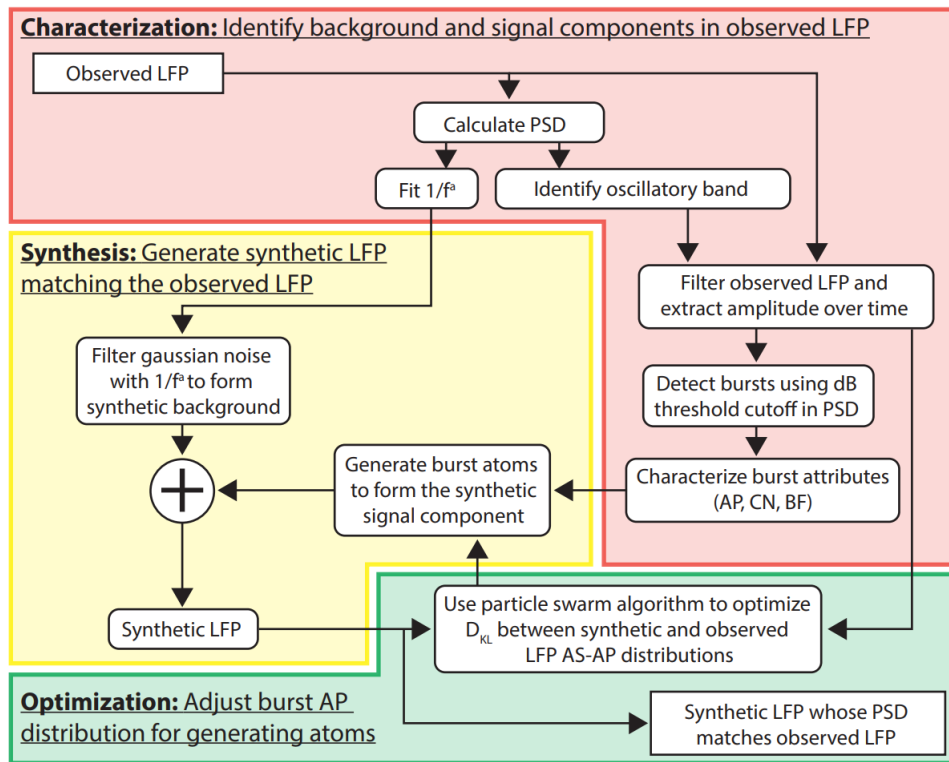# Design of Brainwaves Detection - GUI Instructions

