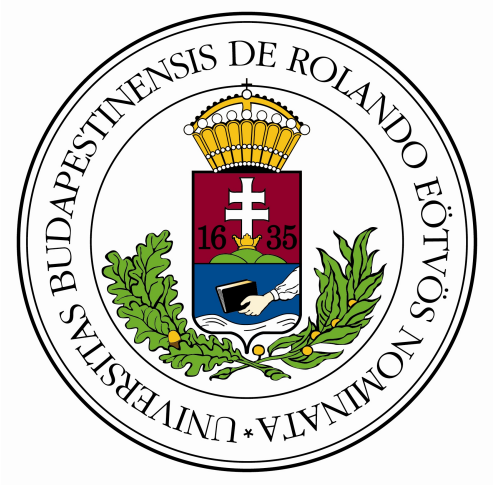
Machine Learning on EEG Signals

Data Mining and Machine Learning
Physics MSc.

Kristóf Furuglyás

January 2020



Contents

1 Introduction

The goal of this project is to familiarize myself with the use of machine learning in electroencephalographic signals. This project was originally made for the Data Mining in Physics course but I plan to work on it later also. Methods used here are more or less very rudimentary, most of them are only to show different aspects.

My topic is about somehow applying machine learning and data mining methods on EEG data. Since I have a lot of data and different measures, my main goal is to find such a method that can either group these measures or at least validate the theories we already have.

2 Experimental Background

Mismatch negativity is a well-known phenomena that is present in the brain when a deviant signal is present after many standard has been shown. By somewhat trying to learn the incoming signal and/or predicting it, the local connections (rather neural activities) form in such a way that the measures EEG signal drops down i.e. the brain gets used to the same input.

Let us say, the brain gets used to stimulus A, and after a lot of A (>20 but may vary) we present stimulus B. The response we get is not the same as if we just randomly shuffled the responses (equiprobable experiment). Two main ideas lay beneath the surface; first that the difference comes from the fact that the received signal is not the same as the predicted one, second is that when processing the signal other areas of the brain may join into the decomposition i.e. we do not measure the same population of neurons.

However, in our case we have two types of stimulus: visual and auditory. This setup allows us to measure whether conditional mismatch negativity exists. Combining the two types of stimulus with the two kinds (horizontal or vertical grating to the eyes and low or high pitch noise to the ears) gives us all in all 4 types of stimulus, let us call them A-D. From this one can identify the mismatches: only visual, only auditory or bimodal. A short summary can be seen on figure 1.

The experiment has been done many times, both on anaesthetised and awake mice, data was acquired in the visual and anteriolateral cortices of the brain by Holland biologists. During my job I always worked on only one anaesthetised mouse, but I can access the data easily if needed. For the main experiment the stimulus lasted for 500 ms with 1.5 sec of ITI (inter-time interval). One session (when one was the standard) meant approximately 600 stimulus with a $32kHz$ of sampling frequency, all in all I have the access to nearly a terabyte of raw EEG data.

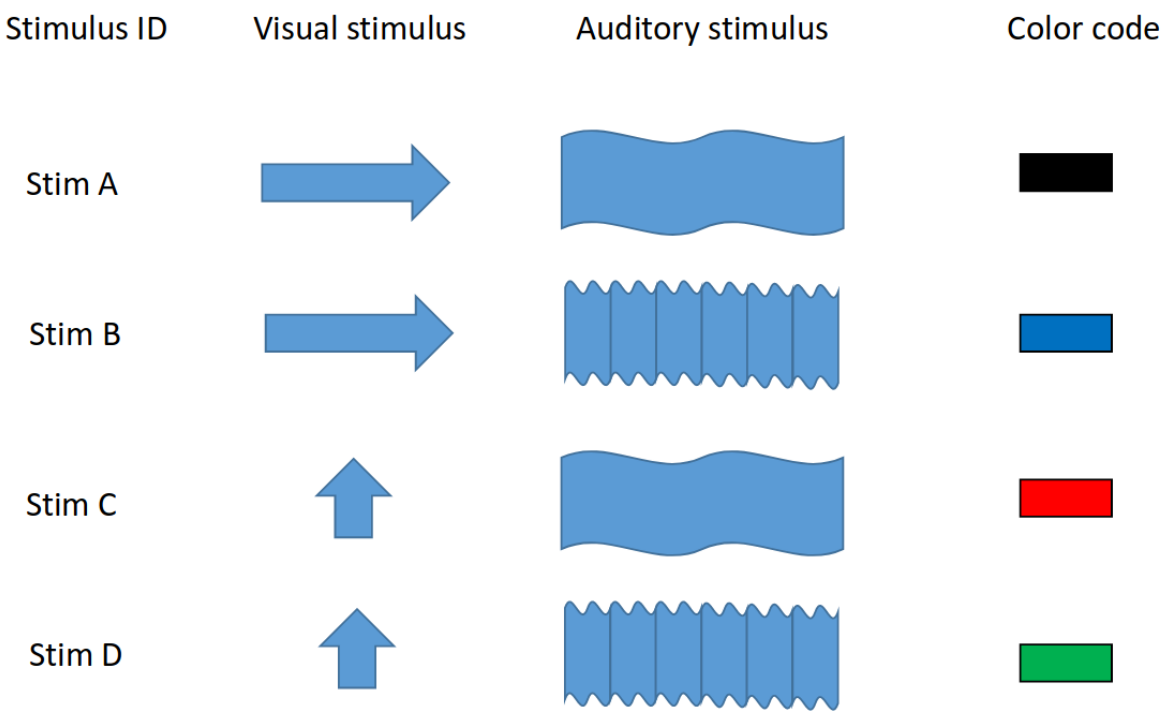


Figure 1: Stimulus description

3 Goals

As it was mentioned in the introduction data clustering can be useful and maybe dimension reduction. For the latter TSNE and UMAP are available, yet, I have no precise idea about the former but maybe Andrew NG's publications (which was mentioned by Bálint) might help – use of ML in neurology. All the results I have so far can be tested both by supervised and unsupervised learning, both in time and frequency domain.

4 Data preparation

Since we are talking about large files, here I am only going to load in one epoch, where the A stimulus was the standard (10 ms after the beginning and before the end to avoid the on/off effects). This means that there are way more stimulus from type A. The following lines are showing the cardinality of each one.

- There are 426 trials for STIM_a.
- There are 19 trials for STIM_b.
- There are 20 trials for STIM_c.
- There are 20 trials for STIM_d.
- There are 485 trials all in all.

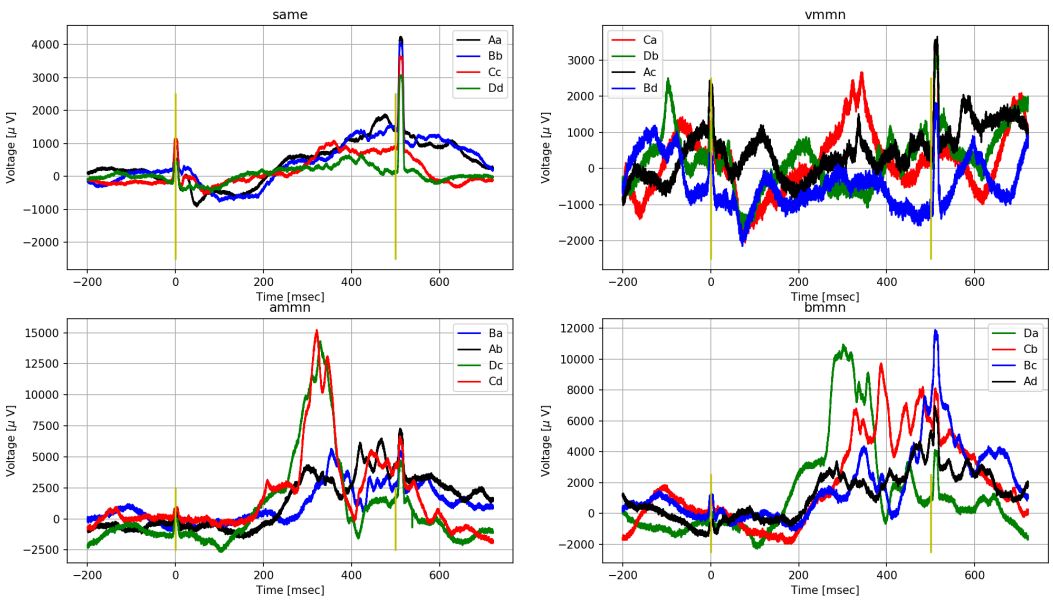


Figure 2: A new underlying phenomenon, the *conditional mismatch negativity* is present on the bottom left figure. There is a significant difference in the signals based on the stimuluses.

