Contents

Overview 3

Installation 3

File Menu 3

- Data types Definition & Supported file formats 3
- Open 4
- Save 4

Visualization Menu 4

- Plot 4
- Live Time 5
- ISI 6
- 3D Plot 6
- Raw Data 7

Toolstrip 8

- Detection 8
  - Filtering 8
  - Thresholding 8
- Auto-Sorting 10
  - Alignment 11
  - Statistical filtering 11
  - t-Distribution Sorting 12
  - Run 13
- Manual Sorting 13
  - Clusters Index 13
  - Merge 13
  - Remove 13
  - Resort 13
  - Denoise 13
  - Assign to nearest 14
  - Tag 15

PCA remove 15
PCA group 16
Back 16
Main Window 16
Detected Spikes PCA 16
Sorted Spikes PCA 16
Sorted Waveforms 16
Log 17
Legend 17
Update figures 17
Demo Video 17
References 17

ROSS: Offline Spike Sorter V1.0.0

Overview

Spike sorting is the process of extracting the activity of the single neurons from the extracellular recording, which is a vital step for investigating brain behavior in neuroscience. ROSS is a fast and accurate offline spike sorting software developed based on the methods described in [1], and is made freely available.

The software gets the raw extracellular recording as input and all steps including filtering, alignment, detection and sorting could be done in a fully-automated manner. Several configuration options are provided in each of the mentioned phases, enabling user to change the parameters of the algorithms. Moreover, efficient manual tools are provided along with the useful visualization methods in order to fine-tune the sorting results by user interference.

Installation

Recommended System:

- MATLAB > 2018a
- 12GB RAM

Download the package to a local folder. Run matlab and navigate to the folder. Then you could easily start Offline Spike Sorter by running “main.mlapp” or typing “main” in the command window.

File Menu

Data types Definition & Supported file formats

Software input and output data are in the form of “.mat” files, which might include one or more of the following variables, depending on the processing level of the data.

a. SpikeMat:
   An MxN matrix, where M is the number of waveforms (or spikes), and N is the number of time samples in each of the waveforms, i.e. each row contains a spike waveform.
b. SpikeTime:
   An Mx1 vector, where the m’th element is the corresponding time of the occurrence of the m’th spike in the SpiekMat.
c. optimal_set:
   A structure, containing the soring results. It has two fields: cluster_index and typicality. cluster_index contains the index of corresponding cluster of each waveform. Typicality is a criterion, which denotes the sample quality of being representative of its assigned cluster.

d. tag:
A Cx1 vector, where C is the number of clusters. Each element is a user-defined label for the corresponding cluster (e.g. multi-unit, single-unit …).

Open

a. Raw Data

Opens Raw Data, which is the basic data file accepted by the software. It is a Mat file that must contain a vector of the extracellular recording signal. None of the above-defined variables is needed in Raw Data. In a case there exist multiple variables in your selected Mat File, a window will open which asks you to select the desired variable.

b. Detected Data

Opens Detected Data, which is a Mat file containing two variables: I) SpikeMat, and II) SpikeTime.

c. Sorted Data

Opens Sorted Data, which is a Mat file containing three variables: I) SpikeMat, II) SpikeTime, III) optimal_set.

Save

a. Detection Result

Saves detection results including the spike waveforms and spiking times in a Mat file with two variables, named SpikeMat, and SpikeTime, respectively.

b. Sorting Result

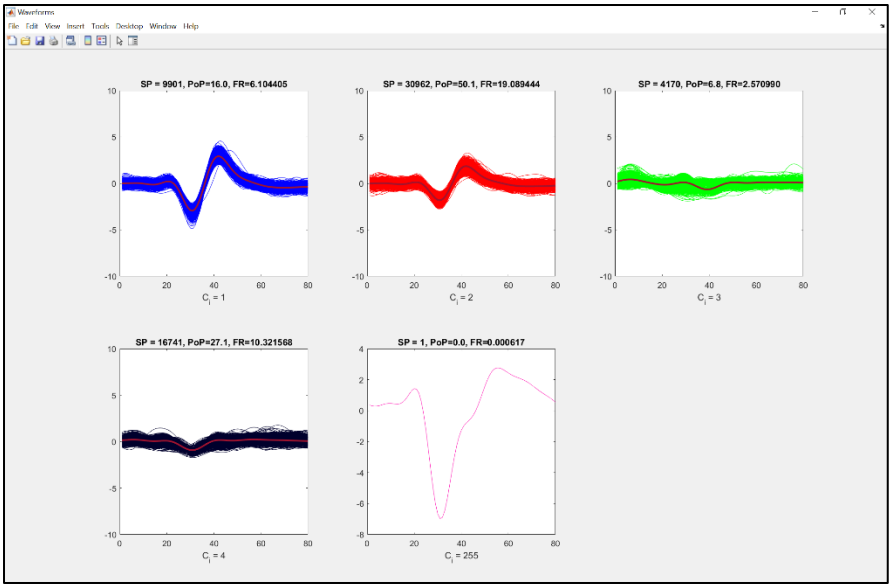
Saves sorting results including the spike waveforms, spiking times, sorting information and tags in a Mat file with four variables: SpikeMat, SpikeTime, optimal_set, and tag.

