**Supplementary Materials**

**S1. DEEP User Manual**

Contents

[Overview 2](#_Toc74648663)

[Setup and Run (DEEP\_main) 5](#_Toc74648664)

[[1] Data Import (DEEP\_main\_1) 7](#_Toc74648665)

[[2] Preproc I: Bad Channel Detection, Filtering (DEEP\_main\_2) 9](#_Toc74648666)

[[3] Cleaning prior to ICA (DEEP\_main\_3) 11](#_Toc74648667)

[[4] Preproc II: ICA-based Artifact Correction, Bad Channel Recovery, Re-referencing (DEEP\_main\_4) 12](#_Toc74648668)

[[5] Automatic and Manual Artifact Detection (DEEP\_main\_5) 14](#_Toc74648669)

[[6] Narrow Band Filtering and Hilbert Transform (DEEP\_main\_6) 16](#_Toc74648670)

[[7] Estimation of Phase Locking Values (PLV and cross-frequency PLV) (DEEP\_main\_7) 18](#_Toc74648671)

[[8] Power analysis (TFR, pWelch) (DEEP\_main\_8) 19](#_Toc74648672)

[[9] Averaging over dyads (DEEP\_main\_9) 20](#_Toc74648673)

# Overview

The pipeline was initially designed to analyze data from a mother-infant hyperscanning study conducted in our lab. We reasoned that more researchers could benefit from the pipeline, and started the process of adjusting the functionalities of the pipeline to address the needs of a general user. Because this is work in progress, in the current version of the pipeline, users would need to adjust the “DEEP\_generalDefinitions.mat” script to be able to proceed with the pipeline. This script includes information specific to our study such as the EEG markers used to indicate the events or the duration of the conditions. In Table S1, we present information on the DEEP\_generalDefinitions.mat file used to analyze the data in the paper.

Table S1: The attributes of the DEEP\_generalDefinitions.mat file that are relevant to the analyses reported in the paper.

|  |  |  |
| --- | --- | --- |
| ***Attributes*** | ***Content*** | ***Description*** |
| artfctMark: | {'S 4' 'S 5' 'S 3'} | Markers indicating periods that will be deleted from the analyses such as breaks. S4: Pause, S5: Resume, S3: Quit |
| artfctNum: | [4 5 3] | Numeric form of artfct |
| artfctString: | {3x1 cell} | Cell array of string, artfct |
| condMark: | {'S 11' 'S 13'} | Markers indicating the conditions analyzed in the current paper. S11: Free play condition, S13: Resting state condition Please note that the original experiment includes other conditions, which are not relevant to the analyses reported in this paper. |
| condMarkDual: | {'S 11' 'S 13'} | Dual condition markers used for the analyses of the example dataset. |
| condNum: | [11 13] | Numeric form of condMark |
| condNumDual: | [11 13] | Numeric form of condMarkDual |
| condString: | {2x1 cell} | Cell array of string, cond |
| duration: | [150 45] | The duration of the free play and resting state conditions, respectively, in seconds. |
| trialNum1sec: | [150 45] | The total number of trials in the free play and resting state conditions, respectively, in one second epochs. |
| trialNum5sec: | [30 9] | The total number of trials in the free play and resting state conditions, respectively, in five second epochs. |

The pipeline is based on functions of the FieldTrip toolbox (<https://www.fieldtriptoolbox.org/>). Users are expected to download the FieldTrip toolbox independently. Once it is downloaded and set up correctly (see startup.m), one can type ***“ft\_version”*** in the MATLAB command window to get the latest version of the FieldTrip toolbox. One can also clone and use a certain version of the FieldTrip toolbox to operate the pipeline.

In order to operate the pipeline and analyze the data, each researcher can create their own session. In one session, data for each processing step is saved. Several sessions for the same study can operate in parallel without interfering with one another. Thus, several researchers can work simultaneously on the data analysis. Preprocessing steps completed in one session can be copied to another session using the ***“DEEP\_cloneSession”*** script. Within each session, the user can choose from several processing steps. Whereas preprocessing steps 1 to 6 have to be done in a fixed order, users can proceed with the subsequent steps (e.g., steps 7 or 8) depending on the analysis of their choice.

The user launches the pipeline by typing the “***DEEP\_main***” command in the MATLAB command window. Following this, the user is asked to select the default path or define a new path on their computer where the raw data are stored. The processed data and the final exported data will also be saved under this path. Next, the user chooses the participants’ data to be processed among three options 1) all available dyads 2) all new dyads 3) a specific subsample of dyads. This gives flexibility to users especially when they want to exclude some of the participants’ data from the start or to perform the processing steps with a subsample of participants reducing preprocessing time and effort.

After the dyads are selected, the pipeline allows the user to choose the processing steps that are available for the imported data files (e.g., if no data was imported for a certain participant, no later processing steps can be selected). This function also helps the user to remember which preprocessing steps are completed for which dyads. In the current version of the pipeline, the same preprocessing steps are applied to both the mother and the infant data. For each preprocessing step, first, the data of the adult is processed followed by the data of the infant/child. When processing data of two adult participants, the pipeline first processes data of participant 1 followed by data of participant 2. After each processing step, data files are restored separately as \*.mat files. Below, we will illustrate each processing step in detail and introduce the main functions used at each step.