To prove that there is an underlying phenomenon going on, please take a look at the figure below. The figure show a channel (number 45) from the anteriolateral cortex. The subplots are based on the types of mismatches negativities (no mismatch, visual, auditory, bimodal). The lines on one subplot mean the relation of the standard and the deviant stimulus, for example, on the auditory mismatch negativity's (ammn) subplot the Ba means that the response for the signals that were averaged was A, but many signals before that had been the stimulus B. The number of signals that were averaged are around 20, except for the no mismatch subplot, where there were more than 400. All subplots show total zero nowhere, therefore we can assume that the averaged signals can be somehow grouped into one. Furthermore, there is a new phenomena called *conditional mismatch negativity*, which can be seen on the subplot ammn. The responses Cd and Dc showed a significantly higher response than Ab and Ba between 200 and 400 ms. A and B stimuluses were visually horizontally grating whilst C and D were vertically.

Moreover, since there are a lot of scales in which the data are, the normalisation of them had been done in such a way that the mean of each bunch of data – given to the machine to learn – has a mean of 0 with a standard deviation of 1.

5 Time domain

As it was mentioned, a total of 485 trials with $32kHz$ of sampling frequency through $480ms$ would be an unnecessarily huge amount of data. Due to the limitation of computing capacity, for each signal, I am going to re-sample the signal in such a way, that the new sampling frequency is going to be only $2kHz$. There has been a discussion with Zoltán Somogyvári (my boss) about this, and according to him, there is no significant change in time domain between the two sampling frequencies, moreover, this

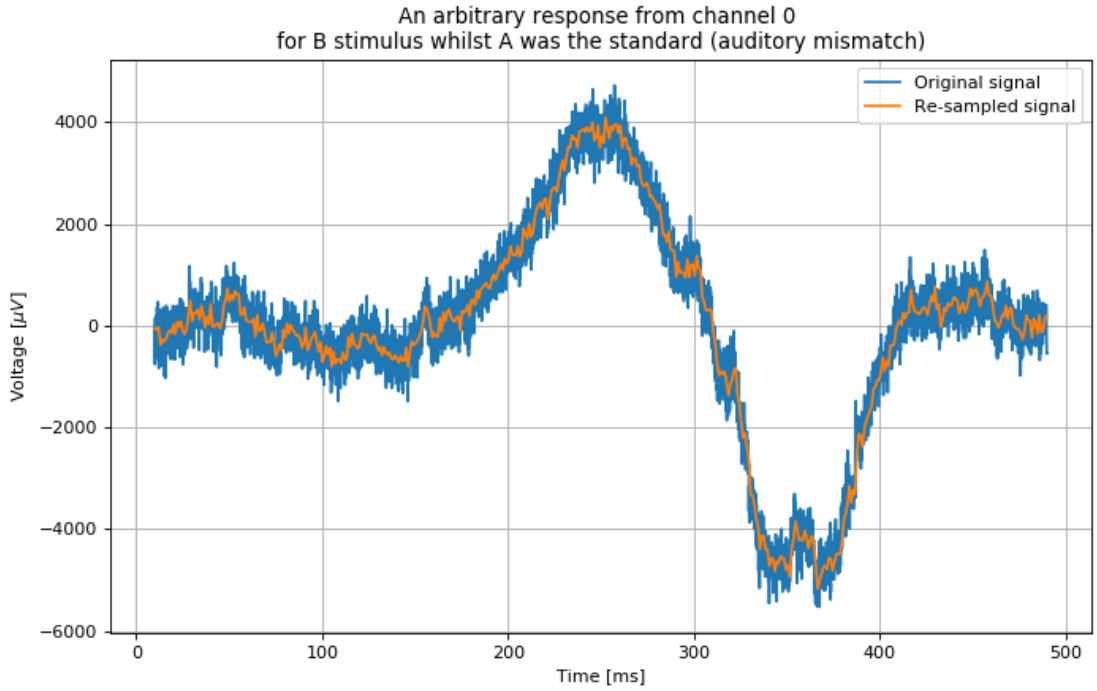


Figure 3: The original and the re-sampled signal from the same source.

would somewhat eliminate some of the noise – see fig 3. Further information about re-sampling can be found on python's official website.

5.1 Dimension reduction

Two main dimension reduction methods have been used to get a deeper meaning of the data which are the TSNE and the UMAP embedding. For further information, click on the links provided in their names.

Three types of trials have been made to separate the data:

- **by stimulus type**: in this version every data point is the response to a given stimulus on one channel, meaning there shall be a total number of 485 points on one figure,
- **by channels**: in this version, the data is grouped together by the stimulus type, and one point means the response to one epoch by one channel, so there are $\sim 20 \times 64 \approx 1240$ dots on each figure;
- **by epochs**: one epoch means one particular stimulus (which are by the way labelled with their timestamps and last for half a second). The reason behind this type of separation is to see whether there is a fatigue-factor in the mice. Since these are grouped together by the same means as the latter, there are the same number of points also.

All of the figures can be found in the appendix, in the core of this document I listed the ones which are meaningful. These are the following divided into bullet points based on the separation types discussed above:

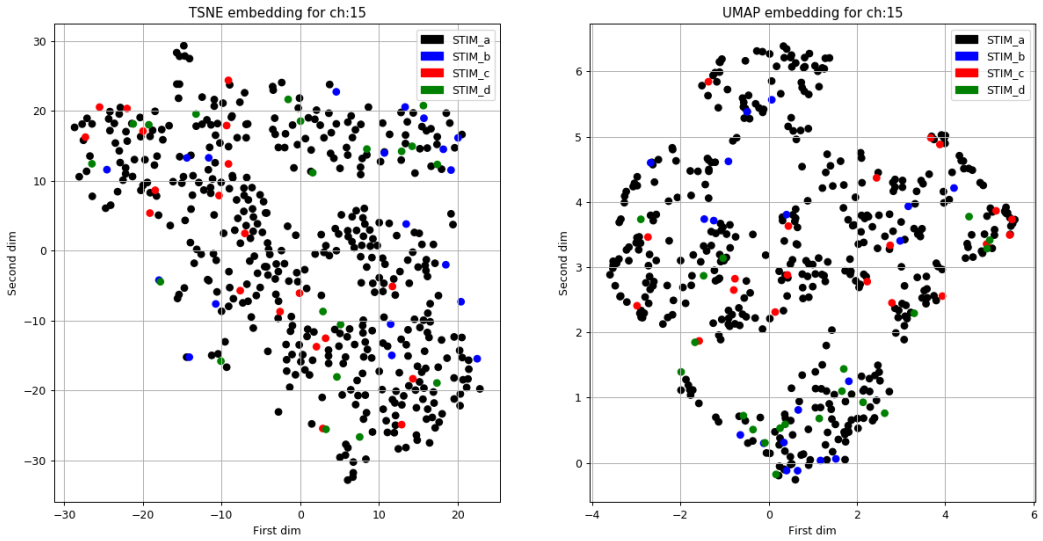


Figure 4: The dimension reduction techniques for the channel 15 – from the visual cortex. As one can see, no real distinction can be made based on the type of stimulus.

