Methods:

### Experimental Design:

* Rats were subjected to sessions involving the investigation of a novel rat stimulus within a chamber.
* Sociability was determined by measuring the length of bouts exhibited by the subject rat during the investigation.
* Sessions were divided into two groups: affiliative (sessions with longer bouts) and aversive (sessions with shorter bouts).

### Selection of LFP Pairs:

* LFP recordings were obtained from various brain regions.
* A minimum of five sessions were required for each LFP pair in both the affiliative and aversive groups.
* Change in coherence between each pair was calculated by subtracting the mean coherence before interaction before stimulus insertion and the mean coherence during the interactions after stimulus insertion and before stimulus removal
* Mann-Whitney U tests were conducted to compare the coherence of each pair between the two groups.
* The following pairs were selected for analysis:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **area1** | **area2** | **stat** | **pval** | **freq** |
| CeA | MeD | 12 | 0.024456 | 4-12 |
| BMA | MePV | 36 | 0.327672 | 4-12 |
| MeD | STIA | 27 | 1 | 4-12 |
| AA | BMA | 7 | 0.177489 | 4-12 |
| AA | MeD | 30 | 0.020729 | 4-12 |
| EA | MeD | 45 | 0.967849 | 4-12 |
| BMA | MeD | 60 | 0.629796 | 4-12 |
| CeA | STIA | 17 | 0.370962 | 4-12 |
| CeA | MeD | 17 | 0.083064 | 30-80 |
| BMA | MePV | 28 | 0.954645 | 30-80 |
| MeD | STIA | 32 | 0.661172 | 30-80 |
| AA | BMA | 10 | 0.428571 | 30-80 |
| AA | MeD | 32 | 0.028108 | 30-80 |
| EA | MeD | 45 | 0.967849 | 30-80 |
| BMA | MeD | 54 | 0.945211 | 30-80 |
| CeA | STIA | 9 | 0.055278 | 30-80 |

### Data Filtering and Subset Selection:

* The significant area pair-frequency range combinations were identified as follows:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **freq** | **area1** | **area2** | **stat** | **pval** |
| 4--12 | AA | MeD | 30 | 0.020729 |
| 4--12 | CeA | MeD | 12 | 0.024456 |
| 30-80 | AA | MeD | 32 | 0.028108 |

* The dataset was filtered to include only sessions that contained measurements from at least one of these area pairs.
* After filtering, a total of 31 sessions remained for further analysis.

### Train-Test Split:

* The 31 sessions were divided into two subsets: train and test, based on the following criteria:
* Sessions containing recordings from all three areas (CeA, AA, MeD) were selected for the test subset (n=11).
* Sessions with missing files were chosen for the training subset (n=20).
* To facilitate data imputation for missing values, all recordings from rat number 23 (with 2 sessions) and rat number 4 (with 3 sessions), which contained measurements from all areas, were moved from the test set to the training set.
* Additionally, recordings from rat number 19 (with 4 sessions, 1 aversive and 3 affiliative) were transferred from the training set to the test set to balance the distribution of classes in the test set.

To summarize:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **test** | | **train** | |
|  | **affiliative** | **aversive** | **affiliative** | **aversive** |
| **Rat number** |  |  |  |  |
| 3 | 0 | 0 | 2 | 2 |
| 4 | 0 | 0 | 1 | 2 |
| 10 | 0 | 4 | 0 | 0 |
| 12 | 0 | 1 | 0 | 0 |
| 15 | 0 | 0 | 1 | 2 |
| 16 | 0 | 0 | 1 | 2 |
| 17 | 0 | 0 | 0 | 3 |
| 19 | 3 | 1 | 0 | 0 |
| 20 | 1 | 0 | 0 | 0 |
| 21 | 0 | 0 | 1 | 0 |
| 23 | 0 | 0 | 1 | 1 |
| 26 | 0 | 0 | 2 | 0 |
| sum | 4 | 6 | 9 | 12 |

To impute missing data in both the training and test subsets, we employed the MissForest algorithm. MissForest is a machine learning algorithm specifically designed for imputing missing values in datasets. It utilizes a random forest approach to predict missing values based on the observed values and other variables in the dataset. The algorithm iteratively imputes missing values until convergence is achieved.

Remark: The MissForest algorithm was introduced by Stekhoven and Buehlmann in their paper titled "MissForest—Non-parametric missing value imputation for mixed-type data" (2012).

After imputing the missing data, we trained a random forest classifier using the imputed training set and evaluated its performance on the imputed test set. It is important to note that imputation was performed separately for each dataset (training and test) to prevent data leakage and maintain the integrity of the evaluation.

