases viewing which frequencies have higher or lower task-related activations when compared to the brain not doing any non task related activity.

4.2.2 Results

The result the time-frequency analysis of power is knowing the relation of power that the signal has at each frequency range at any point in time compared to the baseline for all electrodes, as shown in Figure 9. With this result it is possible to study the differences in power between types of trials. In order to do so, we select, via visual inspection, different time and frequency windows according to the peaks of activity. We then averaged the power values for each of the windows for all electrodes in one of the four brain regions for each condition. Additionally the three centerings (stimulus, reaction time and movement time) are compared, as the task-related activity can vary with different time locking.

The selection of the frequency bands of interest was done by visual inspection. The observed frequency bands are within the range of frequencies known to be linked to brain activity in previous works so the activity we see is in accordance to well-established neuroscience literature.

There are automated methods that allow for frequency band separation [gedBounds]. However, the selected bands were very similar to bands presented in previous literature so the activity we see is in accordance to what well-established neuroscience literature suggests it should look like.

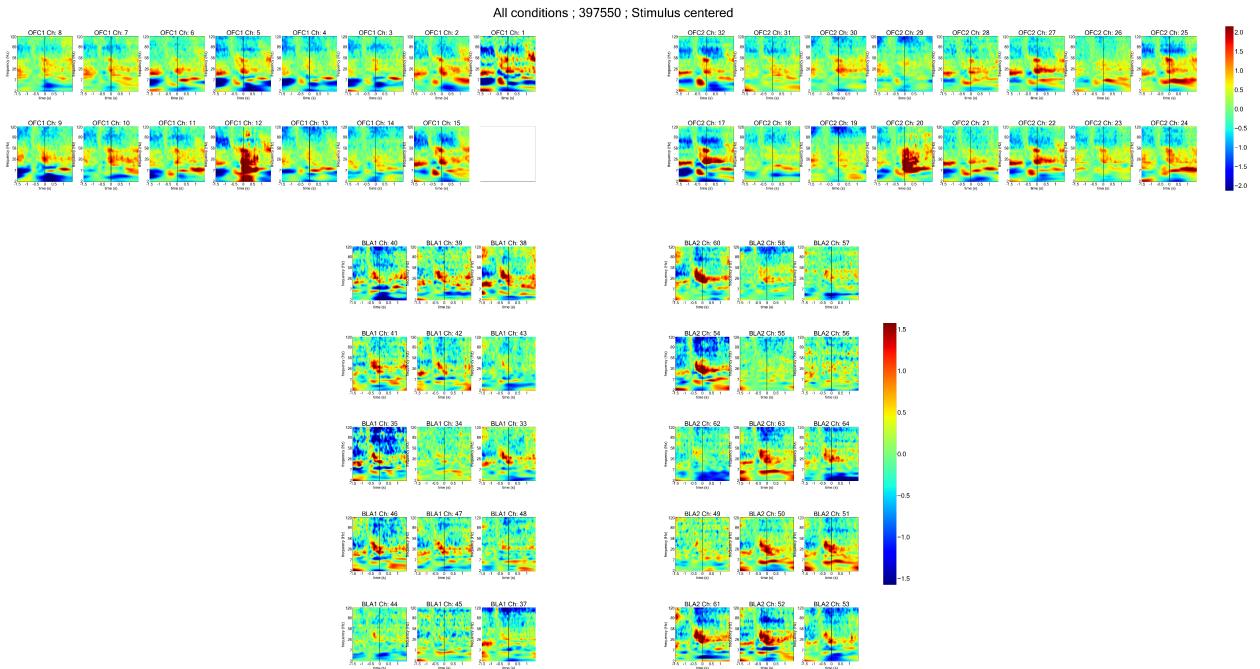


Figure 9: Time frequency plots for the 63 electrodes for all conditions with the data centered in the stimulus

The results of the previously described process can be seen for 2 rats in Figures 10 and 11.

In Figure 10 it is very notable that in both OFCs for all centerings the low gamma range (60-90 Hz) power in CS- correct trials is consistently decreased from the two other conditions. Other frequency bands such as theta (8-12 Hz) also present the same difference, with CS- C power being the lowest for many of the regions and centerings. Finally in both BLAs, the beta band (15-30 Hz) presents the difference for all centerings.