Overview

The work flow has 3 major steps: 1. Characterization, 2. Synthesis, 3. Analysis. Each step does following things:
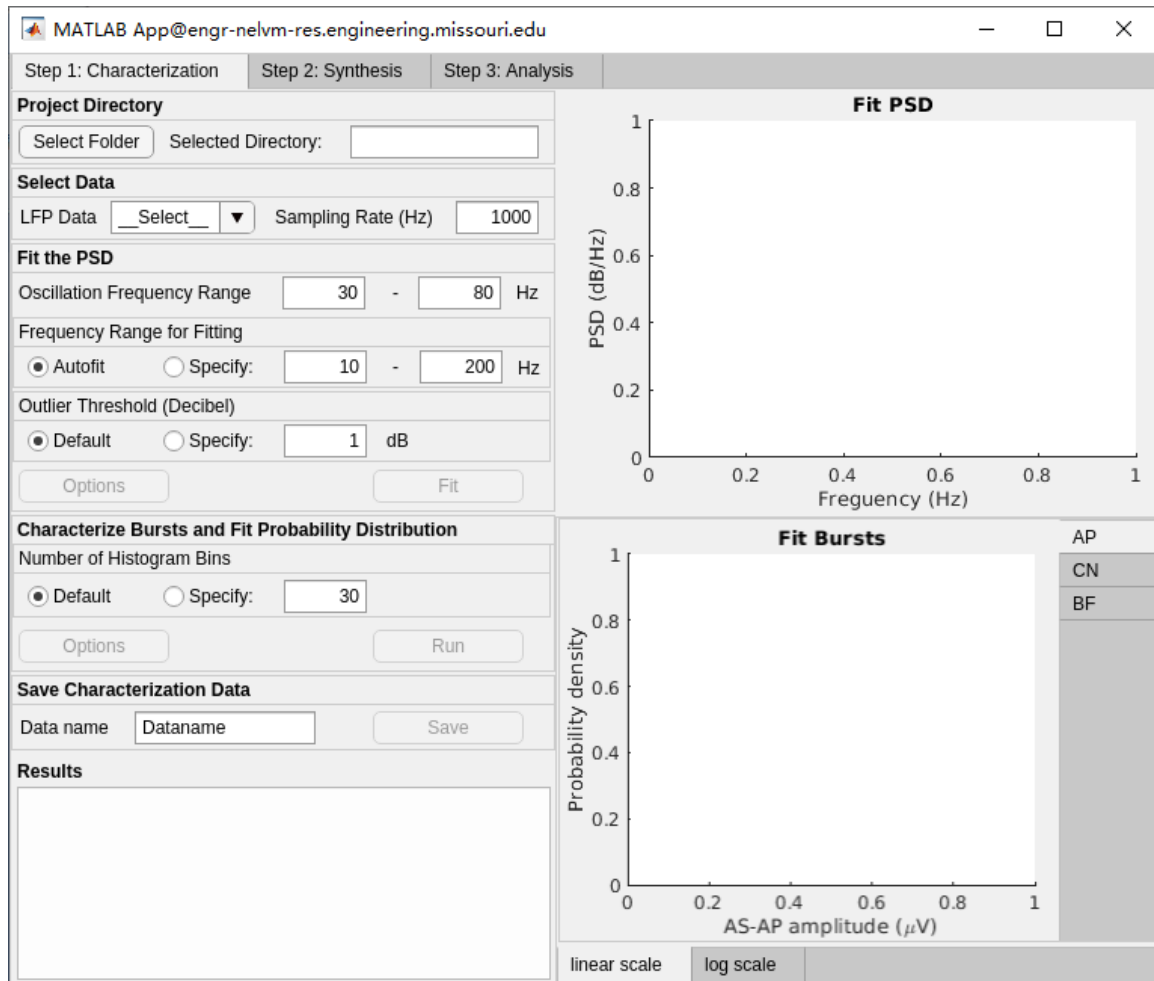
1. Select local field potential (LFP) recording as input. Run the algorithm to fit power spectral density (PSD), calculate signal-to-noise ratio (SNR) and other power spectral properties of the LFP. Then characterize bursts properties, such as distributions and their fit parameters of amplitude peak (AP), cycle numbers (CN) and burst frequency (BF). Finally save the characterization results.

2. Load results from previous step. Specify parameters for synthesizing LFP. Synthetic background component and oscillatory signal component will be generated. It's required to run an optimization for AP distribution parameters of burst atoms to be generated in the signal component. This step is essential for the synthetic data to reproduce the analytic signal amplitude peak (AS-AP) distribution of the observed data. There are three candidate types of distribution for the burst atom AP: exponential, gamma, and lognormal. Gamma is recommended. But users can try out all of them and pick the best fitting. The goodness of fit is indicated by Kullback–Leibler divergence (D_{KL}) or Jensen–Shannon divergence (D_{JS}). Lower divergence value indicates better fitting. Then generate synthetic LFP with optimized parameters and visualize the resulting properties including PSD and burst statistics as in step 1. Finally, save the synthesis parameters and data.

3. Analyze the synthetic data generated from previous step for detection problem. Define the ground truth with lower and upper bounds in synthetic signal trace and evaluate detection performance on the composite trace with receiver operating characteristic (ROC) curve. The relation between detection threshold and true/false positive rates will be shown.

Flow chart for step 1 and 2

Characterization: Identify background and signal components in observed LFP

Observed LFP → Calculate PSD

Calculate PSD → Fit 1/f^a

Calculate PSD → Identify oscillatory band

Synthesis: Generate synthetic LFP matching the observed LFP

Filter observed LFP and extract amplitude over time

Filter gaussian noise with 1/f^a to form synthetic background

Detect bursts using dB threshold cutoff in PSD

Generate burst atoms to form the synthetic signal component

Characterize burst attributes (AP, CN, BF)

$+$

Synthetic LFP

Use particle swarm algorithm to optimize D_{KL} between synthetic and observed LFP AS-AP distributions

Optimization: Adjust burst AP distribution for generating atoms

Synthetic LFP whose PSD matches observed LFP

GUI Instructions

Step 1: Characterization

Project Directory:
- Select the project directory where the result files will be saved. Use "Select Folder" button to select or directly type the path in the text box.