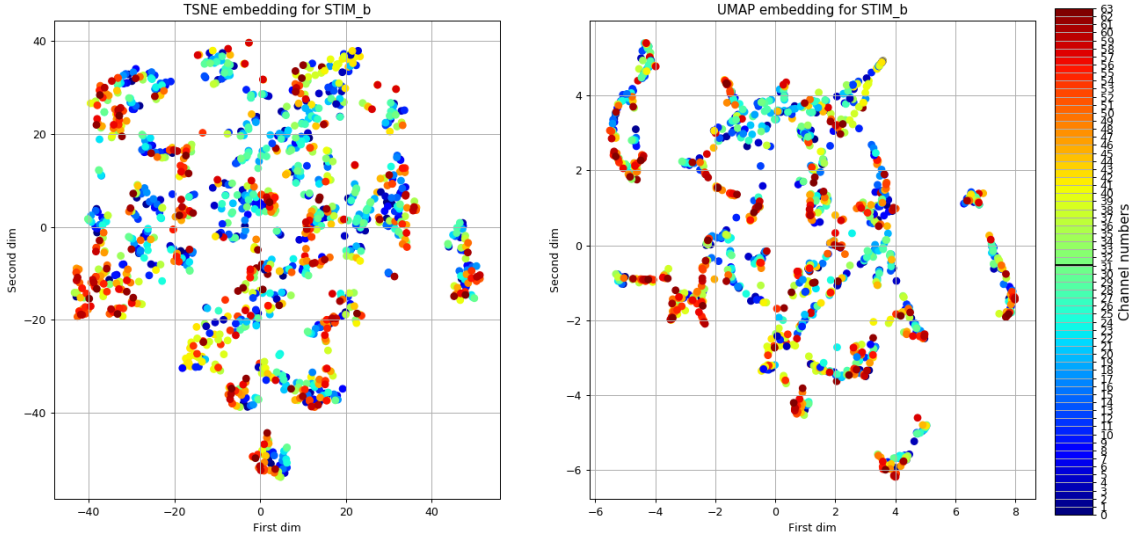


Figure 5: The dimension reduction techniques for stimulus B. The different colours mean different channels and below 31 there are the channels from the visual cortex and above that are the ones from the anteriolateral.

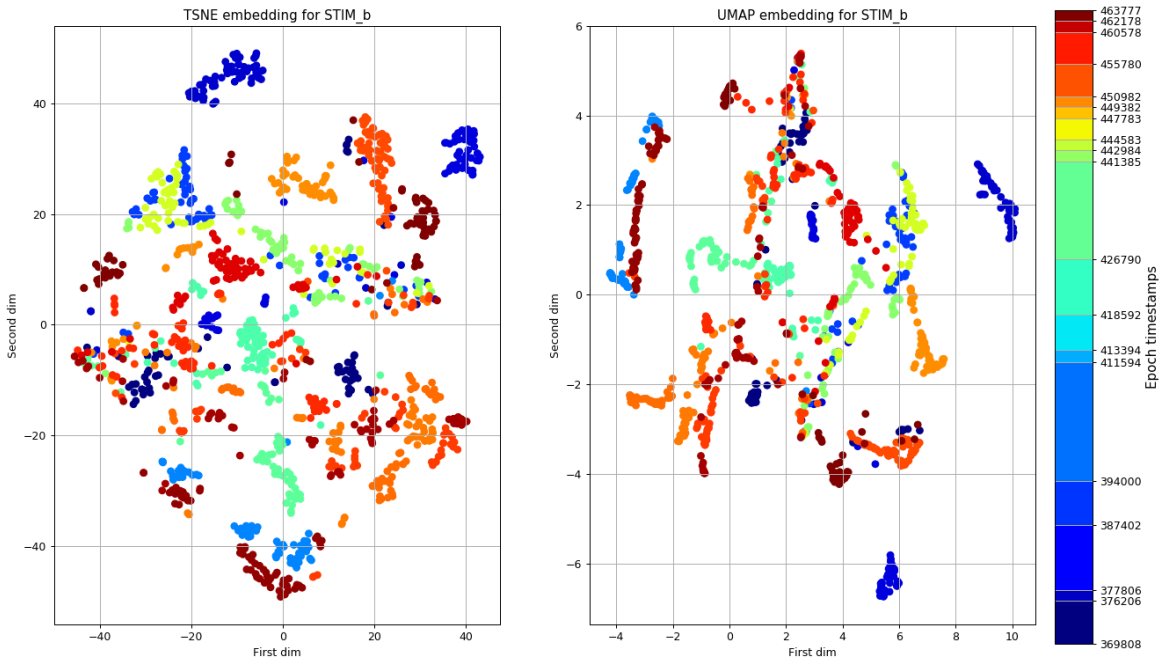


Figure 6: The dimension reduction techniques for stimulus B. The different colours mean different epochs. The spacing between the epochs are related to the time there was between them. There shall be 64 points to every colour.

- **by stimulus type** on figure 4. one can see that there are no meaningful separation on one channel only from the visual cortex. Reminder: the number of black points is higher, since in this trial, the standard stimulus was A.
- **by channels**: fig 5. shows us every response to stimulus B. In my opinion, there is no real significant gain of information on these figures.
- **by epochs**: separation of signals by epochs was more or less successful – see fig 6. The different colouring did show us how the groups are formed. Clear groups are visible in both methods, though, they might overlay and the presence of time cannot be seen here. Further improvements could be grouping the signals first by epochs then by the depth in the brain.

6 Frequency Domain

Analysing the signal in time domain is costly, therefore the Fourier representation of the signal is more convenient. Furthermore, even though the sampling frequency was $32kHz$, we do not need frequencies above $1kHz$ for biological reasons.

The resulting table consist all the incoming signal from all the channels (64) for all the trial (485), therefore having $64 \cdot 485 = 31040$ rows – same as in the time domain. The coloumns are mostly the amplitudes for the given frequency, for the stimulus I already created a dummy table, we have a channel number (ch) a timestamp (ts) and two colors; one for the stimulus type, the other for the place in the brain (anteriolateral (red) or visual (blue) cortex). Since low frequencies cannot be measured exactly (see warning message above), I am thresholding them at $11Hz$.

6.1 Dimension Reduction

Here the original dimensions are the frequencies – $\sim 2.08Hz$ apart – and the embedding dimensions are also 2, same again as it was for the time domain. Here I again used TSNE and UMAP for dimension reduction, and the separations were also the same.

The resulting figures carry important information based on their splitting:

- **by stimulus type** on figure 7. one can see that there are no meaningful separation on one channel only from the visual cortex. Reminder: the number of black points is higher, since in this trial, the standard stimulus was A.
- **by channels**: fig 8. shows us every response to stimulus B. First thing that is impressive is the heart-shaped object that is present for all of the stimuluses with TSNE. The brain areas are clearly separated on the two halves of the heart. This is a great success. UMAP also did separate the signals which therefore is an indicator that t changing to Fourier domain is an effective way to separate the signals.
- **by epochs**: separation of signals by epochs was successful – see fig 9. However, even though the separation of groups is visible, the presence of time is not there, unfortunately. For example, if the tail of the heart would be the stimuluses at the beginning and the bubbles of the end would be last ones, then we could say that there is a temporal learning in the mouse.