# Setup and Run (DEEP\_main)

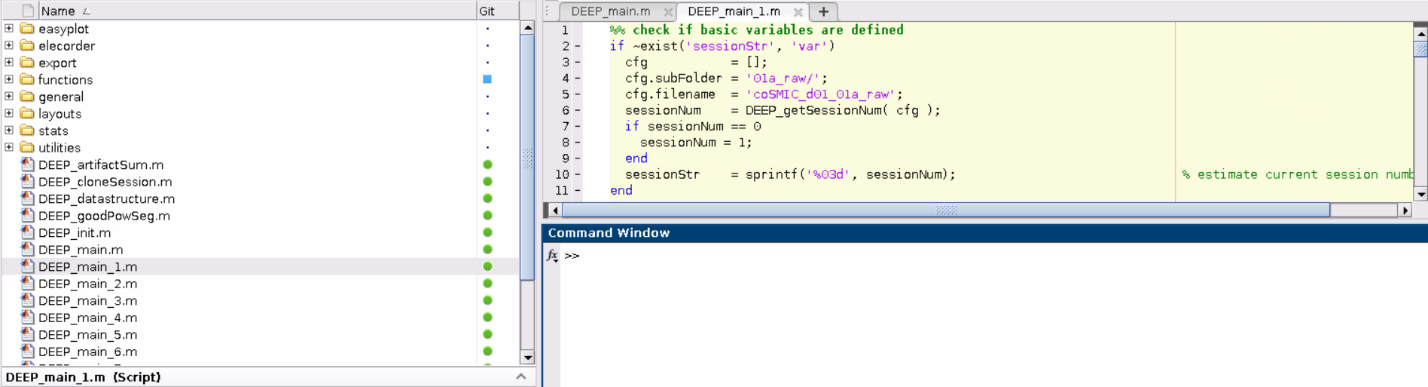
After downloading the project files from the GitHub repository (<https://github.com/dmatthes1982/MPI_CBS-DEEP>), run MATLAB. After opening MATLAB, make sure that you are working with the right directory. Otherwise, switch to "…/DEEP\_eeg\_pipeline directory". Check the following:

**>> pwd**

**ans =**

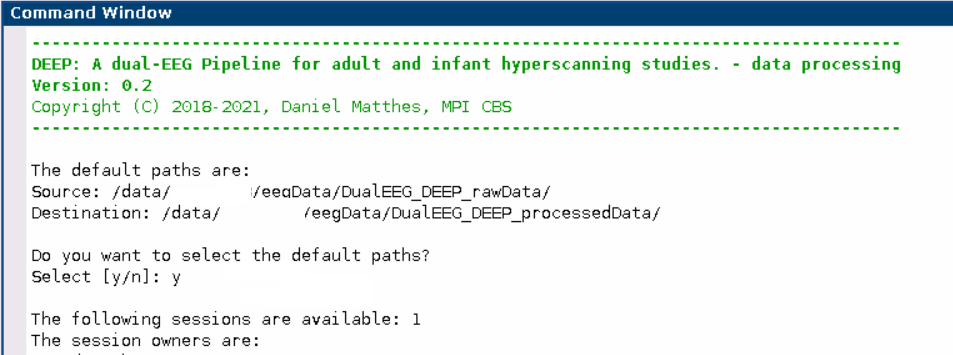
**'/data/UserName/MATLAB/scripts/ DEEP/DEEP\_eeg\_pipeline'**

DEEP scripts should be ready to run in the MATLAB command window when you are in the right directory. If not, check if you have successfully installed the FieldTrip toolbox and the project files.

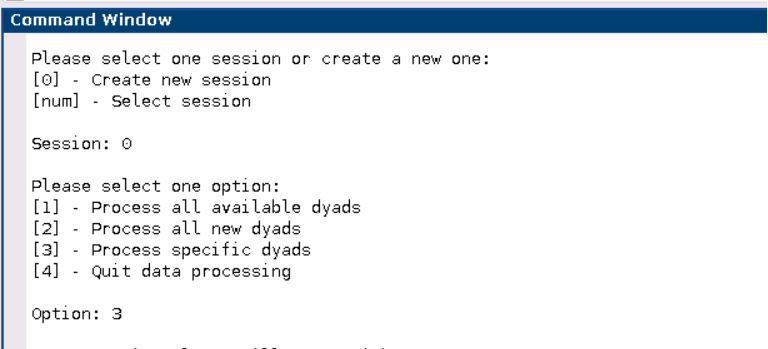


To launch the pipeline, type the following in the command window in MATLAB:

**>> DEEP\_main**



This command runs the “***DEEP\_main”*** script. “***DEEP\_main”*** selects sessions, data processing steps and dyads. Select ***"y"*** for default paths. Select ***"n"*** to add a new path. Enter a previous session number (i.e., enter session number such as 1 or 2) or create a new one (i.e., enter 0). Please note that each session data will be saved separately. A user can create more than one session to operate them simultaneously, if needed. You have to select the same number to continue your previous work in that session. **Importantly**, if a previously completed step is rerun within the same session, previous data will be overwritten. In the default path, the pipeline generates a “00\_Settings” file that lists the parameters and user choices for preprocessing within a particular session.



For illustration purposes, we will process a few dyads. Select **"3"** to process a specific dyad.

**DEEP\_main functions:**

* ***DEEP\_init:*** Adds directory and subfolders to the path.
* ***DEEP\_main\_1 to DEEP\_main\_9:*** Subscripts for data processing.

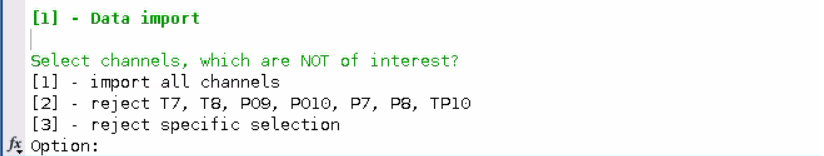
Next, we will explain each main data processing step and its associated functions in detail.