Visualization Menu

Visualization tools needed for manual analysis on spike detection and sorting results appear in this menu.

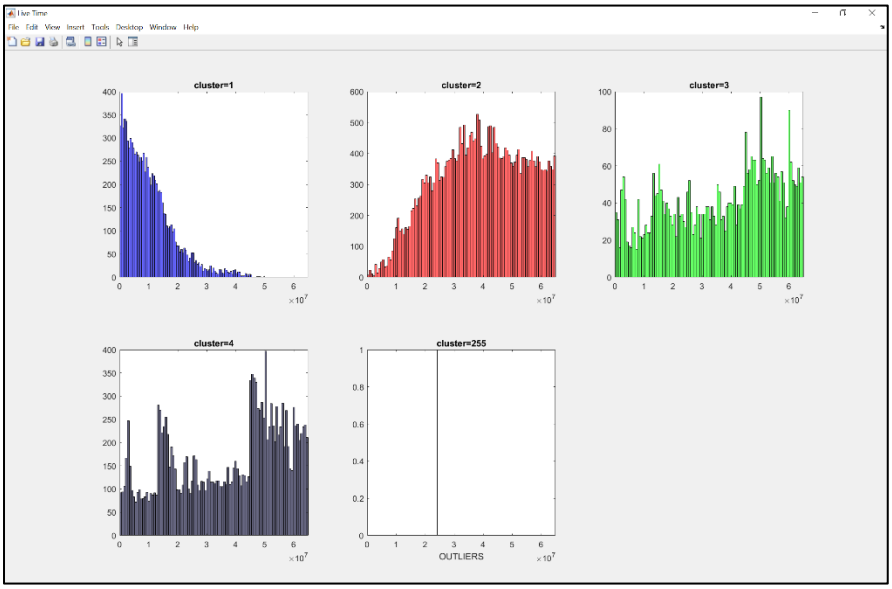
Plot

When clicked, a window containing multiple time plots will open. Each plot represents waveforms assigned to one cluster versus time. In the bottom sample, the sorting result contains five clusters.



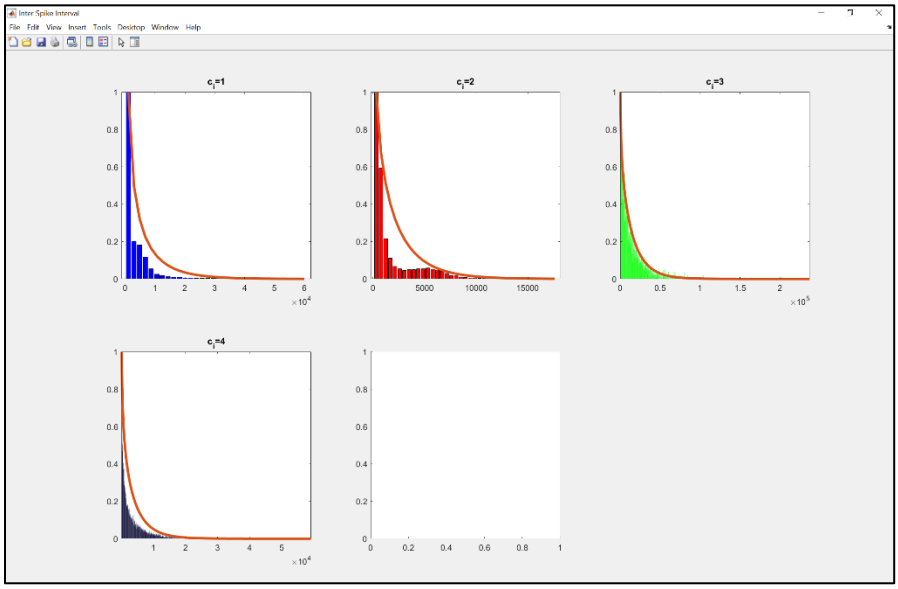
Live Time

Clicking on Live Time will open a window containing multiple plots, each for one of the clusters. Each plot represents the histogram of the occurrence time of the spikes corresponding to that cluster. Live Time is useful for analysis on the firing rate of the neurons during the recording. A sample Live Time for five clusters is shown below.