The results for the second rat are presented in Figure 11. We observe some power differences across conditions but not as notable as for the previous subject. For instance, beta (15-30 Hz) and low gamma (60-90 Hz) in both OFCs and beta in both BLAs for stimulus centered data present a decreased CS- C power compared to the other two conditions. The other centerings do not seem to evoke much task-related activity in such a systematic way.

The results suggest that the rats, specifically the rat from Figure 10, have a reward expectation. The power for trials where they are provided water (CS+ C) and those where it is not provided but the rat still goes for the reward side (CS- I) have a different power distribution than trials where the rat deliberately goes for the side that does provide a reward to complete a trial (CS- C).

As preliminary results it's possible to say that the expectation of a reward generated different power patterns across regions, but further results are required to confirm since as now only two animals' data has been analysed.

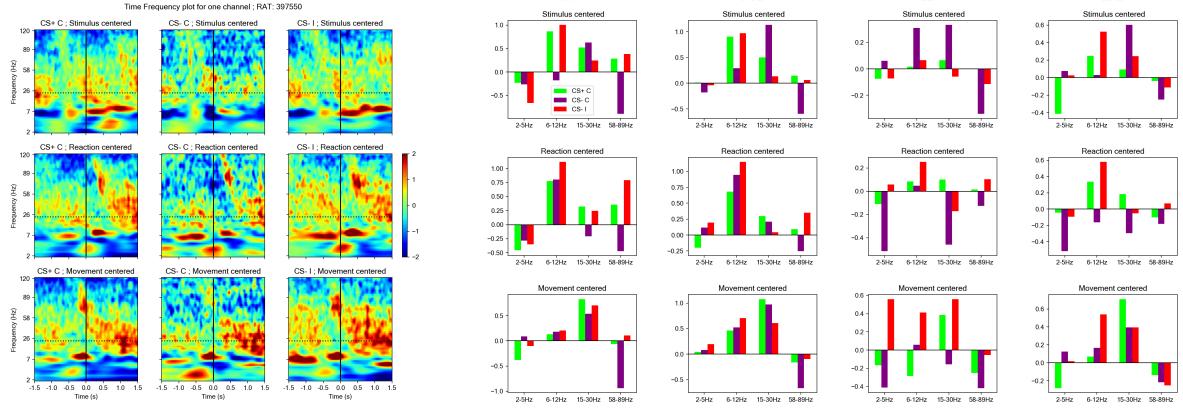


Figure 10: (Left) Time-frequency plots for the three centerings and three conditions. (Right) Averaged power for 4 different frequency bands, all regions, all conditions, all centerings

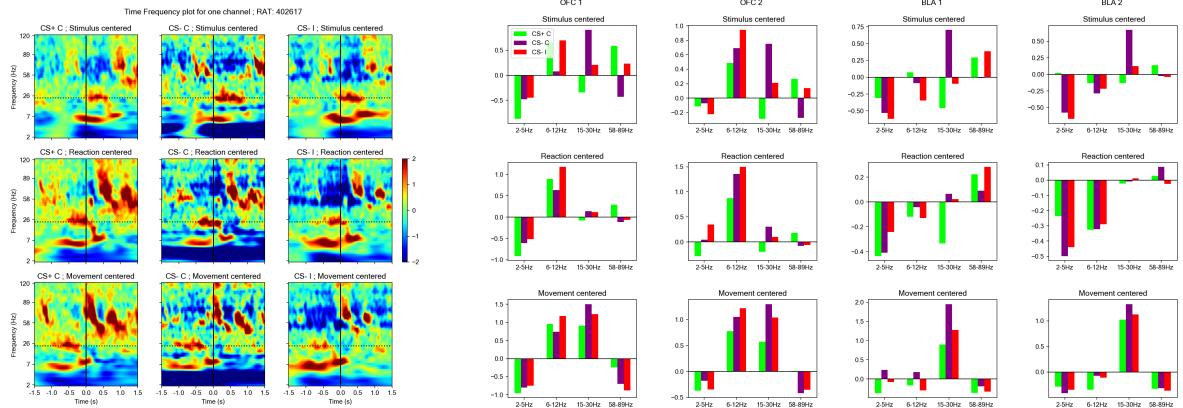


Figure 11: (Left) Time-frequency plots for the three centerings and three conditions. (Right) Averaged power for 4 different frequency bands, all regions, all conditions, all centerings

5 Machine Learning

5.1 Introduction

Claiming that two experimental conditions have a different power distribution at a given frequency band for task requires the use of all available information to reduce the noise and variance present in the LFP data. On the contrary, to the naked eye, individual trials don't usually have features as salient as the aggregate of all trials do, and differentiating them is almost an impossible task.