# [1] Data Import (DEEP\_main\_1)

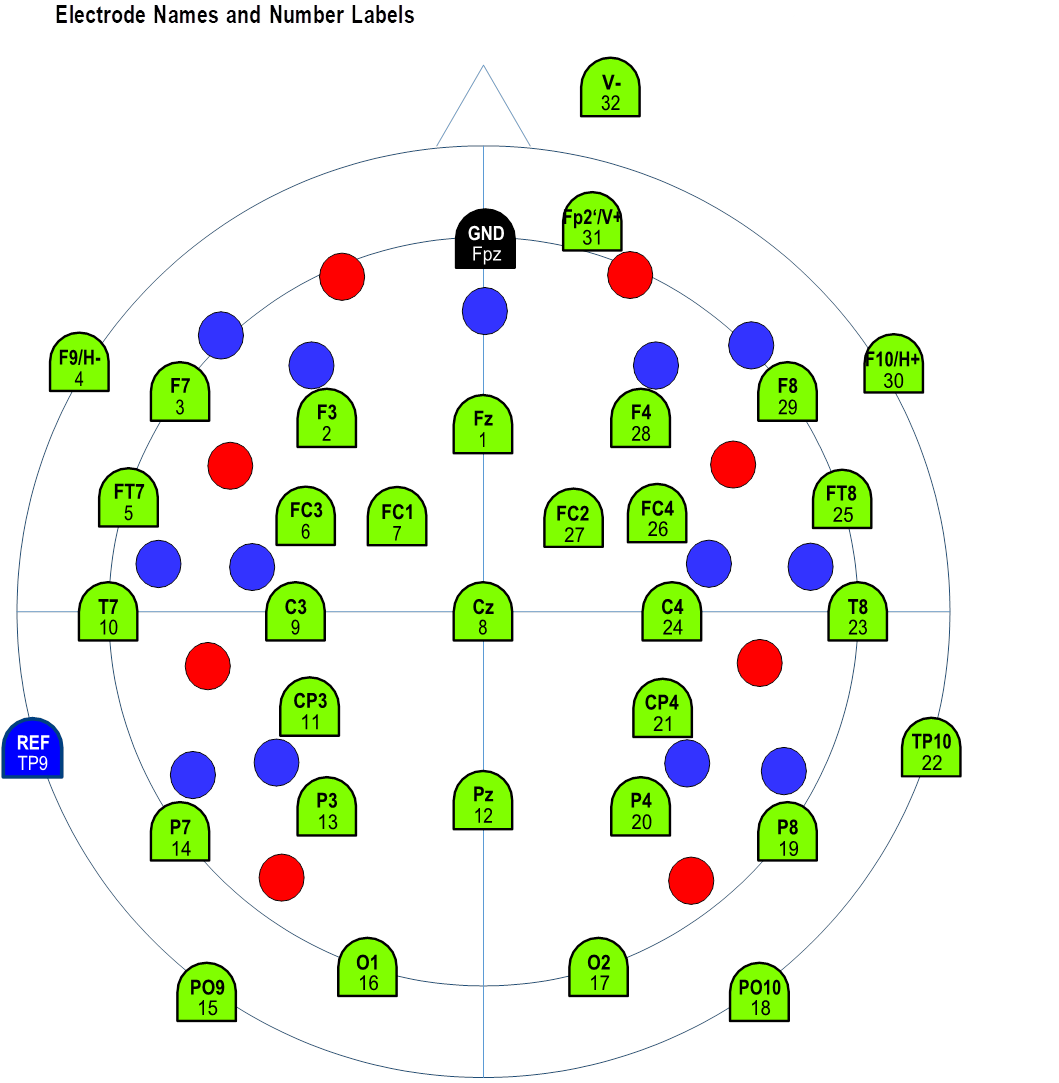
Select **"1"** to import data.



Then, select the dyads listed in the pop-up window (e.g., 25). This selection runs the “***DEEP\_main\_1”***script. “***DEEP\_main\_1”*** imports EEG files from the source location. The user is asked to choose the electrodes to include in the data import. Select ***"1”*** to import all of the channels. Select ***"3"*** to exclude a specific set of channels (e.g., the ones on the periphery). In the pop-up window, choose the channels you would like to exclude from the start. Click **"OK."**



As shown below, we used a customized channel configuration (*EASYCAP GmbH.*, Germany) to record the data presented in the paper. Green points illustrate the EEG channel names and numbers. “V” and “H” refer to vertical and horizontal eye electrodes, respectively. Blue and red points represent optode locations for functional near-infrared spectroscopy (fNIRS), which were kept empty during the recordings of the data used in the paper, as we only measured dual EEG in this experiment.

****

**Important note:** In the current version of the pipeline, only data recorded with the Brain Products system can be analyzed. If you are using another system, you can convert your data and process it outside of the GUI. In this link (https://github.com/dmatthes1982/MPI\_CBS-DEEP/tree/master/dataconversion), we provide two example scripts: 1) to convert data obtained with a BioSemi system (i.e., in *.bdf* format) 2) to call DEEP sub-functions to operate the pipeline afterwards. Users can open the “***DEEP\_importBdf”*** function and copy their dataset into a folder on their computers and change the directory to the destination of the dataset. Further instructions can be found in the scripts.

**DEEP\_main\_1 functions:**

* ***DEEP\_getSessionNum:*** Estimates current session number.
* ***DEEP\_createTbl:*** Creates settings file.
* ***DEEP\_importDataset:*** Imports the data of each dyad.
* ***DEEP\_saveData:*** Stores the data in a *.mat* file.