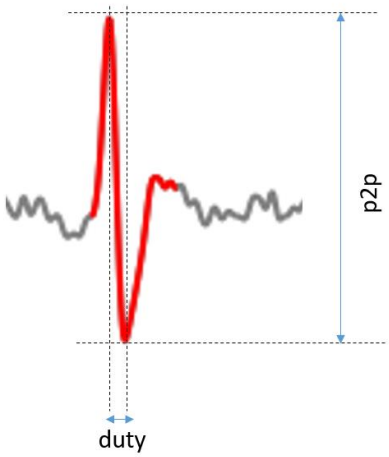
ISI

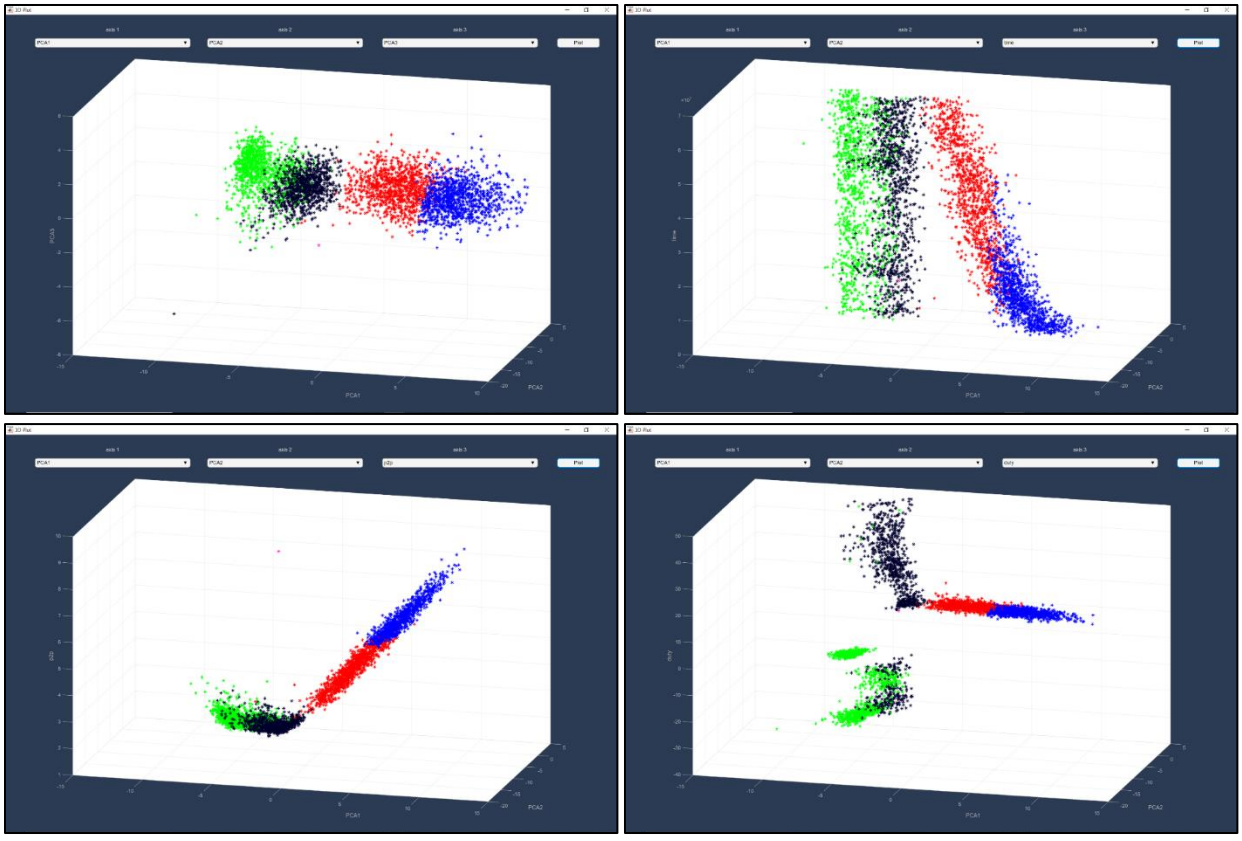
The distribution of inter-spike interval for each of the clusters will appear using this tool. Moreover a fitted gamma distribution is plotted for each cluster.



