**EEG RECORDINGS HANDLING AND ANALYSIS HOW-TO**

This how-to aims at describing how:

1. Multi-channel iEEG recordings from the Johns Hopkins Epilepsy Center (JHU-EC) are stored in our workstation;
2. Power-, and coherence-based adjacency matrices are stored in our workstation;
3. Singular values and vectors (both right and left) are stored in our workstation.

**Root folder:** recordings and results of computation are stored in external hard drives, under the folders **/media/ExtHDD01**, **/media/ExtHDD02**, **/media/ExtHDD03**, and **/media/ExtHDD04**.

**Folder tree:** there is a folder for each data set and a sub-folder for each recording session. The folder name is **PY##N@@@**, where **##** is the year (last 2 digits) when the recordings have been collected and **@@@** is the ID of the patient. This notation comes from JHU-EC. Please, don’t change it in the common folder!

**Raw iEEG recordings:** multi-channel recordings are sampled at 1kHz and stored into \*.rec files. Data type is “short” (i.e., 2 bytes per recorded sample) and byte order is “big endian”. For each recording session, there is a folder with several (consecutive) \*.rec files and a text file named **“eeg.lst”**. The text file explains the content of each \*.rec file. If you need to read raw data from a \*.rec file in the MATLAB environment, please type the following:

% set the path to the folder where the \*.rec file is

**cd** /media/ExtHDD…/PY… %(state the full path)

% read the data-specific information from a \*.hdr file

hdr = **readhdr**('eeg.hdr');

% extract the number of recording channels

num\_channels = hdr.file\_fmt.Numb\_chans;

% extract the format of the data in the \*.rec file

format\_file = hdr.file\_fmt.File\_format;

% extract the offset from the beginning of the \*.rec file

offset = hdr.file\_fmt.Data\_offset;

% read the data

fid = fopen('filename','rb'); % replace filename with the name of

% the \*.rec file to be opened

fseek(fid,offset,'bof');

data = fread(fid,[num\_channels inf],format\_file);

fclose(fid)

The MATLAB function **readhdr.m** is in the folder of utility software on the PBWORKS website.

**Adjacency matrices:** For each \*.rec file, 12 types of adjacency matrices are computed. They are all stored in the sub-folder **adj\_matrices/** located in the same folder where the \*.rec file is. The adjacency matrices depend on:

* Which statistic has been used (power spectrum or coherence);
* Which frequency band has been considered (there are: “theta”, “delta”, “alpha”, “beta”, “gamma”, and “high” thus far, where “high” covers the range [90,200) Hz and is stored for future analyses). **NOTE: *in order to save space on the workstation, matrices and vectors computed for the “theta” and “delta” frequency band are stored in the cluster but not in the workstation*.**

For each combination of statistic and frequency band, there is a sequence of adjacency matrices, one per every second of multi-channel recordings stored in the \*.rec file, and all the matrices in the sequence are stored in the same \*.dat file. See the cartoon below, where:

**filename.rec**

***time***

***adj matrices***

**adj\_xxx\_filename\_fff.dat**

**xxx =** type of statistic (it can be either **pwr** for power or **chr** for coherence)

**fff =** is the frequency band (as above).

If you need to open the adjacency matrices in the MATLAB environment, please use the function **read\_adj\_matrix.m**, which is in the folder of utility software on the PBWORKS website. Code is commented and endowed with a help page.

**Singular Value Decomposition:** For each adjacency matrix, the singular value decomposition (SVD) is computed and the correspondent output (i.e., singular values *S* and left singular vectors *U*) are stored. In particular, the singular values *S* are stored in a \*.dat file in the sub-folder **svd\_values/**, while the singular vectors *U* are stored in \*.dat files in the sub-folder **svd\_vectors/**. The name of the \*.dat files follow the following rule:

|  |  |  |
| --- | --- | --- |
| **ADJ MATRIX FILE** | **SVD FILES** | **LEGEND** |
| adj\_xxx\_filename\_fff.dat | svd\_l\_xxx\_filename\_fff.dat  svd\_v\_xxx\_filename\_fff.dat | xxx = type of statistic  fff = frequency band  l = left singular vectors  v = singular values |

If you need to open the SVD matrices in the MATLAB environment, please use the functions **read\_svd\_array.m** (for values *S*) and **read\_svd\_matrix.m** (for vectors *U*). Both functions are in the folder of utility software on the PBWORKS website. Code is commented and endowed with a help page.

**SEIZURE EVENTS AND TEMPORAL CONSTRAINTS**

In the folder “JHU-EC Dataset” on the PBWORKS website is now available **“datainfo.zip”**. It contains two .mat files which store information about the temporal constraints and annotated seizures in the iEEG recordings from the Johns Hopkins Epilepsy Center. Here the details:

**infoevent.mat** – It contains a MATLAB struct for each subject in the dataset. Each struct has name **eventPY##N@@@**, where **##** is the year (last 2 digits) when the recordings have been collected and **@@@** is the ID of the patient. Each struct has the following fields:

1. **code**. It is a vector of *N* integer values, where *N* is the number of annotated events. Each value is the code associated with the correspondent event (i.e., value in position **#*k*** corresponds to the ***k*-**th event;
2. **desc**. It is an array of *N* cells. Each cell stores the description associated with the correspondent event code;
3. **dur.** It is a vector of *N* integer values. Each value is the duration (in sec) of the correspondent event;
4. **time.** It is a *N*x2 array of real positive values. Each row reports the onset (first column) and termination (second column) of the correspondent event. Onset and termination time are in sec.

**infotime.mat** – It contains a MATLAB struct for each subject in the dataset. Each struct has name **PY##N@@@**, where **##** is the year (last 2 digits) when the recordings have been collected and **@@@** is the ID of the patient. Each struct has the following fields:

1. **filename**. It is an array of *N* cells. Each cell stores the sub-directory and the name of the original file of iEEG recordings;
2. **time.** It is a *N*x2 array of real positive values. Each row reports the onset (first column) and termination (second column) of the recordings stored in the correspondent file in **filename**. For example, the row **#*k*** of **time** stores the onset and end time of the recordings stored in the file whose name is in **filename(*k*)**. Onset and termination time are in sec.