**Additional functions:**

***DEEP\_loadData:*** Loads data of a previous step.

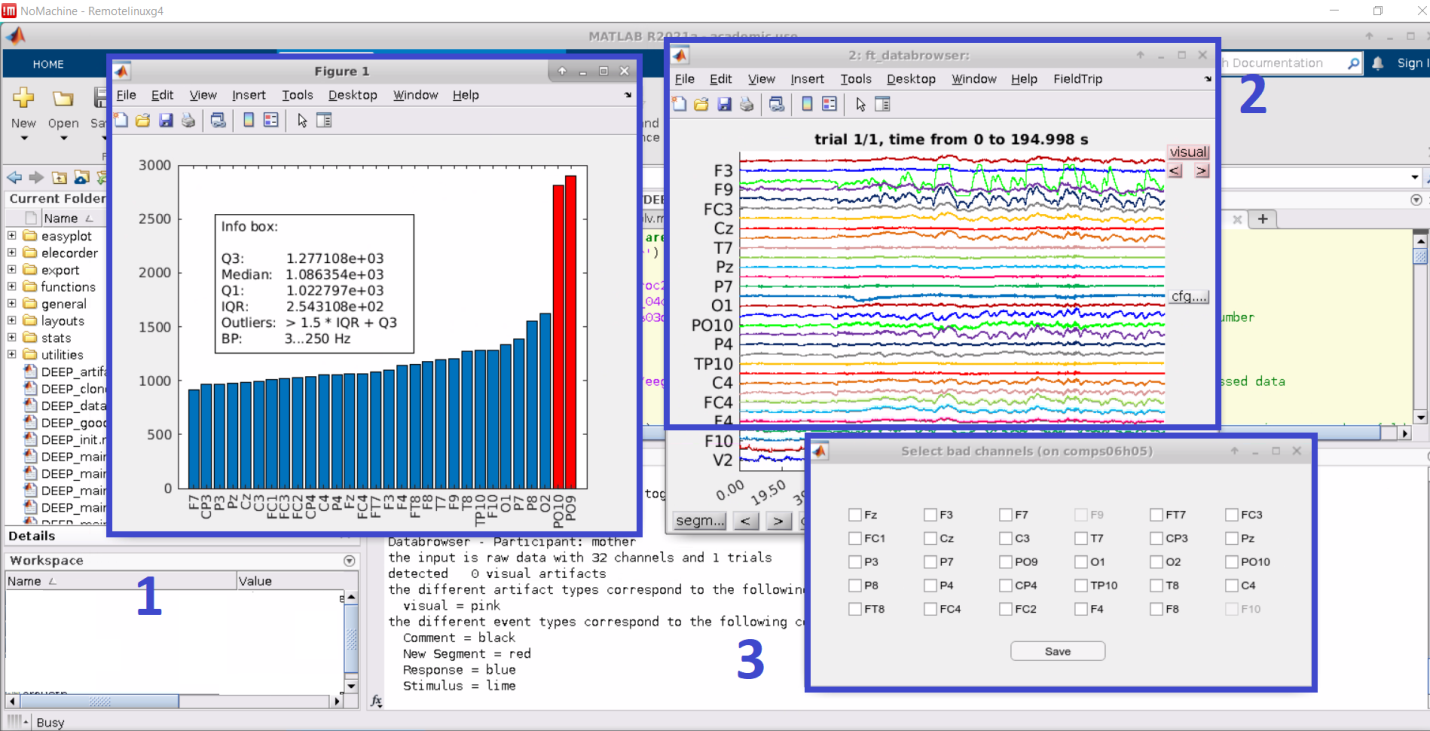
***DEEP\_selectData:*** Selects data.

Some of these basic functions are used in the upcoming steps as well. To avoid redundancy, we only describe them here.

# [2] Preproc I: Bad Channel Detection, Filtering (DEEP\_main\_2)

Continue data processing by selecting ***"y"***. If you have stopped the processing after step 1, select step ***"2"*** to proceed. This selection runs the “***DEEP\_main\_2”*** script. “***DEEP\_main\_2”*** has two sub-steps:1) selection of bad/noisy channels 2) applying a basic bandpass filter for an initial cleaning of channels. By default, one can mark a maximum of 10 % of the total channels as noisy. If needed, this limit can be deactivated in the code. Three pop-up windows appear to help you choose the channels to be interpolated:

* Figure 1 shows the total power of each channel and highlights outliers by marking them red. Channels are marked as bad/noisy channels when their total power from 3Hz (to suppress eye artifacts and baseline drifts) is below Q1 - 1.5 \* IQR or above Q3 + 1.5 \* IQR.
* Figure 2 uses the ft\_databrowser function of FieldTrip to plot the raw signal for each channel across time.
* Figure 3 shows the pop-up window for selecting the channels to interpolate.



Select bad/noisy channels that you would interpolate later and click "**Save."** When chosing the channels to be interpolated, we recommend the users to check the neighboring channels that are used for interpolation (for more information on neigboring channels, please see https://github.com/fieldtrip/fieldtrip/tree/master/template/neighbours). Note that horizontal and vertical electrooculograms (EOGs), namely EOGH and EOGV, cannot be interpolated, thus, they cannot be selected.

**DEEP\_main\_2 functions:**

* ***DEEP\_concatData:***Concatenates all trials of a dataset to a continuous data stream.
* ***DEEP\_estNoisyChan:*** Detects noisy channels automatically.
* ***DEEP\_selectBadChan:*** Selects corrupted channels.
* ***DEEP\_preprocessing:*** Applies a basic bandpass filter [1-48 Hz] to the raw data.

# [3] Cleaning prior to ICA (DEEP\_main\_3)

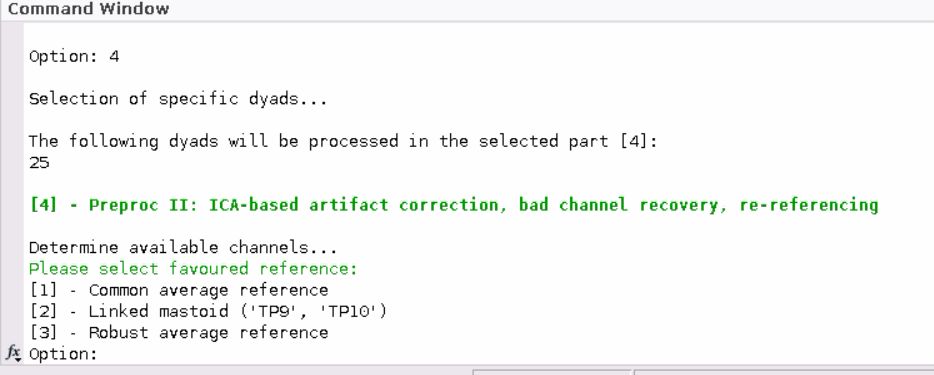
Continue data processing by selecting ***"y."*** If you have stopped the processing after step 2, select step ***"3"*** to proceed. This selection will run the “***DEEP\_main\_3”*** script. “***DEEP\_main\_3”*** (1) concatenates preprocessed trials into a continuous stream, (2) detects artifacts and rejects transient artifacts (criterion: within a 200 ms segment with a 50 % overlap [i.e., 100 ms], the difference between the maximum and the minimum voltage cannot be more than 200 μV), (3) concatenates cleaned data into a continuous stream, (4) runs ICA decomposition, (5) extracts EOG channels from the cleaned continuous data.