3D Plot

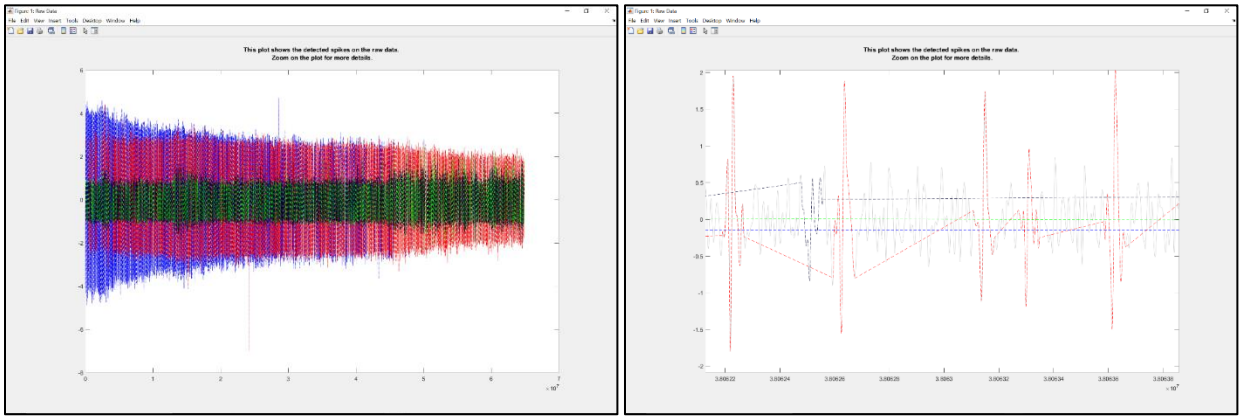
The tool provides different 3D plots to user. Each dimension of the 3D plots could be chosen arbitrary between the following statistics: “PCA1”, “PCA2”, “PCA3”, “time” (time of the spike), “p2p” (amplitude distance between maximum and minimum of the spike) and “duty” (time distance between maximum and minimum of the spike). Sample 3D plots are shown below. User could also explore between different 2D views of the 3D plot by making the third axis “none”.





Raw Data

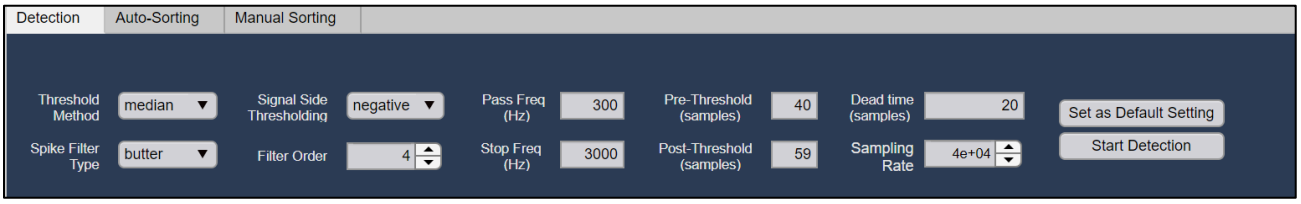
This option opens a window containing time plot of the raw data, where detected spikes could become visible by zooming.



Toolstrip

Detection

Detection tab in the toolstrip provides a fully automated and powerful spike detection tool described in [1]. The detection process is done in two steps: filtering, and thresholding. The following manual configurations are made available to user in the detection tab.



Filtering

Filtering is the first step to separate spike activity from the LFP and background noise. A zero phase IIR filter is implemented by filtering the signal in forward and reverse orders in an offline mode. Configurable settings are as followed:

Spike Filter Type

The filter type for zero phase filtering of raw data. Only “butterworth” type is supported in the current version.

Filter Order

The order of the IIR zero phase filter is selected using this option.

Pass Freq (Hz)

This option determines pass frequency of the band pass filter.

Stop Freq (Hz)

This option determines stop frequency of the band pass filter.

Sampling Rate

The sampling rate of the input raw data is determined here.

Thresholding

Spike detection is done by a thresholding method in this step. Settings of this section are:

Threshold method

Threshold method must be chosen between three options: “median”, “wavelet” or “plexon”. This selection determines the method used to estimate the threshold. Each of the methods is described in the following.

Median: The standard deviation of noise (σ_n) is estimated according to Donoho's rule, then the spike detection threshold is set three times σ_n. The formula for "both" side thresholding would be:

$$\sigma_n = \frac{median(|r_F|)}{0.6745}$$

$$t_s = 3 \times \sigma_n$$

Where r_F denotes filtered signal, and t_s is the calculated threshold.

Wavelet: First the wavelet decomposition of unfiltered signal at level 2 using "db3" wavelet is calculated. Then the noise of wavelet coefficients are estimated. The detection threshold is defined as five times the estimated noise. "wavedec" and "wnoisest" matlab functions are used:

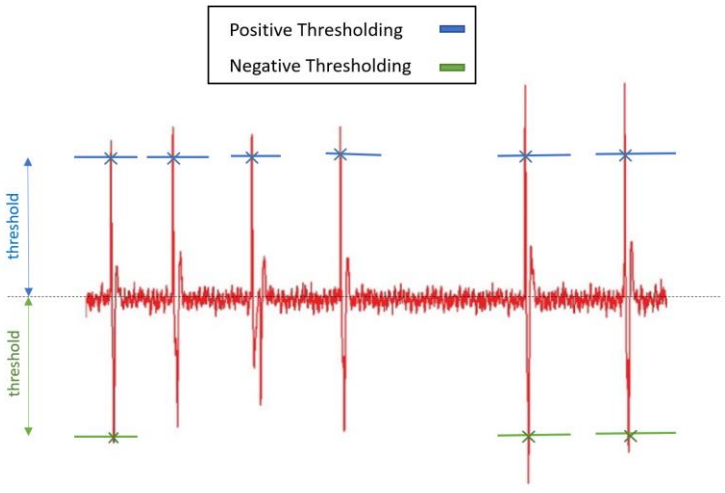
```
[c,l] = wavedec(Spike,2,'db3');

thr = 5*wnoisest(c,l,2);
```

