**Data and Code structure for LFP analysis of data collected by Dr. J Patrick Mayo**

This file summarizes the data structure and codes for doing spike and LFP analysis on data collected by Dr. J Patrick Mayo in Dr. John Maunsell’s lab. The monkeys performed an attention task with valid, invalid and neutral cues, while data was recorded from both hemifields of area V4. For details of spiking data analysis, see Mayo and Maunsell, 2016, JNS.

Within the parent project folder that contains everything related to this dataset, there is a Data folder. In addition, there is one programs folder for each project. Hence, the contents of this parent folder are as follows:

* Parent\_Folder
  + Data
  + MayoProjectPrograms (<https://github.com/supratimray/MayoProjectPrograms>)
  + MayoProject2 (<https://github.com/surya0021/MayoProject2>)
* The “Data” folder contains the following
  + extractedData – original data provided by Patrick, along with extracted LFP data. Data from 2 monkeys and 26 sessions (of which 25 were used). Size: 120 GB.
  + segmentedData – following a particular convention used in our lab, data is subsequently segmented in a specific format and saved in this folder. Size: 24.4 GB.
  + savedData & savedDataSummary – see details under MayoProjectPrograms below.
* MayoProjectPrograms – this was the first project on this dataset, details of which can be found in Prakash et al., 2021, Cerebral Cortex. This folder has a lot of files, which are used to convert extractedData to segmentedData (for example, by running runSaveSegmentedDataMayo.m which runs saveSegmentedDataMayo.m), and to generate intermediate data that is subsequently used for all analyses. However, this intermediate data is not stored within this folder but instead in the main Data folder (in savedData and savedDataSummary folders; see above). The display programs in this folder read from these intermediate folders to show the results. They do not work unless these intermediate data are available, which are several GB in size.
* MayoProject2 (current folder) – this is the second project on this dataset. Here, we will study how different neural measures – spikes, LFP power, LFP phase, phase consistency, noise correlations etc can be used to predict behavior (hit versus miss). Unlike the first project, intermediate data will also be stored within this folder. One set of programs generate and save this intermediate data. Another set of programs then analyze this data and show results.

The programs in MayoProject2 folders are described below.

**Codes and data files for MayoProject2**

## I – Codes to locally save data

**How to get “intermediate data” for further analysis?**

**Data files:** To start with, the following data folder/files are required and should be present in “Data” folder inside the parent folder within which this folder is kept. For example, if the location of this folder is X/MayoProject2, then we expect the Data folder to be X/Data (X being the parent folder).

1. ***extractedData***: contains codes files and DAT files which are used to obtain task related and behavioral data.
2. ***segmentedData***: contains LFP, spikes data for each attention conditions segmented between -0.5 s to 0.1 s relative to target onset.
3. ***electrodeArrayList***: contains the list of good electrodes of all the 25 sessions. (This file is present in the savedDataSummary folder inside the Data folder). It is also saved locally in X/MayoProject2/savedData later.

**Data codes:** These set of programs are used to save “intermediate” data which are subsequently used by analysis codes.

1. ***getAttentionExperimentDetails.m*** – This code gives the filenames of each session data.
2. ***getBehavior.m***: This program gives the behavioral performance of the monkey. This also saves the behavioral data locally in savedData/behaviorData if it is not present there already.
3. ***saveDataForAnalysis.m*** – This program concatenates LFP and spikes data for one single orientation of all attention conditions. This datafile is saved locally in a new *savedData/neuralData* folder and used by the analysis codes.
4. ***runDisplayBehaviorAndSaveData*** – It is the main program to locally save data. It also displays the orientation that is used for each condition.

If you have access to the Data folder (in the parent folder as discussed above), you should first run runDisplayBehaviorAndSaveData. It will generate the folders neuralData (831 MB) and behaviorData (240 KB) within the X/MayoProject2/savedData folder. Once this data is saved, the programs mentioned above (except the first one) are not needed. The remaining programs read from this locally saved data so access to the original Data folder is also not needed. The savedData folder which contains neuralData and behaviorData is provided along with the codes in the MayoProject2 GitHub repository.

## II – Analysis Codes

1. ***getAnalysisMeasuresSingleElectrode*** – This program gives PSDs computed using multitaper (MT) method as well as firing rates for every trial and electrode. This function is called by displayResultsSingleElectrode (see later). It also locally saves the data in the savedData folder as singleElectrodeMeasureX where X is the TW factor used in MT analysis (number of tapers is 2TW-1; all analyses are done for X=5). Note that this program saves the mean MT power per trial (averaged across all tapers). It does not save the MT phases since they cannot be averaged across tapers. It also has an option to save FFT power and phase per trial. It is saved as singleElectrodeMeasuresFFT in the savedData folder. Please run this code first as *getAnalysisMeasuresSingleElectrode(5,1)* to get MT estimate for TW=5 as well as the FFT estimates, which should get saved locally as singleElectrodeMeasure5 and singleElectrodeMeasureFFT, respectively.
2. ***getMTValsSingleElectrode*** – This program is similar to the one above, except that for each trial, MT power and phase for each taper is saved separately. This is required for doing single trial PPC and power correlation analysis, where PPC or correlation is computed across tapers for each trial. Please run this code as *getMTValsSingleElectrode(5,’phase’)* to get both phase and power estimates (saved separately as singleElectrodeMTPhase5 and singleElectrodeMTPower5, respectively). These files are very large (For example, 9 times larger than the files generated in 1 above when TW=5). Note that this program makes the output of 1) redundant, since it can be obtained by simply averaging the MT power across tapers for each trial. But it is cumbersome to deal with large file sizes, so we have kept the output of the first program as well.
3. **getFFTValsSingleElectrode** – This program generates multiple estimates of power and phase per trial by dividing the time duration in numDivisions segments. Power and phase are estimated using FFT. Note that the frequency resolution is reduced by numDivisions here, so the file size is not as large as 2. Please run this program as *getFFTValsSingleElectrode(10,’phase’)* to get singleElectrodeFFTPhase10 and singleElectrodeFFTPower10 files. Here, we get 10 estimates per trial. It is comparable to TW=5 in the MT code above, which yields 9 estimates.
4. **getHilbertPairwiseMeasures** – This program computes and saves the single-trial estimates of shuffle corrected and uncorrected LFP power correlations and PPC using the Hilbert transform method. Here correlation and PPC are calculated across time points. Please run this program as *getHilbertPairwiseMeasures(5)* - where 5 is the half of frequency bandwidth of the bandpass filter used to filter LFP signal - to get ElectrodePairwiseMeasuresHTW5 file saved.