**DEEP\_main\_3 Functions:**

* ***DEEP\_concatData:*** Concatenates all of the preprocessed trials into a continuous data stream.
* ***DEEP\_autoArtifact:*** Detects transient artifacts given the criterion above.
* ***DEEP\_mergeThArtResults:*** Merges two artifact detection operations (i.e., based on all electrodes vs. eye electrodes) into a common one.
* ***DEEP\_rejectArtifacts:*** Rejects previously detected transient artifacts.
* ***DEEP\_ica:*** Conducts an independent component analysis (ICA) on the data using all channels except eye electrodes and the reference electrode.
* ***DEEP\_selecdata:*** Extracts EOG channels from the cleaned continuous data.

# [4] Preproc II: ICA-based Artifact Correction, Bad Channel Recovery, Re-referencing (DEEP\_main\_4)

Continue data processing by selecting ***"y."*** If you have stopped the processing after step 3, select step ***"4"*** to proceed. This selection will run the “***DEEP\_main\_4”*** script. “***DEEP\_main\_4”*** (1) finds ICA components for eye movements, which show a minimum of 80 % (i.e., default value) correlation with EOGV and EOGH, (2) verifies the estimated components by using the *ft\_icabrowser* function, and allows the user to add additional components to the selection, (3) corrects the EEG data, (4) recovers bad channels, and (5) applies an offline re-referencing method.

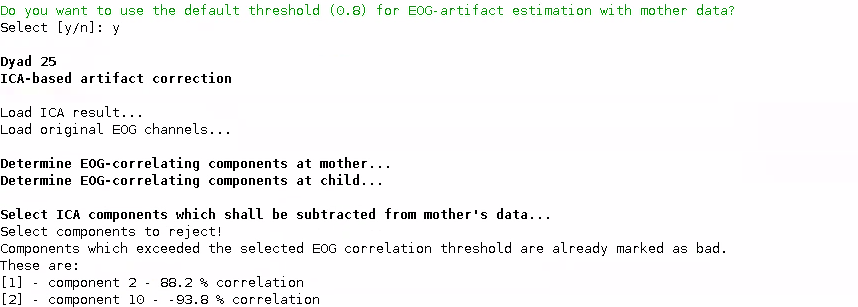


Select ***”1”*** for re-referencing using the common average method.

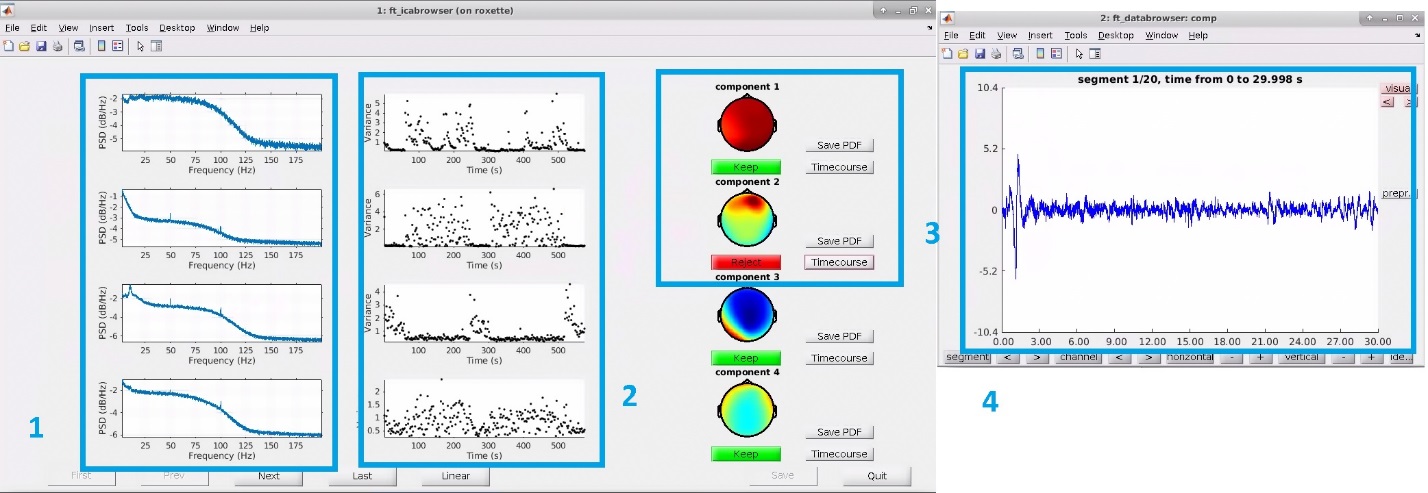
Select ***”2”*** for re-referencing using the linked mastoids method.

Select ***”3”***for re-referencing using the robust averaging method.

Next, select ***”y”*** to use the default threshold (i.e., 0.8) for the estimation of eye movement components. If you want to change the threshold value, select ***”n”***and define another value.



This selection opens the *ft\_icabrowser* window to inspect eye movement artifacts. The pipeline highlights detected components by marking them as **"Reject"**. This can be undone by clicking on the **"Reject"** tab once, which turns into **”Keep”**.As illustrated below,*ft\_icabrowser* (for more information, see <https://www.fieldtriptoolbox.org/reference/ft_icabrowser/>) presents a GUI showing the 1) power spectrum, 2) variance over time, 3) the topography of the components, and 4) time course of the rejected components. There is an option to save this information as a *.pdf* file. Click **"Quit"** after checking the components.