Plexon: First the samples with amplitudes higher than 2.7 times standard deviation of the full signal, are considered as outlier and would be ignored in threshold computation. Then the threshold is defined as four times the standard deviation of the filtered signal without considering the detected outliers.

Signal Side Thresholding

This option is used to determine how the amplitude of the signal would be compared to the threshold. User could select "positive", "negative" or "two" amplitudes to be accounted for thresholding. Also by choosing "mean", an automatic decision between "positive" or "negative" options is made based on the sign of the signal mean.



Pre-Threshold (samples)

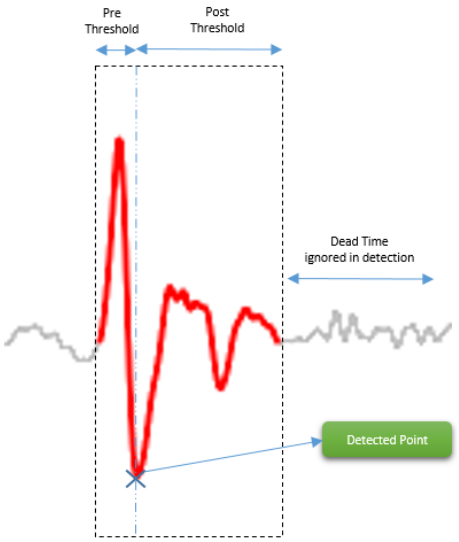
Number of samples before the detected (threshold crossing) point to be kept in the spike waveform.

Post-Threshold (samples)

Number of samples after the detected (threshold crossing) point to be kept in the spike waveform.[1]

Dead time (samples)

Number of samples after the detected spike waveform, which are ignored in detection.



Set as default settings

Saves the current settings as the default.

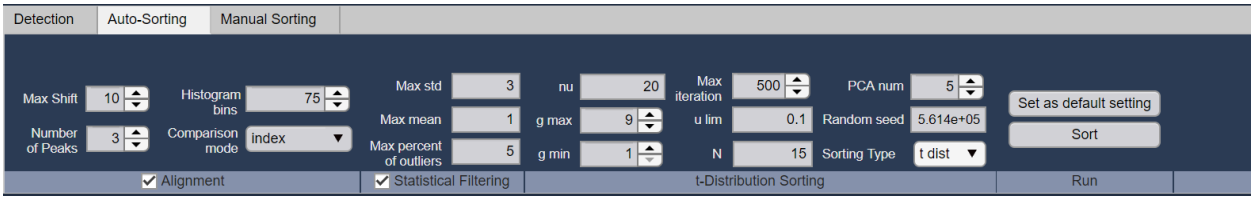
Start Detection

Starts the detection process.

Auto-Sorting

Auto-Sorting tab provides useful tools for automatic spike sorting in four sections. Alignment tool is used to perform waveform alignment. Statistical Filtering is a noise removal process. The third section contains settings related to t-Distribution sorting or skew-t , and the last part includes Run commands.

[1] The total number of samples is “Pre-Threshold” + “Post-Threshold” + one.



Alignment

Alignment is an effective step before clustering to improve results. Waveform alignment is implemented based on the method described in [1]. By checking “Alignment” the alignment method would run before the sorting process, otherwise, a simple alignment algorithm would be considered which aligns all spikes based on their minimum. Four settings are made available to user for configuration in the alignment algorithm:

Max Shift

Maximum number of points that the waveforms are allowed to be shifted in the alignment process.

Histogram Bins

Number of histogram bins for plotting histogram of the waveform extremums.

Number of Peaks

Number of histogram peaks which are considered as aligning points.

Comparison Mode

Determines the mode for grouping the waveforms. If “magnitude” is selected, the waveforms are divided in two groups based on the amplitude of their extremums. (Samples with higher amplitude in maximum than minimum, are assigned to one group, and the remaining ones to the second group). Else, if “index” is selected, the waveforms are divided based on the indices of their extremums. (Samples with their maximum index lower than their minimum index, are assigned to one group and the remaining ones form the second group).

Statistical filtering

An automatic noise removal process is designed in [1] based on data statistics. Settings related to this phase lie in this section. Detected noise samples would be assigned to cluster number 255 in the sorting phase.

Max std

Magnitude of the threshold used for waveform standard deviation. Waveforms with higher standard deviation values would be considered as noise.