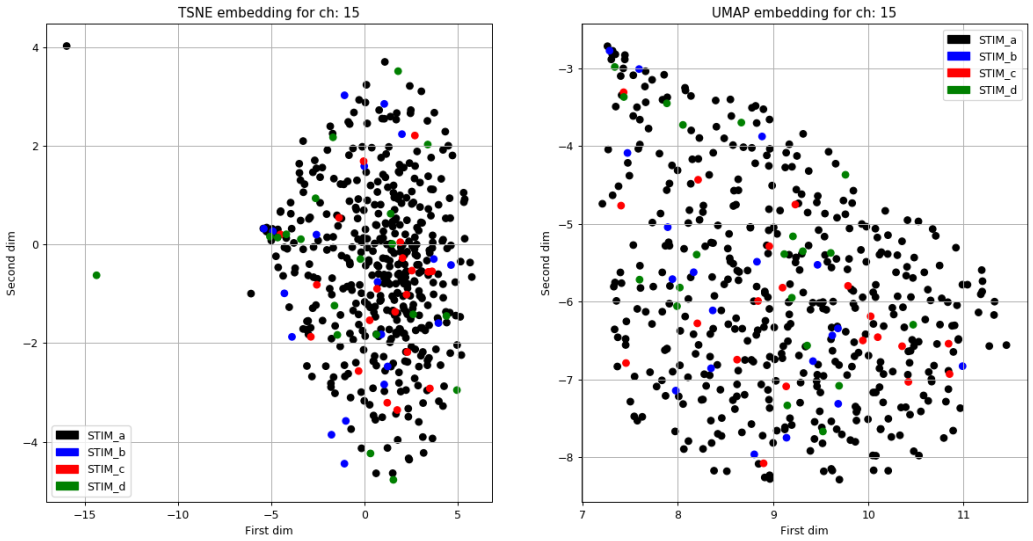


Figure 7: The dimension reduction techniques with frequencies for the channel 15 – from the visual cortex. As one can see, no real distinction can be made based on the type of stimulus.

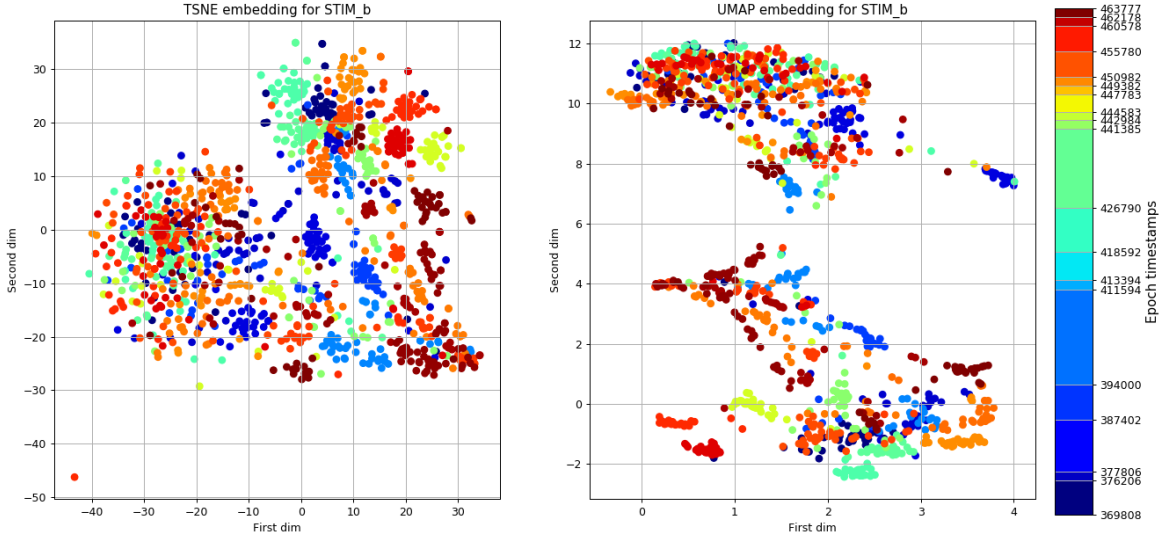


Figure 8: The dimension reduction techniques for stimulus B with frequenciey. The different colours mean different channels and below 31 there are the channels from the visual cortex and above that are the ones from the anteriolateral.

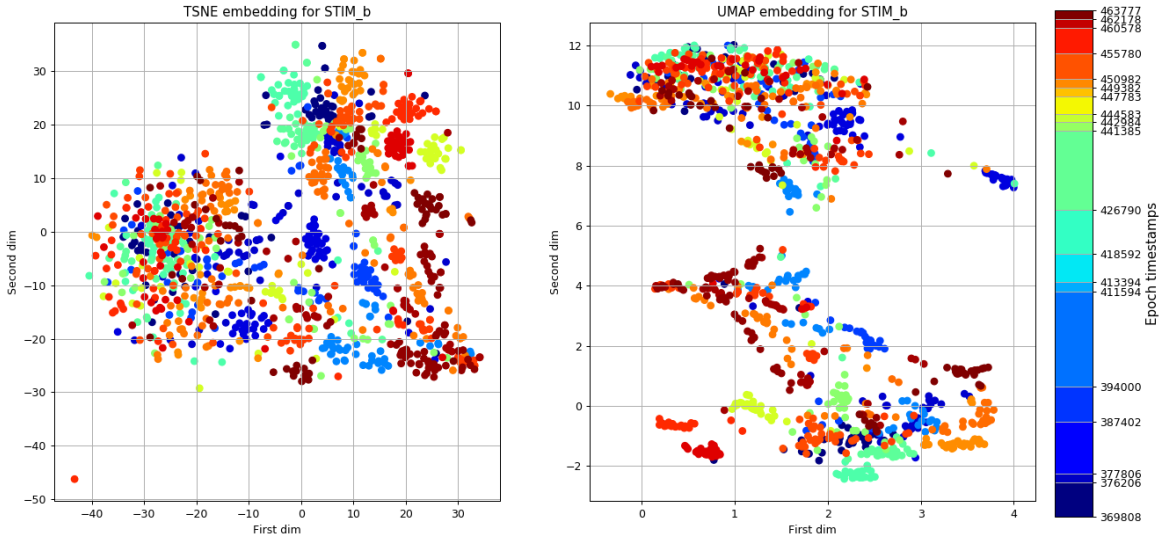


Figure 9: The dimension reduction techniques for stimulus B for frequencies. The different colours mean different epochs. The spacing between the epochs are related to the time there was between them. There shall be 64 points to every colour.

6.2 Linear Regression

Since it seemed appealing to stay in the frequency domain, I tried to use a simple linear regresssion method on the data. For further information on linear regression, I encourage you to visit the Wikipedia.

I used a simple linear regression method in which I tried to predict the type of the stimulus, the channel number and the timestamp of the epoch. Therefore, the separation of the signals are the same again, and the results are:

- **by stimulus type** on figure 10. one can see that there are on one channel from the anteriolateral cortex, the coefficents are higher for lower frequencies. The reasoning behind this might lay in the field of bilogy i.e. I will consult with my boss.
- **by channels**: fig 11. shows us the coefficients to stimulus B. In my opinion, there is no new information that is useful for us.
- **by epochs**: separation of signals by epochs was successful – see fig 12. The presence of the periodical high values is eye-catching. Further investigation is needed to evaluate this properly.