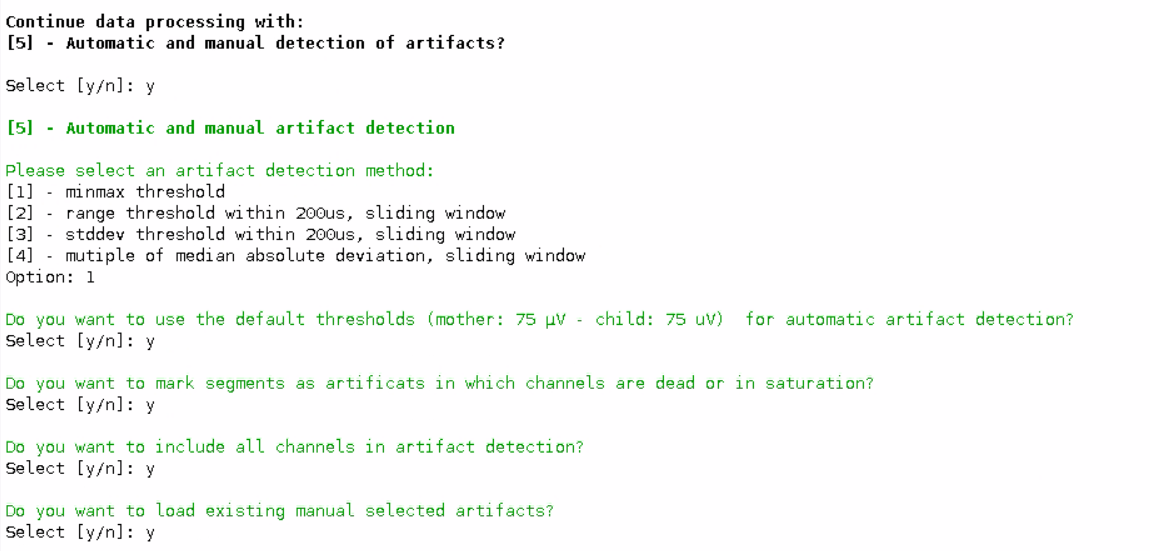
**

**DEEP\_main\_4 Functions:**

* ***DEEP\_detEOGComp:***Determines components that are highly correlated (> 80%) with the eye electrodes (i.e., EOGV and EOGH).
* ***DEEP\_selectBadComp:*** Allows the userto check estimated ICA components and change the predetermined components.
* ***DEEP\_correctSignals:*** Removes artifacts from the data.
* ***DEEP\_repairBadChan:*** Repairs bad channels that were selected at step 2.
* ***DEEP\_robustRef:*** Runs robust average re-referencing method.
* ***DEEP\_reref:*** Conducts re-referencing based on the selection (i.e., either common average or linked mastoids re-referencing methods).

# [5] Automatic and Manual Artifact Detection (DEEP\_main\_5)

Continue data processing by selecting ***"y."*** If you have stopped the processing after step 4, select step ***"5"*** to proceed. This selection will run the “***DEEP\_main\_5”*** script. “***DEEP\_main\_5”***runs(1) automatic artifact detection, (2) manual artifact detection (verification). You can choose among the four automatic artifact detection methods and adjust the thresholds manually:



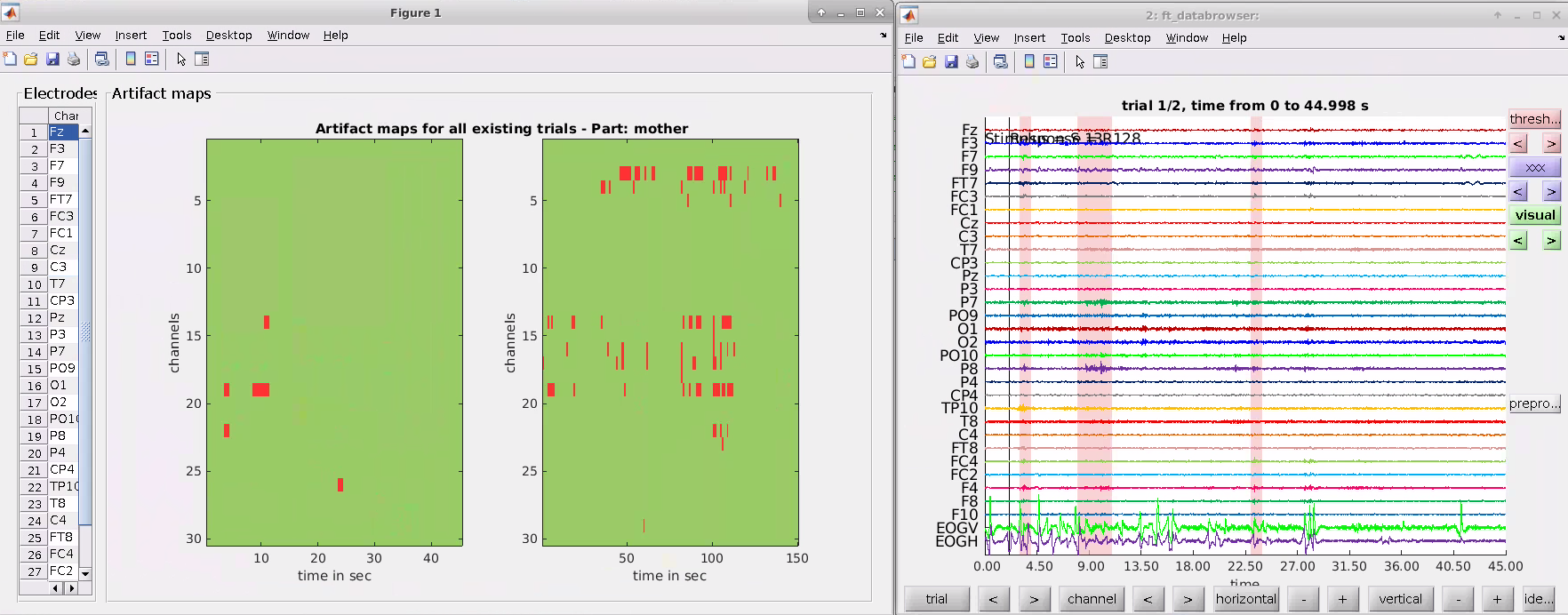
Select ***"1"*** for the min-max threshold (default: 'minmax', +-75 µV).