Machine learning algorithms have the ability to process high volumes of data and find correlations in that data in a way that might not be possible for humans to do. This raises the question about whether a machine learning model could correctly classify spectrograms from deep regions of the brain that apparently don't have a pattern with the stimulus presented.

In order to test this, two approaches were used. Treating spectrograms as images and processing them with a convolutional neural network and feeding the spectrogram itself to a fully connected neural network.

5.2 Convolutional neural network

5.2.1 Materials and methods

A Convolutional Neural Network (CNN) is a deep learning model that is feed and image and by a set of consecutive convolutions of the original image with different groups of filters. It can provide a low dimension vector that compresses the information contained within the image. That vector becomes part of the training set for a classification algorithm that can be used to classify the content of the original image. That is the most basic example of a modern CNN. The most well-known examples of such CNNs is Alexnet [1]. A model that back in 2012 was able to win the ImageNet classification problem for the first time for a model that did not have handcrafted feature extraction protocols. Alexnet also uses a combination of filters and a classifier that operates with a low dimensional vector. It is a common practice to reuse already trained CNN for purposes that they were not intended to be used by extracting the low dimensional vectors from the images one want to classify and only train the a new classification model. This approach will also be followed here.

The initial 5 convolutional layers from the Alexnet model provided by the *pytorch* library were used to generate the low dimensional vectors from spectrograms of single trials obtained in the previous section.

The dataset consists of 560 spectrograms for 3 classes. Their distributions is [CS+ Correct: 257, CS- Correct: 109, CS+ Incorrect: 194] each one with a size of 7000x50 pixels. The target is a 3x560 one hot encoding matrix for the class to which each spectrogram belongs.

The initial 2000x50 pixels of the spectrogram, corresponding to the frequency analysis for the 2 seconds previous to the stimulus presentation, were discarded. The 5000x50 pixel spectrogram left was resized to a 224x224 pixel image using python's *cv2* library method *resize* with default parameters. Then they were made into a 3x224x224 image, where each channel had the same information. This image is not normalized to the values recommended by the Alexnet developers because the normalization made the loss values grow up until a NaN result was provided. Finally each spectrogram is passed through Alexnet.

The result is a 1x9216 vector for each one of the original spectrograms. The resulting vectors were splitted in train a test datasets with proportion of 66%/33%.

Two classifiers were trained. The first one was a fully connected neural network with intermediate sizes [9216, 1000, 512, 3] and ReLu layers between each Linear layer. The training was performed using Stochastic Gradient Descent with a learning rate of 1e-3 and a momentum of 0.9 for 150 epochs with a batch size of 10, using Cross Entropy Loss as loss function.

The second classifier was a XGBoost model with a depth of 10, learning rate of 0.1 and 100 epochs training.

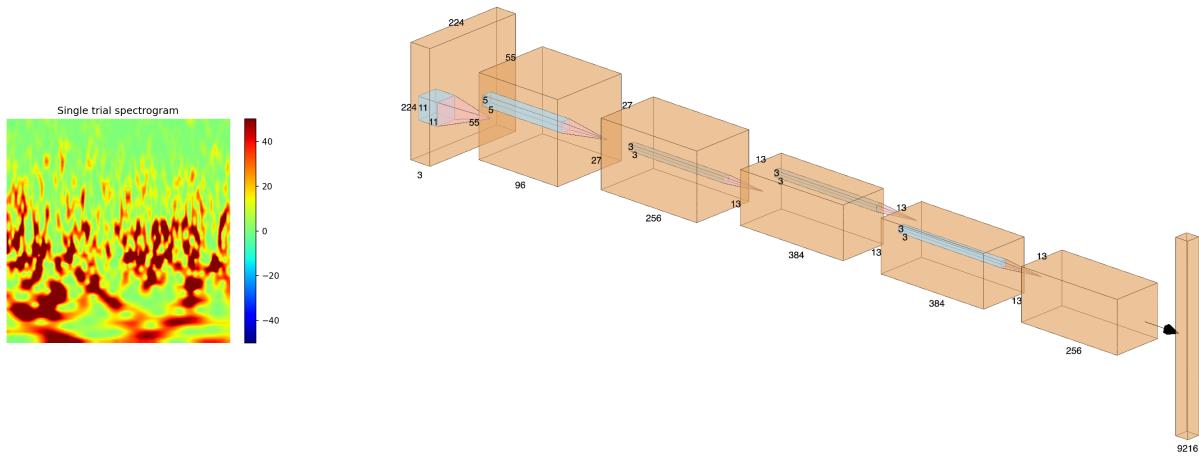


Figure 12: (Left) Example of one of the spectrograms passed as images to the neural network (Right) Alexnet network architecture used to get the feature vector used for classification.