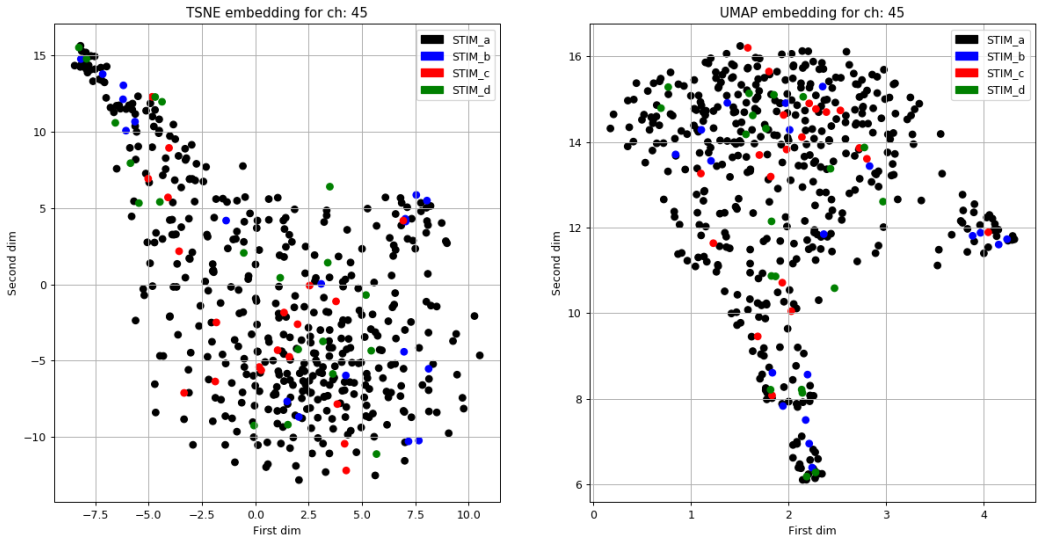


Figure 10: The linear regression coefficients for the channel 45 – from the anteriolateral cortex. What is interesting to see here is that absolute values of the coefficients are higher for lower frequencies.

Figure 11: The linear regression coefficients for stimulus B. The different colours mean different channels. There is no obvious information in either of the figures created with this method

Figure 12: The linear regression coefficients for stimulus B. The different colours mean different epochs. Even though the coefficients are small, there are spikes in there periodically.

7 Discussion

In this project the most useful method I think was the dimension reduction. With the insight of a biologist, there must be some underlying phenomenon going on. The dimension reduction techniques came in utterly handy. Both the time domain representation and the frequency domain representation brought new information to the surface.

In my opinion, the linear regression method did not work as it had been planned. Maybe with noise reduction or more tuned parameters the method would have worked better.

Appendices

A Time Domain Dimension Reduction Results

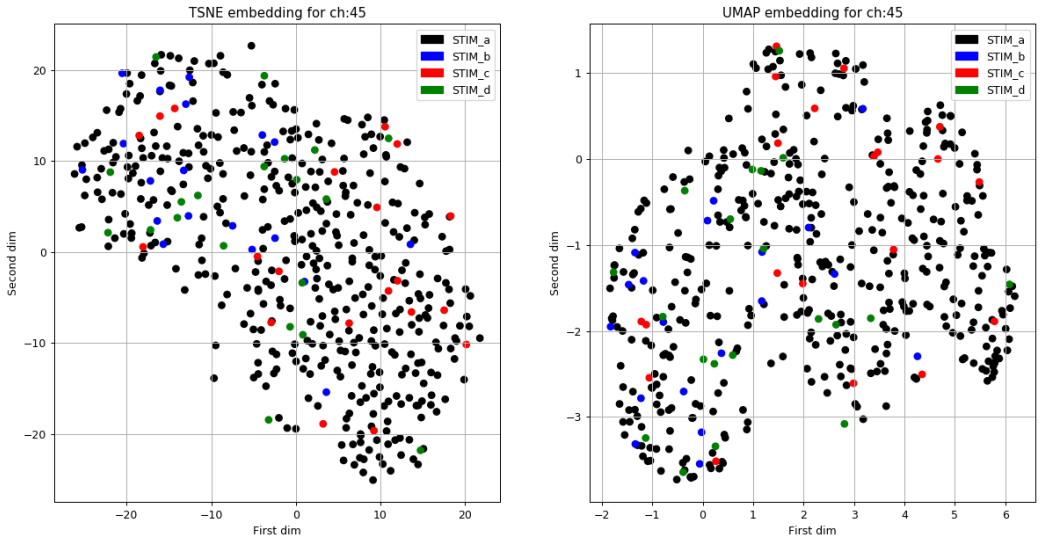


Figure 13: The dimension reduction techniques for the channel 15 – from the anteriolateral cortex. Same again, no real distinction can be made based on the type of stimulus.

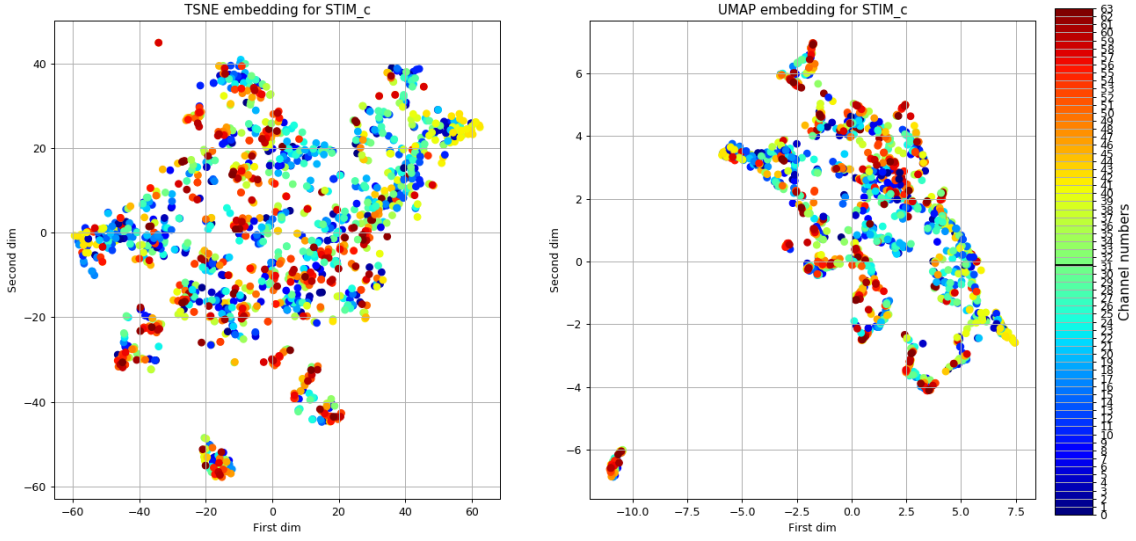


Figure 14: The dimension reduction techniques for stimulus C. The different colours mean different channels and below 31 there are the channels from the visual cortex and above that are the ones from the anteriolateral.

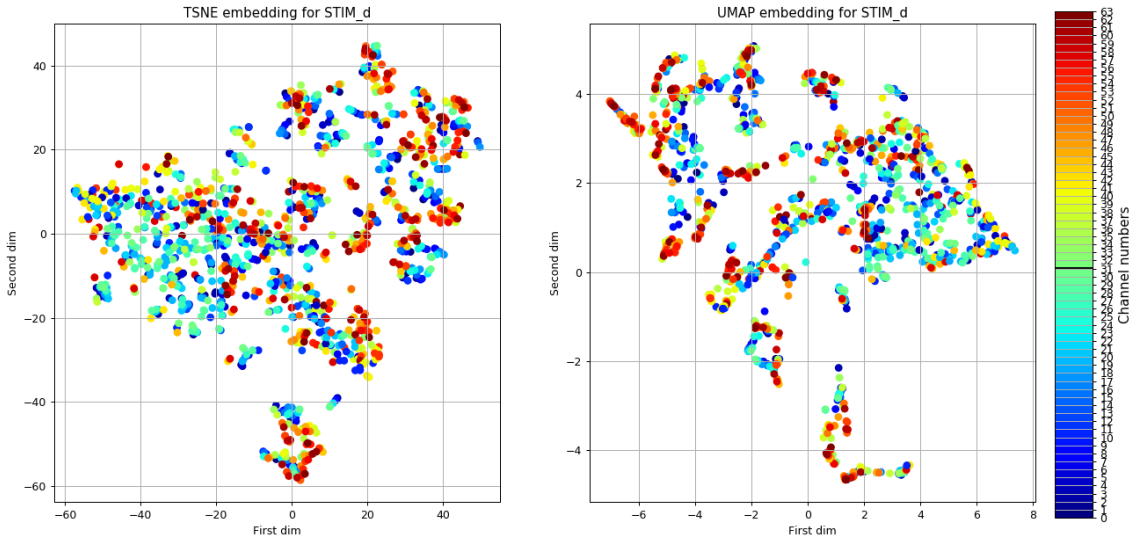


Figure 15: The dimension reduction techniques for stimulus D. The different colours mean different channels and below 31 there are the channels from the visual cortex and above that are the ones from the anteriolateral.