Select ***"2"*** for the range threshold within 200ms, sliding window.

Select ***"3"*** for the std dev threshold within 200ms, sliding window.

Select ***"4"*** for the multiple of median absolute deviation, sliding window.

Select ***"y"*** to use the default threshold of +-75 µV. Select ***"y"***to mark dead channels. Select ***“y”*** to include all channels in artifact detection. Select ***"y"*** to load artifacts that were selected manually at step 2. If you haven’t marked any artifacts manually, type ***“n”***.



This selection opens artifact maps for each condition across time and channels as well as the *ft\_databrowser* window to inspect artifacts. On the left side of the artifact maps, the number of detected artifacts for each channel is listed. In the *ft\_databrowser* window, segments that contain artifacts are marked red. Users can visually inspect the artifacts, mark certain segments as artifacts manually, or undo a previous selection. Close the tabs by clicking ***“x”*** after checking the artifacts.

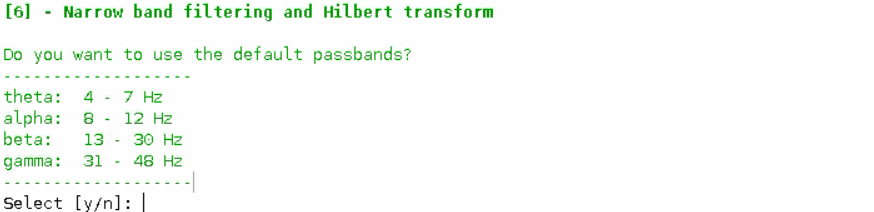
**Note:** Manually selected artifacts from a certain session can be added to another session by duplicating the files named *“\*\_allart\_00OldSessionNumber.mat”* from one session and changing the filename to *“\*\_allart\_00NewSessionNumber.mat”*. We recommend the users to use this functionality carefully to avoid potential errors because the file holds both the automatic and manually defined artifacts. It is only recommended to make this copy if the user wants to run step 5 of the pipeline afterwards. In that case, choose the option "import existing manually defined artifacts."

**DEEP\_main\_5 Functions:**

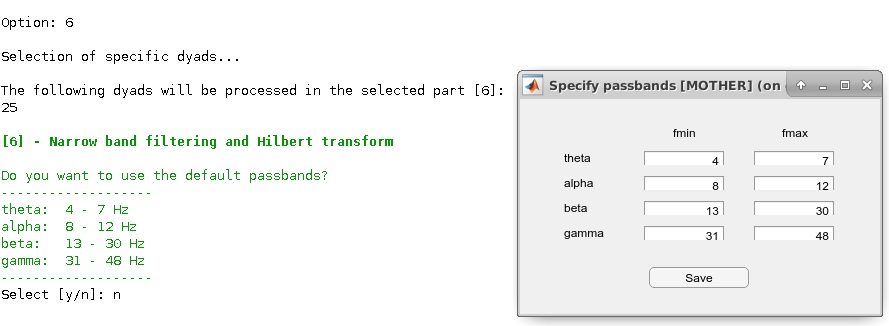
* ***DEEP\_autoArtifact:*** Marks segments as artifacts in which channels exceed a selected threshold based on the chosen method.
* ***DEEP\_manArtifact:***Verifies automatically detected artifacts and marks manually detected artifacts.

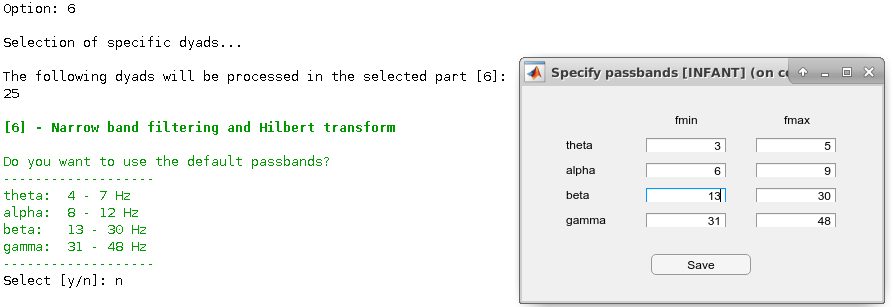
# [6] Narrow Band Filtering and Hilbert Transform (DEEP\_main\_6)

Continue data processing by selecting ***"y."*** If you have stopped the processing after step 5, select step ***"6"*** to proceed. This selection will run the “***DEEP\_main\_6”***script. “***DEEP\_main\_6”***handles narrow band filtering and Hilbert transformation.



Select ***"y"*** to use the default passbands. These passband selections are only recommended when analyzing dual EEG data with older participants such as two adults. To define separate frequency ranges for the adult and the infant/child data, type ***"n"*** and enter passband values manually. Two separate windows will open to define adult and infant passbands.