5.3 Fully connected neural network

5.3.1 Results

For the fully connected neural network, the loss decreased very rapidly to 1 in the first 10 epochs, but then it stayed there for the rest of the training. When the test set of vectors was passed by the network the classification

values for each category were the same for all vectors, which means that the network would classify all vectors with the same label.

To test whether it is a programming error or the result was correct, Gaussian noise images with 0 mean and standard deviation of 1 and size 3x224x224 (all channels were the same image) were passed through the network. The classification values varied in the first decimal, but that only was a 10% of the difference between each of the classification values, which where separated by approximately 1 unit.

The XGboost model results were very skewed and also provided the same label for all of the vectors in the set.

Further tests were performed. A SVC model created with the *sklearn* library was trained. For the test set only 7 of the 185 test vectors were classified in a differently from the other 178.

After much testing and debugging the results seemed to be conclusive, there were no errors. Such results could be a product of 2 factors. The first one is that Alexnet was trained with naturalistic images and using it for something so different as an spectrogram will not produce an acceptable result. It might be that the filters destroy the blobs of activity present in the spectrogram and the vector representation don't have any relevant information. The second factor could be that, due to the high variability in the power for each spectrogram, it really is impossible to do single trial classification.

5.3.2 Materials and methods

Using directly a fully connected network as a classifier for the spectrogram approach exploited the fact that frequency analysis is by itself a very good feature extraction method. For instance, many of the filters in a trained CNN are nothing but Gabor filters, which are very similar to what a 2D Morlet wavelet would look like. That is, an spectrogram is in itself a low dimensional vector that summarizes the information contained in the raw electrical signal.

To test whether this approach would work or not, the spectrogram were processed in a similar was as in the previous section. The initial 2000x50 pixels were removed. The 5000x50 pixel spectrogram left was resized using *cv2* library function *resize* with default parameters down to a 50x50 pixel spectrogram. After that, it was flattened into a 1x2500 vector. This process was performed for the 560 spectrograms. The vectors were splitted into a train and test datasets with a proportion of 66%/33%. The target was a 3x560 one hot encoding matrix containing to which class each spectrogram belongs. The classifier was a fully connected neural network with intermediate sizes [2500, 1000, 512, 3] with ReLu layers between each linear layer. For the training, Stochastic Gradient Descend was used with a learning rate of 0.01 and momentum of 0.9. Cross Entropy Loss was used as loss function. The training consisted of 150 epochs with a batch size of 10.

5.3.3 Results

In this case the outputted results by the classifier network were not the same for all the test vectors, but after using softmax to pick which class had the highest value all the test vectors were classified with the same label anyway.

The similarity of results between this section and the previous one cements the theory that the high variability of the signal for single trial spectrograms is responsible for the them being impossible to classify. It was not just a fluke of Alexnet during the processing of the images, not only, at least.

These results suggest that single trial classification of neural activity during reversal learning is not possible. A result that, most likely, holds for the signals of other high cognitive functions in the deep brain.

6 Discussion

This work shows that it is highly unlikely that electrophysiological data from learning activities presented as single trials, recorded from the cortex and amygdala sites can be classified using machine learning algorithms.

Reversal learning paradigms are tests used to measure cognitive flexibility. The test in our experiment consists of having a rat in a skinner box with three holes. The central one will have a green LED, when lit it's possible to start a new trial. At the start of the trial a tone will be played, either of 2000 Hz or 8000 Hz. Each tone is a Conditioned Stimulus, that is, when heard the rat should perform an action to successfully finish the trial. If the Conditioned Stimulus is associated with a reward (CS+) the rat will need to go to a predetermined side within a short period of time to get a reward, in this case a few microliters of water. If the tone associated with no reward (CS-) is played, the rat needs to go to the opposite side. It won't receive any reward, but will be allowed to start a new trial which might provide one. In case the rat does not go to the side associated with the tone of the stimulus or does not go to any side at all white noise will be played and the same trial will be repeated until done correctly. This is what is called the discrimination phase. In subsequent phases of the experiment phases, called reversal phases, the correct side for each tone is changed 80% of the time. The other 20% are called reversal trials and the side-tone relation is changed back to the previous phase, they are called probabilistic trials and are done to not make the phase change so easy and make it slightly more difficult for the rat to adapt to the change.