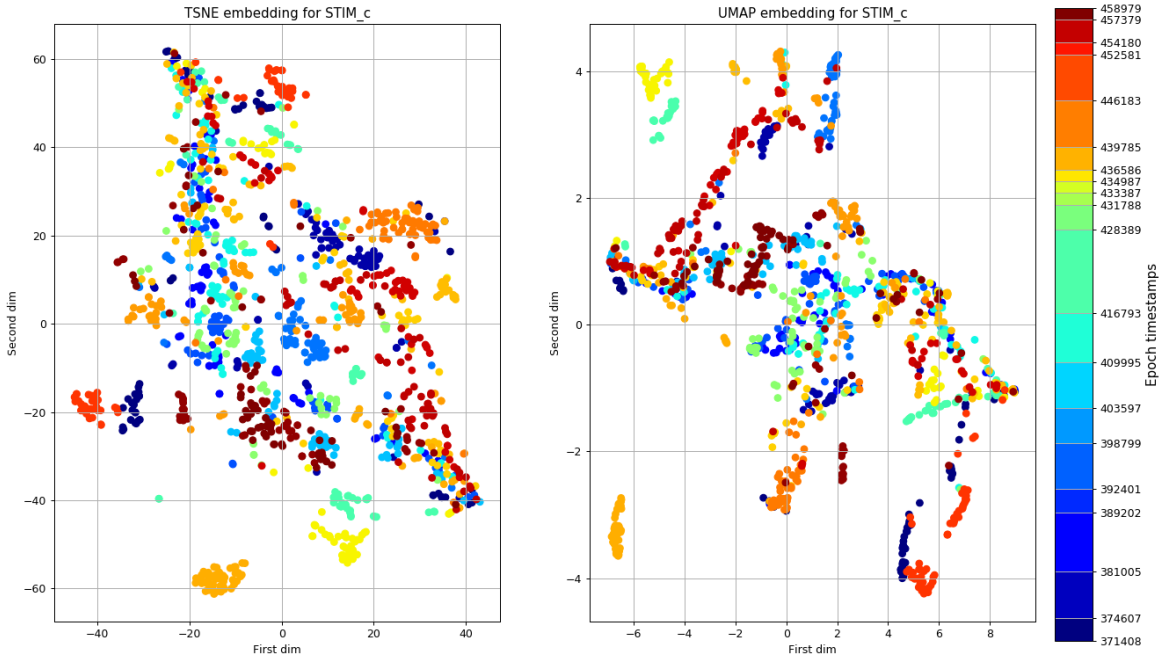


Figure 16: The dimension reduction techniques for stimulus C. The different colours mean different epochs. The spacing between the epochs are related to the time there was between them. There shall be 64 points to every colour.

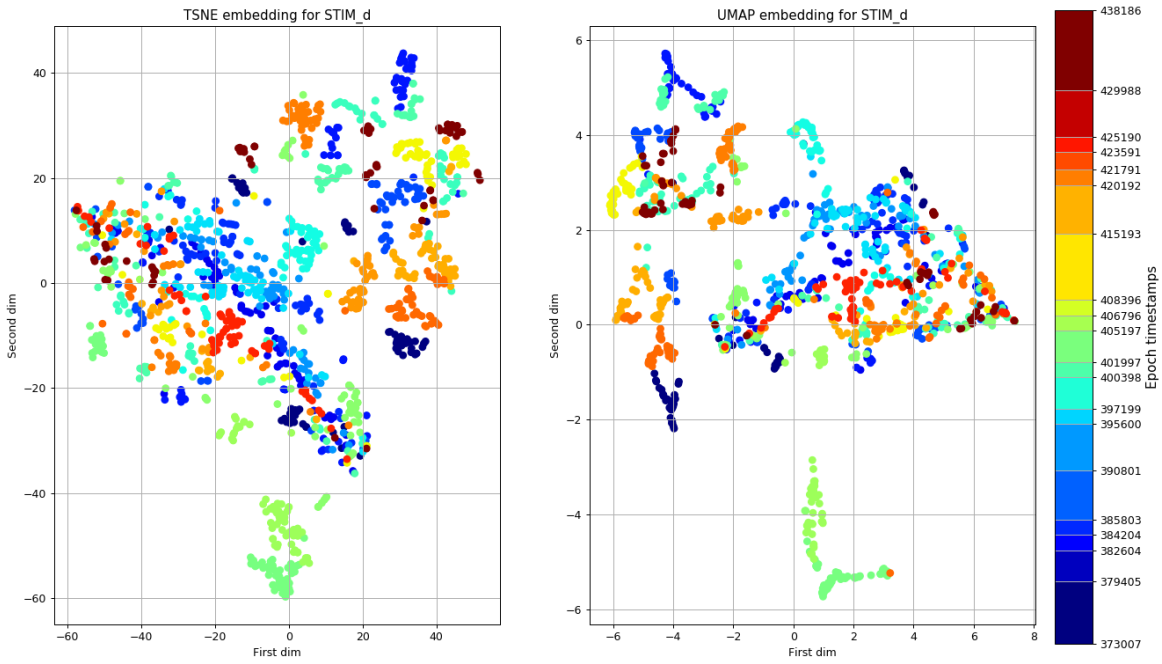


Figure 17: The dimension reduction techniques for stimulus D. The different colours mean different epochs. The spacing between the epochs are related to the time there was between them. There shall be 64 points to every colour.

B Frequency Domain Dimension Reduction Results

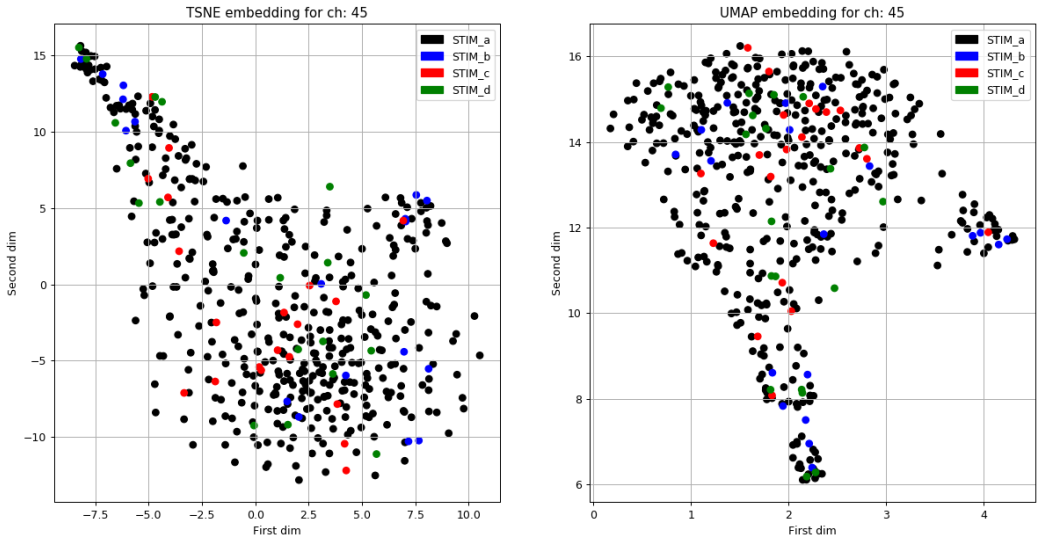


Figure 18: The dimension reduction techniques for the channel 15 with frequencies – from the anteriolateral cortex. Same again, no real distinction can be made based on the type of stimulus.