Max mean

Magnitude of the threshold used for waveform mean. Waveforms with higher absolute mean values will be considered as noise.

Max percent of outliers

This parameter determines maximum percent of the waveforms allowed to be removed as outlier.

t-Distribution Sorting

The clustering tool is implemented based on mixture of distributions modeling method, including t-distribution mixture modeling described in [2], and also skew-t mixture modeling presented in [1]. The provided manual configurations are as followed:

nu

Determines initial value for nu (ν) parameter of t-distribution.

Max iterations

Maximum number of iterations for Expectation Maximization (EM) algorithm.

PCA num

Number of PCA components used as features for clustering.

g max

Maximum number of clusters.

g min

Minimum number of clusters.

u lim / error

The name of this option varies between “u lim” in the case “t dist” sorting type is selected, and “error” for “skew-t dist” sorting.

“u lim” is the limit for determining waveforms typicality. The waveforms with typicality value lower than “u lim” will be accounted as not typical and assigned to cluster number 254.

“error” is a threshold for defining another stop critera in addition to “Max iterations”. The algorithm would stop in the case that maximum change among all parameters is smaller than “error”.

N

Regularization parameter. This option is only activated when “t dist” is selected as the sorting type.

Random seed

Random seed for MATLAB random generator engine.

Sorting Type

There are two options for clustering method, which are assumption of “t dist” or “skew-t dist” distributions for clusters in the sorting algorithm. It should be noted that selecting skew-t would turn “u lim“ option into “ error“, and deactivate “N “.

Run

Set as default settings

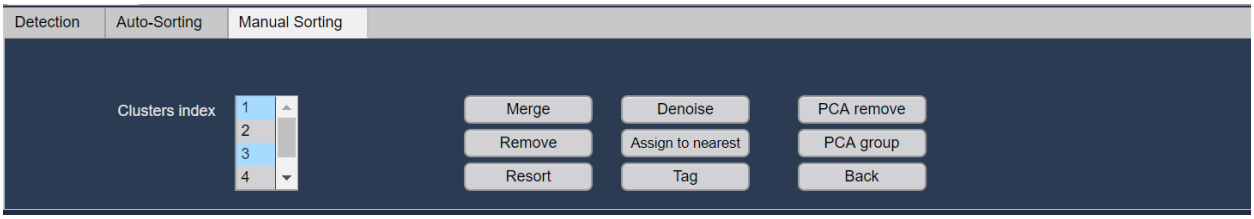
Saves the current configurations as default settings.

Run

Starts automatic sorting process.

Manual Sorting

Manual Sorting tab provides several manual tools for modifying the auto-sorting results.



Clusters Index

One or a number of clusters could be selected from this list for applying the manual changes.

Merge

Using this button, the selected clusters in the “Clusters Index” list will be merged to form one cluster.

Remove

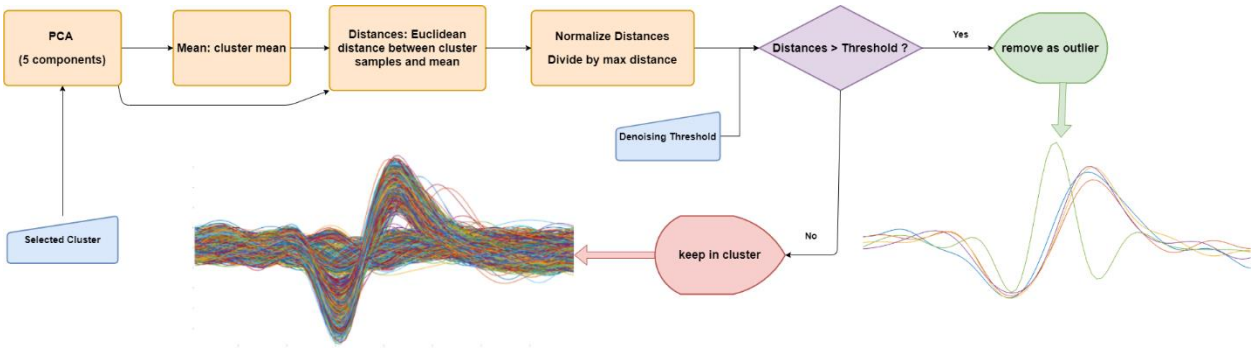
This option will remove the selected clusters in the “Clusters Index” list.