The behavior of the rat allows us to understand how well it performed during the experiment and how fast it was able to learn and adapt to the reversal phases. In the case of the main rat studied here, after it was trained and started with the discrimination phase the performance for the rewarded trials was very high. Reaching performance of around 90% of accuracy in the first try. Meanwhile the accuracy for the non-rewarded trials initially dipped from a 20% to almost a 0% in the first few days of training but later recovered reaching close to 40% by the time the phase is about to end. Reversal phases got an accuracy of around 10% accuracy for non rewarded trials and 90% of accuracy for rewarded ones. Rewarded trials were done correctly in the first try across all phases, non rewarded trials needed between 2 and 3 tries to be done correctly. The rat deemed it worth trying to go to the side providing water in the first attempt, after seeing it had none, it decided to do it correctly.

Electrophysiological data was recorded from the OFC and BLA, regions associated with reversal learning in previous studies. The recordings were performed on those two regions for both hemispheres, adding up to a total of 4 regions with 16 electrodes on each region. Two in-house built probes were used for each rat.

Event Related Potentials was the first analysis performed with the electrophysiological data. The presence of a burst of activity right after one of the events of the experiment across most of the probed sites indicates that the regions probed are the ones of interest and that they actually partake in the reversal learning task. ERPs can't really provide more information as they lack frequency specificity.

In order to extract spectral information from the data, Morlet wavelets were used. Morlet wavelets are a combination of a sinusoid at the frequency of interest and a gaussian that vanishes the sinusoid as it gets further from the central point. This allows the power of a frequency to be extracted, with the vanishing tails due to the gaussians the power extracted is also confined to a time window of interest. A noticeable result of the time-frequency analysis for each electrode is that there's symmetry in the spectrograms with respect to the hemispheres. This suggests that the surgery was very precise and the recorded data will be high quality when it comes to comparing power between regions. Possible future discrepancies in that regard can be explained due to brain function and structure but not experimental error. When looking into detail the spectrograms for each electrode it's possible to see that the activity is present in a range of frequencies, groups of neurons don't fire at a single frequency but in a range. If one were to select frequency bands present in our data, the most common bands present in literature would perfectly fit within that selection. This also points towards the high quality of the data.

Comparing the spectrograms for each task we see that some frequency ranges present changes in power within the 0.5 to 1 second window after the stimulus presentation. After selecting the frequency ranges and time intervals of interest, all the power in the regions of interest were averaged for all electrodes in each region. The goal of the averaging is to avoid cherry picking outlier electrodes and having a more global understanding of the region's activity. The power for the correct CS+ trials, rewarded, and incorrect CS - trials, the rat went to the hole that provides a reward in rewarded trials, was very similar. The similarity was present in both OFC and BLA, and for all centerings. The centerings analysis was performed to focus the power of activities that were spread in time in relation to the stimuli into a single time point.

These results suggest that trials associated with a reward, either provided or wished but not provided, have a different power distribution. When the rat deliberately decides to not go for the reward side a different

process has to be taking place, one where the BLA generally has a reduced activity across all frequencies and hemispheres and where the activity in the gamma band of the OFC decreases.

Finally, it was attempted to classify spectrograms from single trials. Two different approaches were used, treating the spectrograms as images and using a CNN to create a feature vector used to classify the trial or using the data from the spectrogram as a feature vector itself. Following the CNN path, the spectrogram is reduced to the size accepted by a pretrained Alexnet network from the pytorch library and then passed through the network. The result is a feature vector that in Alexnet would be passed by a 3-layer fully connected network. In this work, the feature vector is collected and then passed for classification to a fully connected network and to a boosting algorithm. None of the analysis was successful.

CNNs are fantastic at extracting the most relevant features from an image and summarizing them in a single vector of floating point values. They do this by applying filters to small patches, adding the values and creating a new image, a smaller one for each Alexnet layer. The initial layers have Gabor-like filters, a 2-dimensional filter with some similarities with the aforementioned Morlet wavelets. Later layers are tuned to fire with more intensity when objects present in the test set are in the processed image. Since the test set for the training of Alexnet is formed by natural images which look nothing like the spectrograms, this analysis is definitely suboptimal for our interests.