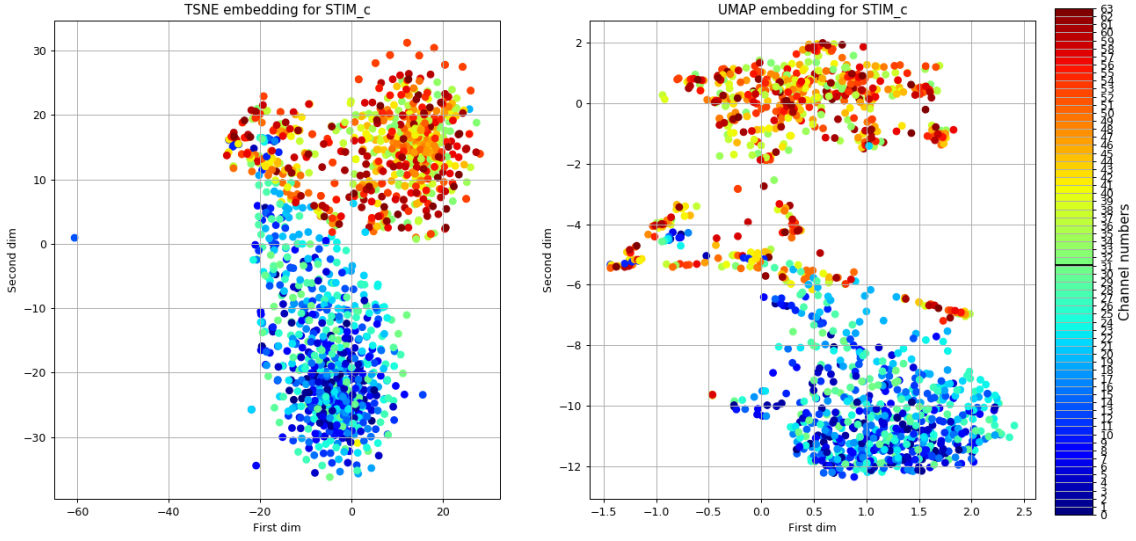


Figure 19: The dimension reduction techniques for stimulus C with frequencies. The different colours mean different channels and below 31 there are the channels from the visual cortex and above that are the ones from the anteriolateral.

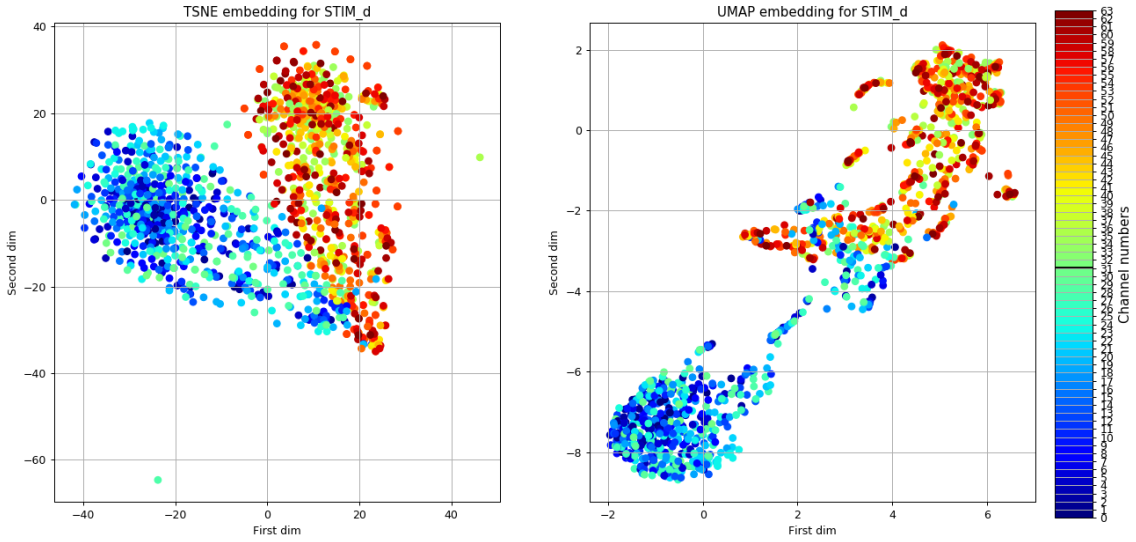


Figure 20: The dimension reduction techniques for stimulus D with frequencies. The different colours mean different channels and below 31 there are the channels from the visual cortex and above that are the ones from the anteriolateral.

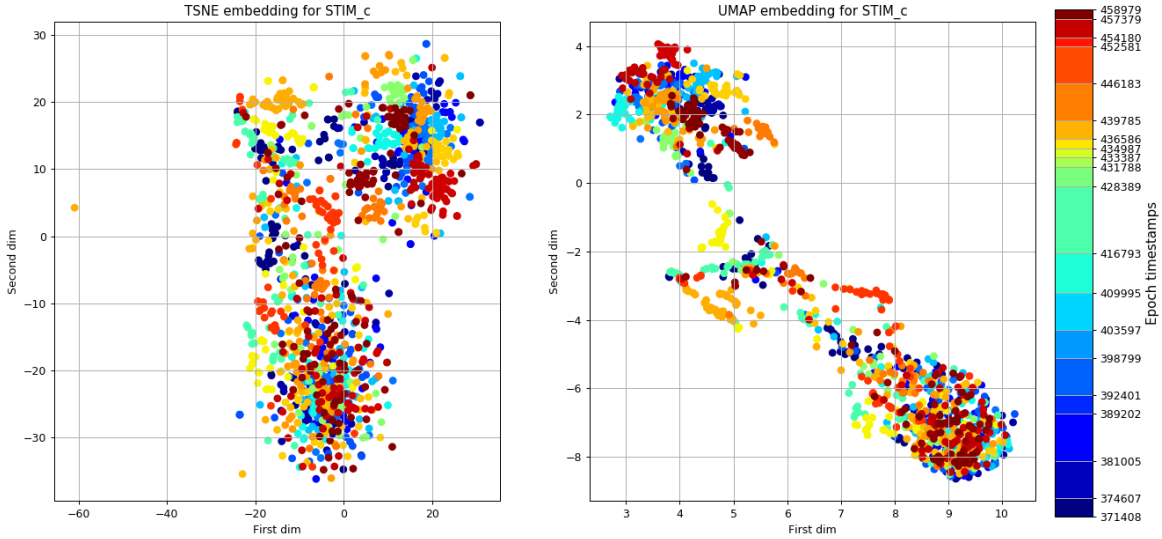


Figure 21: The dimension reduction techniques for stimulus C with frequencies. The different colours mean different epochs. The spacing between the epochs are related to the time there was between them. There shall be 64 points to every colour.

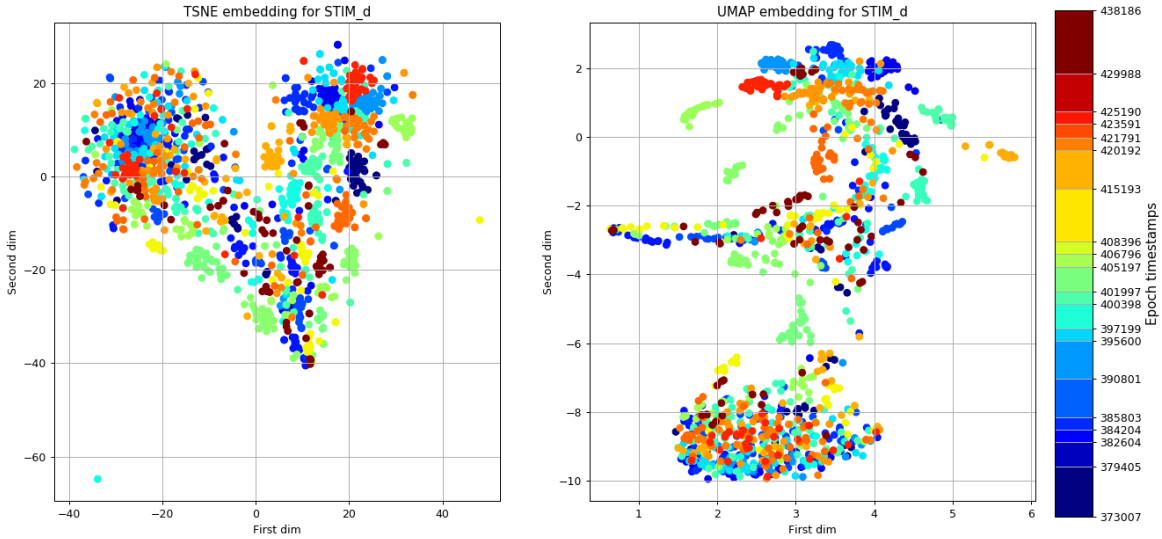


Figure 22: The dimension reduction techniques for stimulus D with frequencies. The different colours mean different epochs. The spacing between the epochs are related to the time there was between them. There shall be 64 points to every colour.