Select Data:
- "LFP Data" specifies the input data, your LFP recording. Select a variable from your Matlab workspace or type in the variable name or an expression. The variable needs to be a **vector** of LFP data or a **cell array** where each element is a vector of LFP segments. The latter format is useful for segmentized data.
- "Sampling Rate" is the recording sampling rate in Hz.

Fit the PSD:
- "Oscillation Frequency Range" is the frequency range of brainwave you are interested in.
- "Frequency Range for Fitting" is the range where the algorithm tries to fit a $1/f^{\alpha}$ curve to the power spectral density. You can adjust the range to find a best fitting

for the $1/f^{\alpha}$ background component. Or you can also use the "Autofit" option to automatically find the range.

- "Outlier Threshold" is the decibel threshold in the PSD used to find a bump in the oscillation frequency range of interest that indicates significant oscillatory signal power and to determine the signal frequency range. Try lowering this value if a bump is too small to be found.
- Click on "Fit" button to run. Fitting results will show up in the plot on the top right.

Characterize Bursts and Fit Probability Distribution:
- "Number of Histogram Bins" specifies the resolution of the distribution histogram. Adjust this number according to your data size.
- Click on "Run" button to run characterization. Burst statistics will show up in the plot on the bottom right. Switch the tabs to view different plots.

Save Characterization Data:
- Type a name for the characterization results and click on "Save" button. A dialog box will open and save the result file under the project directory.

Results:
- A text box where information/message are shown during runtime.

Step 2: Synthesis

Characterization Data:
- Select the file you saved from step 1 using the "Select File" button or type in the file path in the text box

Synthesizing Parameters:
- "Sampling Rate" specifies the sampling rate of the synthetic signal. You can use the same rate as the LFP data which yields best reproduction of it, or you can specify a different one. **Note** that large sampling rate will significantly slow down the process, especially during the optimization.
- "Rng Seed" specifies the seed for random number generator.
- When parameters are set, click on "Load" button before moving to next step.

Optimize Amplitude Distribution Parameters:
- Particle swarm optimization is used. It creates a population of models with different parameters and they will be attracted toward the point in the parameter space with minimal loss found after each iteration. The convergence criteria is that the number of stall iterations (consecutive iterations with no lower loss value found) reaches 5, or maximal number of iterations (30) is reached.

- "Synthetic Data Length" is the duration of generated signal. Shorter ones yield shorter runtime while longer ones yield lower variance in the resulting distribution meaning lower error in the loss.
- "Distribution Types" lists the available candidate distributions for the burst atom amplitude peak distribution that are used to generate the synthetic signal component. You can select multiple items by holding "ctrl" / "shift" + "mouse left click".
- "Population" specifies the particle swarm population for each candidate distribution. Select the candidate in the dropdown menu and type in the number in the text box. Clicking on reset button will restore default number. Large population helps searching the parameter space but takes longer to simulate.
- "Loss Function" is the loss for the optimization. It can be either KL divergence or JS divergence which indicates the similarity between the AS-AP distributions of the observed LFP and synthetic LFP.
- Click on "Fit" button to run. **Note** that the optimization may take several minutes to finish. The resulting AS-AP distributions of synthetic data generated by optimized parameters of selected candidate distribution types will show up in the plot on the top right, together with the distribution of observed data.