Resort

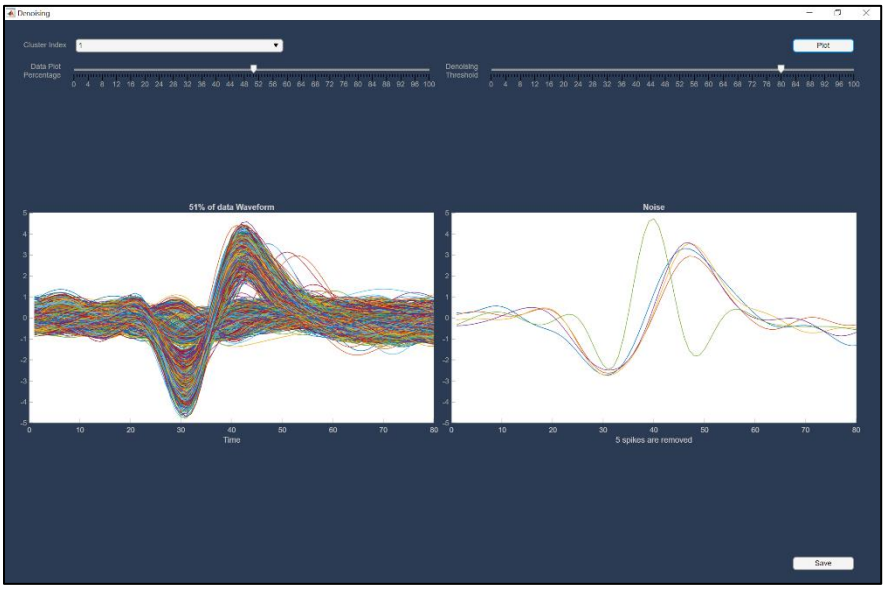
Performs auto-sorting only on the selected clusters in the “Clusters Index” list once again.

Denoise

This option will open a window containing denoising tool. This tool removes the outliers by a thresholding method. The distance between each waveform and the average waveform of the cluster is normalized and compared to a threshold. The denoising algorithm is illustrated in the following figure.

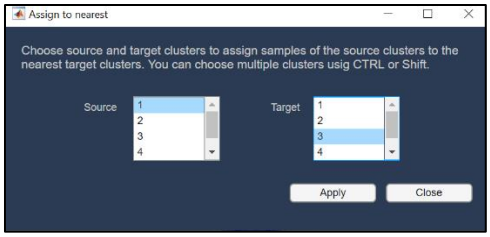


User can select the intended cluster from the “Cluster index” dropdown menu. The percentage of the random plotted data is set by changing “Data Plot Percentage”, and the threshold for denoising algorithm is chosen in “Denoising Threshold”. By clicking on “Plot”, two figures will be plotted. The left figure displays random samples of the waveforms with the desired plot percentage. The right figure shows all of the detected noise waveforms. Clicking “save” will save the outliers under noise cluster (cluster index = 255).



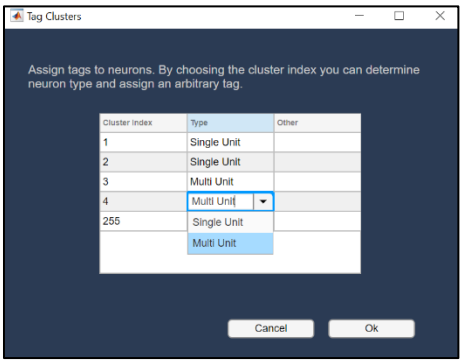
Assign to nearest

This tool assigns all of the samples existing in the “source” cluster to the “target” clusters based on their distance from each of the targets.



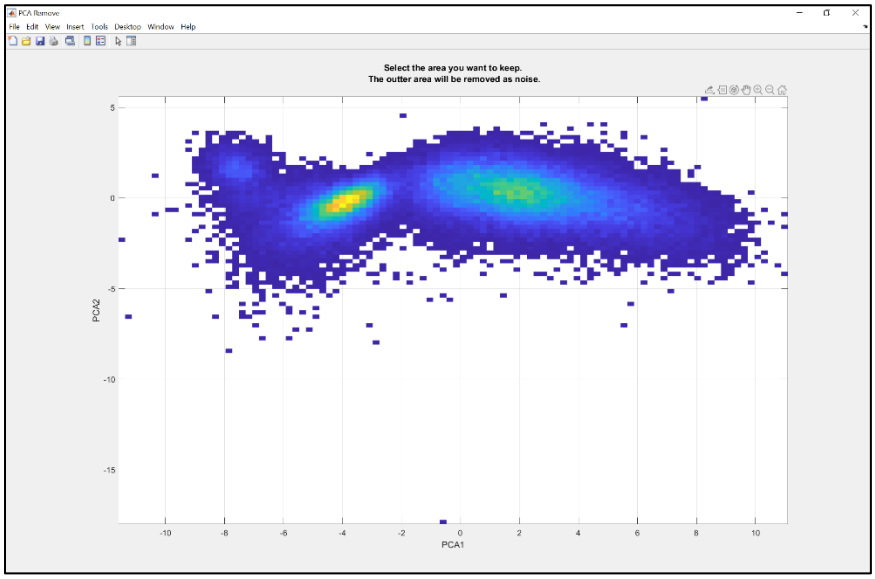
Tag

User could define an arbitrary tag for each of the clusters, or choose the tag from the existing options. There exist a “Type” tag, including two options: “Single Unit” and “multi Unit”. Also, there exist an “Other” tag, which could be any string.