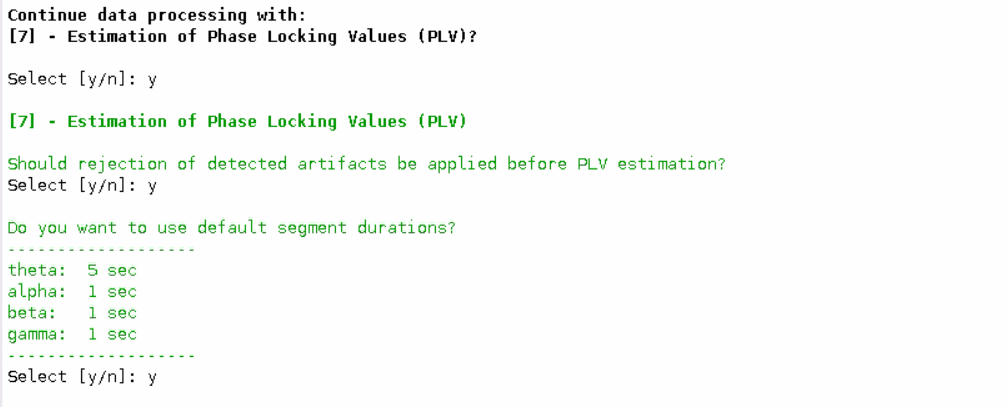


**DEEP\_main\_6 Functions:**

* ***DEEP\_pbSelectbox:*** Displays a GUI window for the specification of passbands. It returns a cell array including the minimum and maximum frequencies of each passband.
* ***DEEP\_bpFiltering:*** Applies a specified bandpass filter to every channel in the data.
* ***DEEP\_hilbertPhase:*** Estimates the Hilbert phase of every channel*.*

# [7] Estimation of Phase Locking Values (PLV) and cross-frequency PLVs (DEEP\_main\_7)

Continue data processing by selecting ***"y."*** If you have stopped the processing after step 6, select step ***"7*"** to proceed. This selection will run the “***DEEP\_main\_7”*** script. “***DEEP\_main\_7”*** (1) segments Hilbert phase data for PLV or cross-frequency PLV estimations and splits the data of every condition into sub-trials of 1 or 5 seconds epochs, (2) executes artifact rejection, (3) estimates PLVs or cross-frequency PLVs, (4) estimates mean PLVs or cross-frequency PLVs.



Select ***“y”*** to use the predefined epoch durations for each frequency band. Type ***“n”*** to define different epoch durations for each frequency band manually.

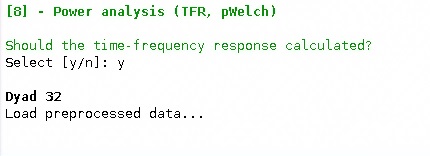
**Important note:** DEEP automatically proceeds with either PLV or cross-frequency PLV estimations depending on the passbands defined at step 6. If there are no differences between the passband values of the two participants, DEEP will run PLV calculations. If differences are detected between the passband values of the two participants, cross-frequency PLVs will be estimated.

**DEEP\_main\_7 Functions:**

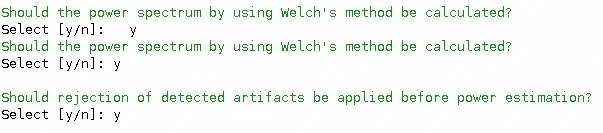
* ***DEEP\_segselectBox:***Specifies the duration of epochs for later segmentation of the data.
* ***DEEP\_Segmentation:*** Segments Hilbert phase data into epochs of predefined length.
* ***DEEP\_rejectArtifacts:*** Executes the removal of segments containing artifacts which were defined at earlier steps.
* ***DEEP\_crossPhaseLockVal:*** Calculates the cross-frequency PLVs for each channel combination in all of the conditions.
* ***DEEP\_phaseLockVal:*** Calculates the PLVs for each channel combination in all of the conditions.
* ***DEEP\_calcMeanPLV:*** Calculates the mean PLVs or cross-frequency PLVs for each condition.
* ***DEEP\_writeTbl****:* Exports the numbers of good trials used for PLV or cross-frequency PLV estimations into a spreadsheet.

# [8] Power analysis (TFR, pWelch) (DEEP\_main\_8)

Continue data processing by selecting ***"y."*** If you have stopped the processing after step 6, select step ***"8"*** to proceed. This selection will run the “***DEEP\_main\_8”*** script. “***DEEP\_main\_8”***calculates 1) time-frequency responses (TFRs), and 2) power spectral densities (PSDs) of the preprocessed data. This step is used to analyze EEG data for each participant individually.



Select **"y"** to calculate the TFRs.



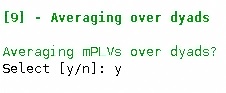
Select ***"y"*** to calculate the PSDs. Select ***"y"*** to reject artifacts that were defined at previous steps.

**DEEP\_main\_8 Functions:**

* ***DEEP\_timeFreqanalysis:*** Performs a time-frequency analysis (for more information, see https://www.fieldtriptoolbox.org/reference/ft\_freqanalysis/).
* ***DEEP\_segmentation:***Segments Hilbert phase data into epochs of predefined length for each condition.
* ***DEEP\_numOfSeg:*** Estimates the number of segments for each condition and participant.
* ***DEEP\_rejectArtifacts:*** Removes segments containing artifacts.
* ***DEEP\_pWelch*:** Calculates power activity using Welch's method.

# [9] Averaging over dyads (DEEP\_main\_9)

Continue data processing by selecting ***"y."*** If you have stopped the processing after step 7 or 8, select step ***"9"*** to proceed. This selection will run the “***DEEP\_main\_9”*** script. “***DEEP\_main\_9”*** averages PLVs, TFRs and PSDs across dyads or participants.



Select ***"y"*** to average PLVs or cross-frequency PLVs over dyads. Select dyads for averaging.



Select ***"y"*** to average TFRs over participants. Select ***"y"*** to average PSDs over participants. Select participants’ numbers to take the average of their data.

**DEEP\_main\_9 Functions:**

* ***DEEP\_mPLVoverDyads:*** Calculates the mean of the PLVs or cross-frequency PLVs for each condition across all dyads.
* ***DEEP\_TFRoverDyads:***Calculates the mean of the time-frequency responses (TFRs) for each condition across all participants.
* ***DEEP\_powOverDyads:*** Calculates the mean of the power spectral densities (PSDs) for each condition across all participants.

**Important note:** We publish a dual EEG dataset of 35 mothers and their infants together with these pipeline instructions.