Generate Synthetic LFP:

- "Synthetic Data Length" is similar to that in previous step. It can be longer since it runs only once. Longer synthetic data yields lower variance in statistics for latter analysis.
- "Distribution Type" is the one among the optimized ones you will use to generate the synthetic data.
- Click on "Run" button to generate. The characterization results of the synthetic data similar to that of the observed data in step 1 will show up in the plot on the bottom right.
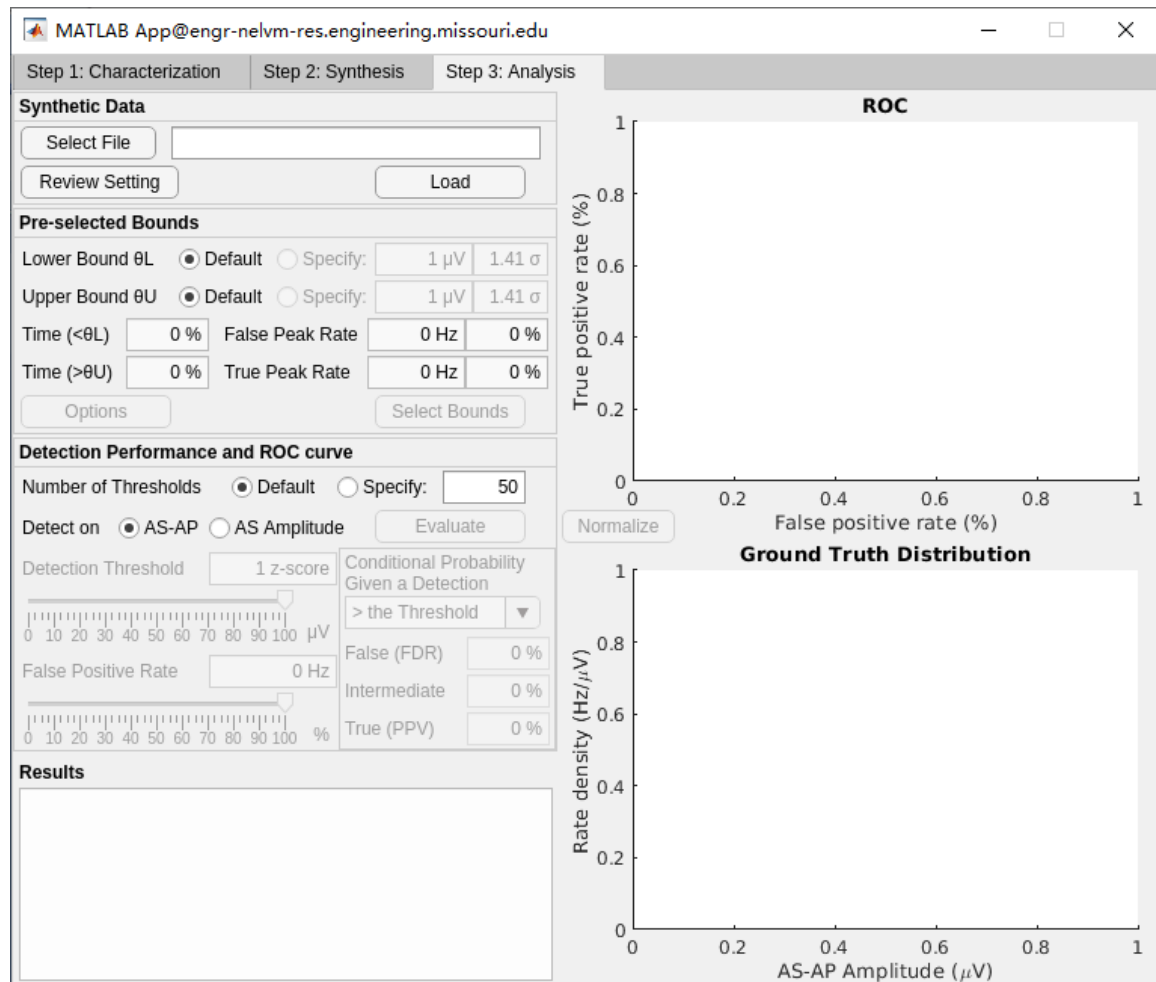
Save Synthetic Parameters:

- Save the synthesizing parameters obtained into a file. You can do it right after the optimization step. You can also check the box "also Save Synthetic LFP" after you run the "Generate Synthetic LFP" step. But the data file size could be large. Saving the data is not essential for step 3 because the synthetic data can be reproduced solely by the synthesizing parameters saved. Data generated by only one "Distribution Type" will be saved in one file. You can rerun "Generate Synthetic LFP" with another distribution type and save the data to another file.
- Type in the case name you want for the file in the text box. The case name will be a suffix following the characterization data file name.
- Click on "Save" button, a dialog box will show up and save the file in the project directory. **Note** that the synthesizing parameter file has to be in the same directory as the characterization data file in order to work properly.