Another approach is directly taking the spectrogram as a vector and using it as the feature vector to be classified. The theoretical explanation of why this approach might yield positive results is based on the time-frequency analysis being a good method to extract useful information from a signal. It's easier to understand a complex signal as a combination of simple signals than it is to understand it as a raw signal. Furthermore, the Gabor-like layers from Alexnet are nothing but a frequency analysis step, so doing a second frequency analysis on top of another whose result will not necessarily be interpretable or it may even degrade the initial analysis altogether. To perform this classification task, a fully connected network was trained on the flattened data, several layer sizes and hyperparameters were tested to see which values yielded the best result. As in the CNN analysis, the results were not very promising. All the samples were labeled the same. Once again the classification attempt failed, in this case not due to an inadequate architecture so we have to interpret the results somehow. The most possible explanation is that the variability of the signal is too high for the spectrograms to have a consistent power distribution for the different tasks. Making it almost impossible or impossible at all to be classified on a single trial basis. To put this data into perspective, the time-frequency analysis that revealed a difference in power between experimental conditions required around 250 trials.

This work is embedded within a larger project studying the effects of serotonin in reversal learning. Two groups of rats were used, a control group and another group of genetically modified rats with a serotonin transporter knock-out gene. The difference between groups will reveal the importance of serotonin in this process. Here, only the initial steps of the project are presented, behavior analysis, ERPs and the time-frequency analysis. However, only for the initial phase of the reversal learning is analyzed and only for two rats of a single group. Therefore this work must be taken as a proof of concept for the analysis pipeline and not as a finished project. Additionally, even though the results are interesting, they lack the statistical power necessary to be considered valid. To overcome this shortcoming, the data from more animals must be recorded, analyzed and compared to be sure it's not a fluke in these animals' data. The classification part originated as a ramification from the main project to study how possible it is to use machine learning to classify data from structure not related to somatosensory regions, which are the main targets of such experiments. It turned out to be unsuccessful, deep brain regions partaking into decision making and similar cognitive processes have their signals more smeared out through several regions and it's encoding is more difficult to understand that it is from simpler brain processes.

7 References

They have different formats? what should i do. everything in APA?

[izquierdo] Izquierdo A, Brigman JL, Radke AK, Rudebeck PH, Holmes A. The neural basis of reversal learning: An updated perspective. *Neuroscience*. 2017 Mar 14;345:12-26. doi: 10.1016/j.neuroscience.2016.03.021. Epub 2016 Mar 12. PMID: 26979052; PMCID: PMC5018909.

[electrodes] França, A. S. C., van Hulten, J. A., Cohen, M. X. (2020). Low-cost and versatile electrodes for extracellular chronic recordings in rodents. *Heliyon*, 6(9), e04867. doi:10.1016/j.heliyon.2020.e04867

[LFP] Herreras, O. (2016). Local Field Potentials: Myths and Misunderstandings. *Frontiers in Neural Circuits*, 10. doi:10.3389/fncir.2016.00101

[Mike's book] Cohen, M. X. (01 2014). Analyzing Neural Time Series Data: Theory and Practice. doi:10.7551/mitpress/9609.001.0001

[ged] Cohen, M. X. (2021). A data-driven method to identify frequency boundaries in multichannel electrophysiology data. *Journal of Neuroscience Methods*, 347, 108949. doi:10.1016/j.jneumeth.2020.108949

[Alexnet] Krizhevsky, A. (2014). One weird trick for parallelizing convolutional neural networks. *CoRR*, abs/1404.5997. <http://arxiv.org/abs/1404.5997>

[pytorch] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., ... Chintala, S. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. In *Advances in Neural Information Processing Systems* 32 (pp. 8024–8035). Curran Associates, Inc. Retrieved from <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>

[cv2] Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*.

[sklearn] Pedregosa, F., Varoquaux, Ga”el, Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... others. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12(Oct), 2825–2830.

[previous reversals] van Wingerden, M., Vinck, M., Lankelma, J. V., Pennartz, C. M. (2010). Learning-associated gamma-band phase-locking of action-outcome selective neurons in orbitofrontal cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 30(30), 10025–10038. <https://doi.org/10.1523/JNEUROSCI.10.2010>