PCA remove

Samples could be removed manually from the feature space using this tool. A window containing 2D PCA plot of the samples is provided to user. User could manually select the desired region of the feature space and throw away the samples out of this region.



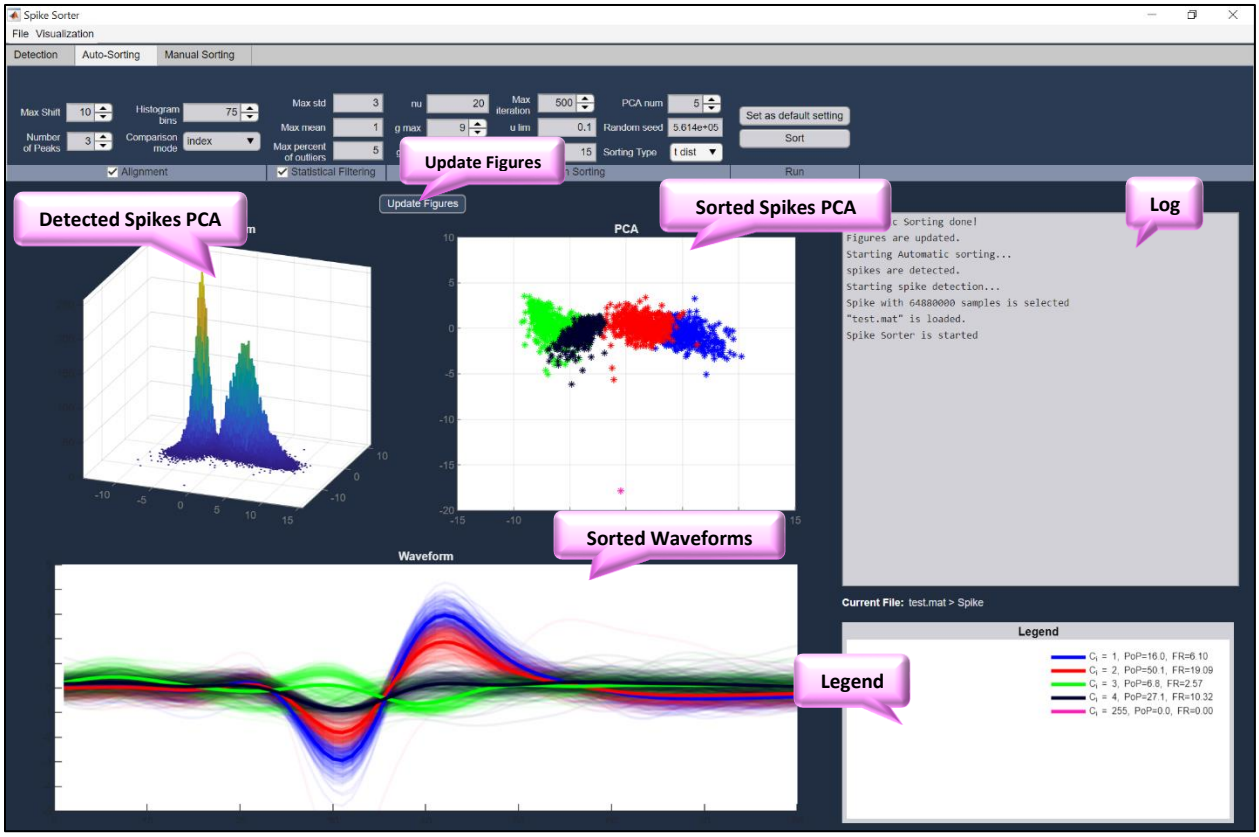
PCA group

User could assign the desired regions in the feature space to a new cluster. A window similar to the previous figure is provided, and user could select the region of the new cluster.

Back

This button is used to undo the recent action. At most five last actions could be undone.

Main Window



Detected Spikes PCA: output of the detection phase, displaying histogram of 2D PCA of the detected spikes.

Sorted Spikes PCA: output of the auto-sorting phase, displaying PCA of the sorted spikes with different colors corresponding to different clusters. For each cluster, randomly selected samples are plotted.

Sorted Waveforms: output of the auto-sorting phase, displaying the waveforms of the sorted spikes with different colors for different clusters.

Log: keeps the history of the performed actions.

Legend: represents the legend for all visualizations.

Update figures: Updates all the plotted figures manually.

Demo Video

References

[1]

[2] Shoham, Shy, Matthew R. Fellows, and Richard A. Normann. "Robust, automatic spike sorting using mixtures of multivariate t-distributions." Journal of neuroscience methods 127.2 (2003): 111-122.