The following table presents the results obtained from the random forest classifier on both the training and test sets:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Accuracy** | **Precision** | **Recall** | **F1 Score** |
| Test | 0.8 | 0.667 | 1.0 | 0.8 |
| Train | 1 | 1 | 1 | 1 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **index** | **files** | **GT** | **predicted** | **correct** | **confidence** |
| 3 | chamber\_Rat10-probe13-day1\_Behavior\_and\_Optogenetics\_TimeStamps.mat | aversive | aversive | TRUE | 0.59 |
| 4 | chamber\_Rat10-probe13-day2\_Behavior\_and\_Optogenetics\_TimeStamps.mat | aversive | aversive | TRUE | 0.82 |
| 5 | chamber\_Rat10-probe13-day3\_Behavior\_and\_Optogenetics\_TimeStamps.mat | aversive | aversive | TRUE | 0.94 |
| 6 | chamber\_Rat10-probe13-day4\_Behavior\_and\_Optogenetics\_TimeStamps.mat | aversive | affiliative | FALSE | 0.64 |
| 7 | chamber\_Rat12-probe16-day2\_Behavior\_and\_Optogenetics\_TimeStamps -.mat | aversive | aversive | TRUE | 0.64 |
| 17 | chamber\_Rat19-Probe18-day1-Behavior\_and\_Optogenetics\_TimeStamps.mat | aversive | affiliative | FALSE | 0.68 |
| 18 | chamber\_Rat19-Probe18-day3-Behavior\_and\_Optogenetics\_TimeStamps.mat | affiliative | affiliative | TRUE | 0.83 |
| 19 | chamber\_Rat19-probe18-Day2-Behavior\_and\_Optogenetics\_TimeStamps.mat | affiliative | affiliative | TRUE | 0.81 |
| 20 | chamber\_Rat19-probe18-day6-Behavior\_and\_Optogenetics\_TimeStamps.mat | affiliative | affiliative | TRUE | 0.81 |
| 21 | chamber\_Rat20-Probe19-day1-Behavior\_and\_Optogenetics\_TimeStamps.mat | affiliative | affiliative | TRUE | 0.81 |

## Figures:

A picture containing screenshot, art

Description automatically generated

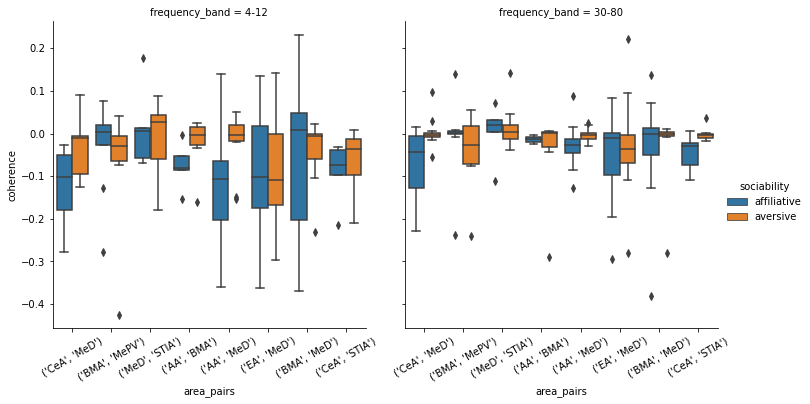


Figure 1. Coherence in LFP changes before and during stimulus insertion between pairs of brain regions.

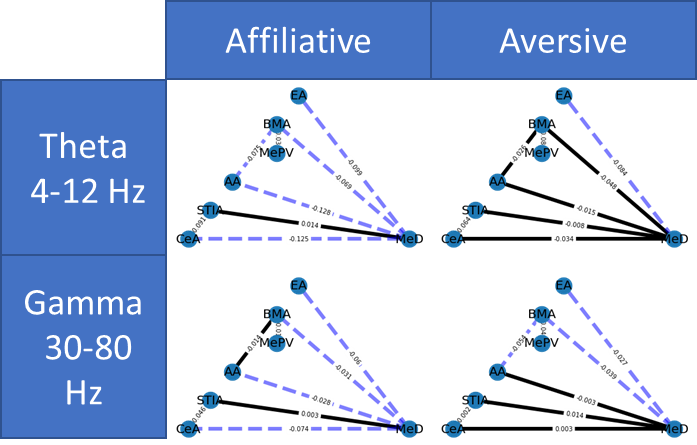


Figure 2. Network graph of the coherence between each pair of regions, dashed lines indicate lower values than the median, and solid lines indicate higher than or equal to the median coherence.

A screenshot of a computer screen

Description automatically generated with low confidence

A screenshot of a computer screen

Description automatically generated with low confidence

A picture containing screenshot, design, illustration, art

Description automatically generated

Figure 3. Visualization of parameters of the training dataset before (nan values removed) (top) and after (middle) imputation using miss forest algorithm and after dimensionality reduction of the imputed dataset using tsne algorithm (bottom).