Once all files are saved, we can use the display codes to generate the plots.

## III – Display Codes

These set of programs are used to display the data. They use the locally saved data from *savedData* folder obtained from the codes listed above. If these saved files are not available, these programs generate the saved files by running the analysis codes above.

1. ***displayResultsSingleElectrode:*** Displays power computed using Multitaper (MT) method for 4 comparisons: (AttIn vs AttOut)Hits, (AttIn vs AttOut)Misses, (Hits vs Misses) AttIn and (Hits vs Misses) AttOut. Options are provided to choose condition (Valid, Neutral or Invalid), and whether target onset times should be matched. This program plots the figure which displays the PSDs, firing rate and dPrime averaged across electrodes. For example, displayResultsSingleElectrode(‘V’,1,10,5) displays the results for Valid condition, no target onset matching, TW=5. For matched target onset times, run displayResultsSingleElectrode(‘V’,3,10,5). Replacing ‘V’ with ‘N’ displays the plots for the Neutral condition.
2. **displayResultsMTPhase:** Displays the PPC computed for absolute phases across trials for individual electrodes. This is not a typical analysis since usually PPC is computed for phase difference between electrode pairs (as opposed to a single electrode). The idea here is that if the absolute phases have a specific relationship with target onset time, it should show a larger PPC value. Indeed, PPC is larger at SSVEP frequencies (20 Hz and harmonics). displayResultsMTPhase('V',3,10,5) or displayResultsMTPhase('N',3,10,5) yield plots similar to the one above (without dPrime analysis since PPC is computed across trials). These results look better with fewer tapers since SSVEP peaks are more prominent.
3. **displayMeanResultsElectrodePairs:** This shows results of pairwise analysis between all simultaneously collected electrode pairs from each hemisphere. For power, correlation is computed across trials. For phase, PPC of the phase difference is computed across trials. Spike count correlation across trials is also computed. These pairwise data are also saved since they take a long time to generate. This program can be run as displayMeanResultsElectrodePairs('V',3,10,5,'power') for valid, target-onset matched condition with TW=5. Replace ‘power’ with ‘phase’ to get PPC. One thing to note that when target-onset times are matched, only a subset of trials are used, so the results vary slightly every time the program is run if the pairwise data are not saved. A better strategy (not implemented yet) is to get the final correlation and PPC values for multiple iterations, take the average mean and average SEM across iterations, and plot that.
4. **displayResultsElectrodePairs**: This shows results of single trial pairwise analysis between all simultaneously collected electrode pairs from each hemisphere. There is an option to either use FFT estimates or MT estimates. For FFT, each trial is segmented into numDivisions segments, and one estimate of FFT power and phase is computed. Firing rate is also computed for each segment. Then, spike count correlation, power correlation, and PPC is computed across the segments to yield single trial estimates of each. The program shows the difference in means as well as dPrimes for different comparisons. When MT estimates are used, power correlation and PPC are estimated across tapers to again yield single trial estimates. Firing rate correlation is not computed in this case.
5. **displayResultsShuffleCorrectedCorrelations:** This shows the results of the FFT power and firing rate correlations computed across segments after shuffle correction to control for factors like adaptation. Correlations shown in displayResultsElectrodePairs are the raw correlations without any correction. This program can be run as displayResultsElectrodePairsHilbert ('V',3,10,10) for valid conditions, target onset matched, a trial cut-off of 10 and 10 segments or bins.
6. **displayResultsElectrodePairsHilbert:** This program displays the results of single-trial estimates of pairwise measures calculated using Hilbert transform. This program can be run as displayResultsElectrodePairsHilbert ('V',5,3,10) for valid conditions, half-frequency bandwidth of 5, target onset matched and trial cut-off of 10. This program calls getHilbertPairwiseMeasures.

All the display codes show the results for one random sample of trials, which match the target-onset distribution of hits and miss trials of each attention condition separately. They also output the mean differences and d-prime values and these codes need to be run for multiple iterations to sufficiently sample the trials to match the target-onset time distributions and then average across the iteration to obtain the final results. While running these codes iteratively please comment out the saving option as well as loading the already saved data.

## Other codes used in this project

**Display**

1. ***getPlotHandles.m*** – creates subplot layout for the figure.
2. ***makeDirectory*** – generic program to create a directory if it is not present already.

**Analysis**

1. ***getGoodStimNums*** – This program is used for mean matching the target onset times.
2. ***getDPrime –*** Computes the dPrime ((mean1-mean2)/(avg std dev)) between two arrays of numbers
3. ***getPPC*** – Takes phase values (in radians) as input and returns the pairwise phase coherence
4. **mtspectrumc\_returnJ** – modified from mtspectrumc in Chronux. This code additionally returns the amplitude and phase of individual tapers.