Notes on the **source code** for

‘*Large-scale brain modes reorganize between infant sleep states and carry prognostic information for preterms*’

General description

Please, also check the 'Analytical pipeline.pdf' file to get idea about all analytical steps and toolboxes (incl. their download links) that were used in this work.

The source code includes three stand-alone Matlab functions:

compute_AAC.m

This function takes as an input multichannel EEG data, filters this data into N = 4 (or 21) frequency bands, computes parcel signals using infant head model (see Tokariev et al., 2019 in Cerebral Cortex), orthogonalizes parcel signals (see Brookes et al., 2012 in NeuroImage) and computes functional connectivity as a pairwise correlation between amplitude envelopes of these signals (AAC = amplitude-amplitude correlation).

With this script we also provide an example of **input** data: infant_eeg.mat (folder 'Data example')

infant_eeg.mat contains a 19-channel EEG (Fs = 100 Hz, average montage).

We also provide pre-computed data needed for source modelling and parcellation (folder 'Head Model').

Folder 'Filters' contains 2 sets of band-pass filters: filters_eeg_4_bands.mat (N = 4: delta, theta, alpha, beta) and filters_eeg_21_bands.mat (N = 21: extended set that was used to assess effect size in frequency domain). Note, that band-pass filters are implemented as a combination of high-pass (top row) and low-pass (bottom row) filters. We used Matlab function designfilt to generate all filters. File filters_amp_15_bands.mat contains filters (N = 15) for amplitude envelopes.

Output. Functional connectivity computed from broadband amplitude envelopes (non-filtered) is stored to the variable AAC_bb (bb = broadband) that is the cell array with dimensions [1 x N of EEG filters] and saved to the folder 'AAC_output' as AAC_bb.mat. Functional connectivity computed from band-pass filtered amplitude envelopes is stored to the variable AAC_nb (nb = narrowband) which is a cell array with the dimensions [N of amplitude filters x N of EEG filters] and saved to the folder 'AAC_output' as AAC_nb.mat. Each cell in AAC_bb or AAC_nb stores symmetric functional connectivity matrix with the dimensions [N of parcels x N of parcels].

compute_Model.m

This script simulated modeled brain activity (for different combination of mode weights) and computes corresponding connectivity matrix and its mean and std (needed in the cost function to find the best fit of the model to real data). User has to define the parameter grid in the *'Eigenmode coefficients'* section of the script.

Output. Script will generate mean (FCmodel_mean) and std (FCmodel_std) values of connectivity matrices computed for all possible combinations of mode weights (params). These variables will be saved to model_params.mat file (into 'Model parameters' folder).

fit_model_to_data.m

This script uses FCmodel_mean and FCmodel_std data to find the best fit between modeled activity and empirical data (stored in folder 'Fit model to data').

Output. Mode weights (paramData = [a1 a3 a4]) that correspond to the best fit (or minimum of the cost function; see section **'Demo'**).

System requirements

These are stand-alone Matlab scripts. All data needed to run the script are provided with this package. The scripts were developed using Matlab (version R2016b) computing environment (developed by MathWorks).

All scripts were tested in Matlab versions R2016b and R2018a and with OS Windows 7 (64 bit), Windows 10, Ubuntu (version 16.04).

Any non-standard hardware is not needed.

Installation

Scripts don't require any special installation. To run the scripts, users need to install Matlab software on their computer and copy the whole folder with all files and sub-folders on their computer.

Add main folder with the scripts to Matlab path (File > Set Path > Add Folder)

Demo

Instructions to run on data and expected output

In the folder 'Data example' users can find the file infant_eeg.mat that contains 19-channel EEG data (similar data was used in the original paper).

First, user have to run compute_AAC.m script first that will read the data file and compute functional connectivity for cortical signals (broadband and narrowband). The resultant connectivity matrices will be saved to 'AAC_output'. There are 2 options to use: a) standard (N = 4) and; b) extended (N = 21) set of filters.

Second, to generate the biophysical model, user have to run compute_Model.m script. This script computes synthetic connectivity data for the model with all possible weighted combinations of cortical modes (mode 1, mode 3, and mode 4; see original paper). Script outputs mean and STD of each connectivity matrix those fill be used further as a variables of the cost function in order to find best fit of the model to empirical data (see description of fit_model_to_data.m script below).

Range of mode weights and the step have to be defined in *'Eigenmode coefficients'* section:

aj = linespace(min value of aj, max value of aj, number of equidistant steps)

Note, that by default, script will use very detailed grid of parameters (aj) such was used in the paper (it has more than 8 million of parameter combinations). It might take few days to compute data for all possible combinations. To reduce computational time, user might consider decreasing the number of steps or using more narrow diapason of weights. Output data for original grid can be obtained from authors by request (file size is 114 Mb). With this package we provide the fragment (to keep the file size reasonable) of the original output data that includes best fit to the test data (see model_params.mat file in the folder 'Model parameters'). The compute_Model.m script generates two variables: FCmodel_mean (mean values of all synthetic connectivity matrices) and FCmodel_std (std for the corresponding matrices). These variables (together with tested coefficients params) are saved to infant_eeg.mat (folder 'Model parameters').

Third, user have to run fit_model_to_data.m script to find the best fit of the model to empirical data. As an input, this script will need: a) real connectivity matrix (that is stored in Connectivity_data_alpha.mat file, which is in the folder 'Fit model to data'); and b) mean/std data of simulated connectivity data and related weights (these data are stored in model_params.mat that is in 'Model parameters' folder). This script is searching for the minimum of the cost function that is:

$$\text{cost function} = (\text{mean of Model} - \text{mean of Data})^2 + (\text{std of Model} - \text{std of Data})^2$$

This script returns paramData vector that stores mode weights [a_1 a_3 a_4] for the best fit (see paper for details).

Run time

(estimated using laptop with the following parameters: processor 2.20 GHz, 16 Gb RAM, Windows 7 64 bit)

compute_AAC.m
For N = 4 EEG filters, run time is about 9.5 minutes
For N = 21 EEG filters, run time is about 47 minutes

compute_Model.m
Computational time depends on the density of the parameter grid (a_1 a_3 a_4). To simulate 1 synthetic matrix (using one combination of parameters a1, a3, a4) it takes about 0.045 sec. Note, that when the parameter grid is big (for example, in the original paper we tested more than 8 million combinations), it might take few days to compute all combinations. For such cases, we recommend using powerful computational servers.

fit_model_to_data.m
To fit the model to 1 subject, it takes about 3.5 sec

Instructions for use

Instructions to run the scripts on the test data (Data example\infant_eeg.mat) are identical to that described in **'Demo'** section. We also provide with this package all precomputed intermediate data for the test file.