C Frequency Domain Linear Regression Results

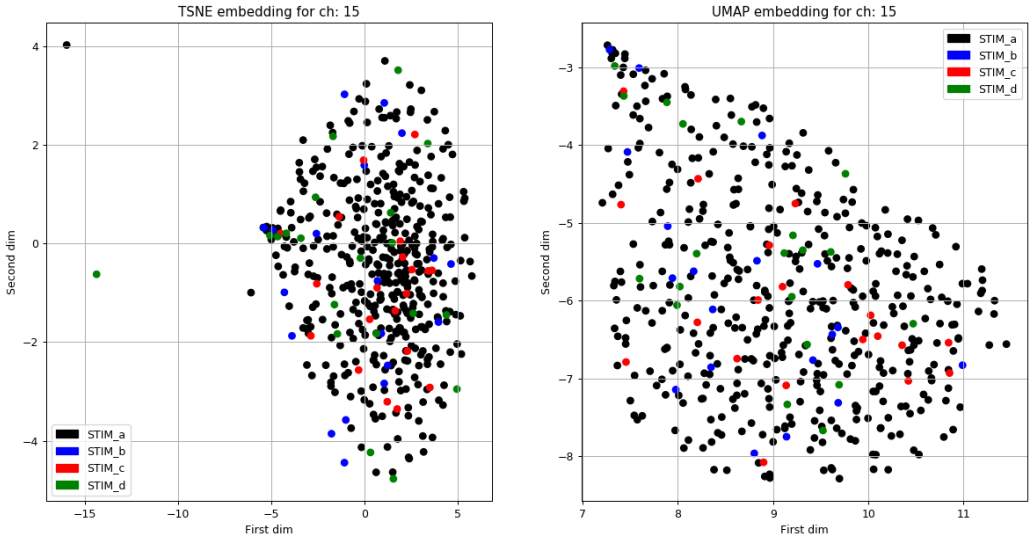


Figure 23: The linear regression coefficients for the channel 15 – from the visual cortex.

Figure 24: The linear regression coefficients for stimulus C. The different colours mean different channels. There is no obvious information in either of the figures created with this method

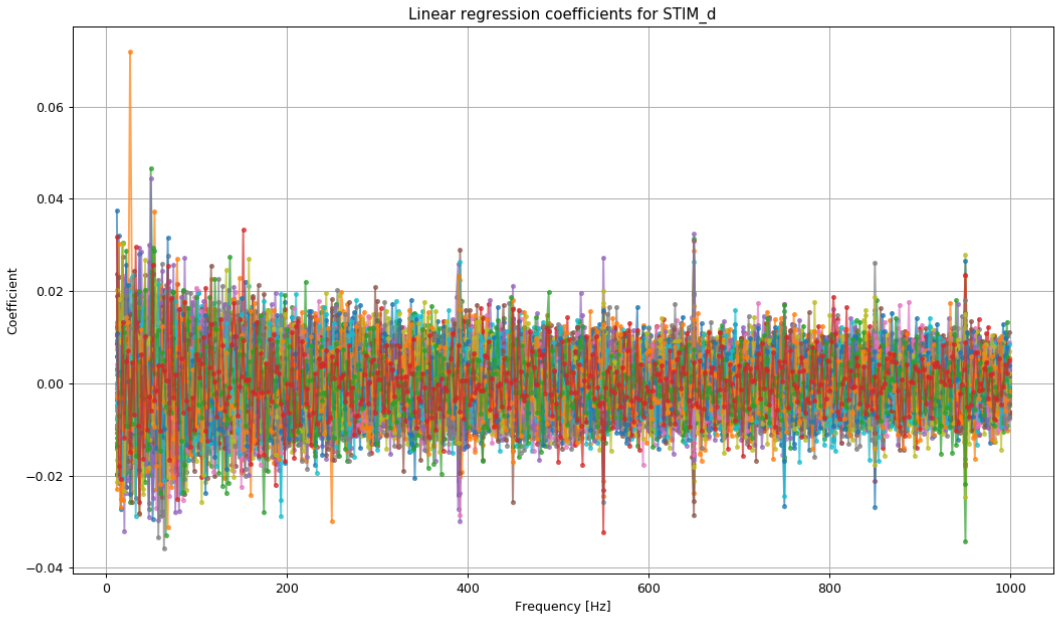


Figure 25: The linear regression coefficients for stimulus D. The different colours mean different channels. There is no obvious information in either of the figures created with this method

Figure 26: The linear regression coefficients for stimulus C. The different colours mean different epochs. Even though the coefficients are small, there are spikes in there periodically.

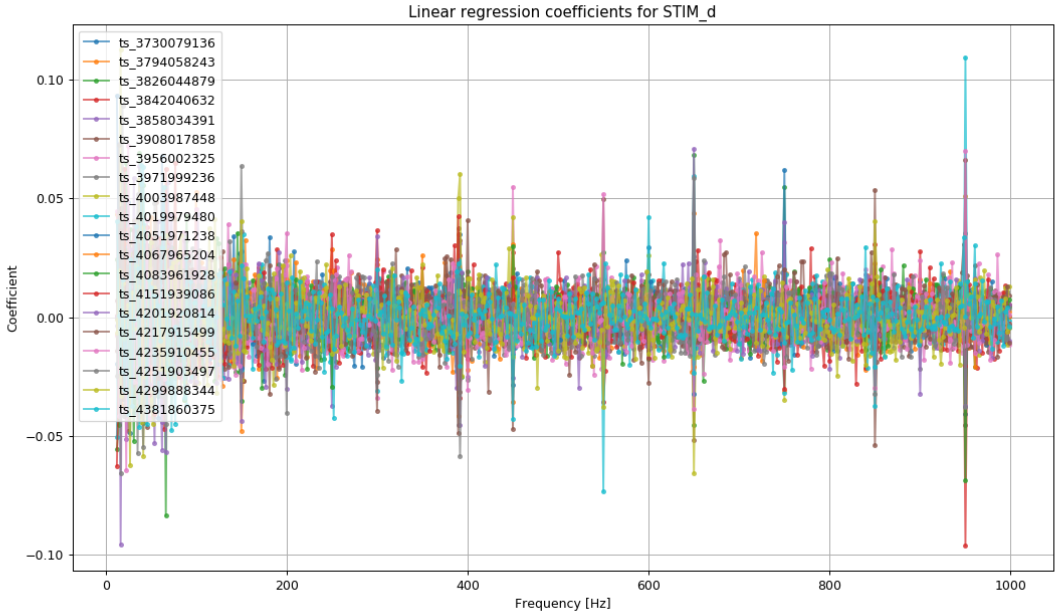


Figure 27: The linear regression coefficients for stimulus D. The different colours mean different epochs. Even though the coefficients are small, there are spikes in there periodically.