Results:

A text box where information/message are shown during runtime.

Step 3: Analysis

Synthetic Data:

  • Select the synthesizing parameter file you saved from step 2 using the "Select File" button or type in the file path in the text box.

  • Click on "Review Setting" to review the synthesizing parameters. The GUI will switch to step 2 and the setting for the synthesizing parameters will be shown on the panels.

  • Once you confirm with the setting, you can click on the "Load" button to load the synthetic data. If the synthetic data was not saved in the synthesizing parameter file, it will take a while to generate the data. The synthetic LFP and its two components, background and signal, will be band pass filtered with the signal frequency range.

Pre-selected Bounds:

  • "Lower Bound" and "Upper Bound" specify the amplitude bounds in the signal trace that are used to define the ground truth of three categories in the synthetic LFP, true, false bursts, and an intermediate category. You can select default to use recommended bounds. Or you can specify the bound in unit of amplitude or in

multiples of the standard deviation of the background trace (σ, also the square root of power).

- Click on "Select Bounds" to evaluate the bounds and create ground truth data. If the upper bound happens to be lower than the lower bound, which could happen in some cases when using default bounds, the upper bound will be set to equal the lower bound. The percentage of durations where the signal trace have amplitude lower than the lower bound, or higher than the upper bound, will be shown. The resulting true/false peak rate in the ground truth data will be shown in both Hz and percentage of total AS-AP in the synthetic LFP. You can adjust the bounds to obtain desired true/false peak rate according to experience.

Detection Performance and ROC curve:

- "Number of Thresholds" specifies the number of detection threshold to be evaluated. Adjusting this number will change the resolution of the ROC curve and distribution plot.
- You can choose the target of detection. "AS-AP" option detects only the peaks in the analytic signal amplitude. "AS amplitude" option detects the analytic signal amplitude at all time points.
- Click on "Evaluate" button to evaluate the detection. The ROC curve will show up on the top right and the distribution of ground truth will show up on the bottom right.
- When detecting on "AS-AP", you can click on "Normalize" button to toggle the false positive rate scale between the percentage of total false peaks and ratio to total true peaks.
- "Detection Threshold" specifies a particular detection threshold you want to evaluate. You can either type in the text box using z-score of the AS amplitude of the synthetic LFP, or use the slider to change the amplitude threshold. A marker will show up on the ROC curve indicating current operation point and a line will show up in the distribution plot indicating the threshold location.
- "False positive rate" shows the resulting false positive rate in the text box and the slider given the specified detection threshold. You can also type in the text box or use the slider to specify a desired false positive rate and the corresponding detection threshold will be shown in the slider and text box above. This is known as the constant false alarm method for selecting a detection threshold.
- "Conditional Probability given a Detection" panel shows the conditional probability of each category a detection belongs to given the selected detection threshold. When you select option "> the Threshold", given each detection with amplitude **above** the threshold, the probability of a true positive is also known as the positive predictive value (PPV), and the probability of a false positive detection is also known as the false discovery rate (FDR). The conditional probabilities correspond to the areas under the distribution curves to the right of the detection threshold line. When you select option "= the Threshold" instead,

given each detection with amplitude ***equal to*** the threshold, the probability of each category that the detection belongs to is shown. The conditional probabilities correspond to the length proportions of each category along the detection threshold line in the distribution plot.

Results:

- A text box where information/message are shown during runtime.

Example data

An example LFP data "LFP_BLA_gamma.mat" is provide in the folder "example data".

After loading it as a "MAT-file" into Matlab workspace, run
    LFP_seg = cellfun(@(x) scale*double(x),LFP_seg,'UniformOutput',false);
to scale it from integer values to microvolts.

The sampling frequency is 1000 Hz, also provided as the variable "fs".

Gamma oscillation is present in this data. Oscillation frequency range from 30 Hz to 80 Hz is recommended for characterizing the PSD.