

## Unsupervised spike sorting documentation

Mingwen & Ankur

The current documentation is for “Wave\_clus\_Ming\_v1.0”.

**Readme:** **this is still a test version**, so I suggest users first test the script with data you have manually sorted and compare them with automatically sorted results. In this way, you would have feeling about whether it’s worth your time to try this script.

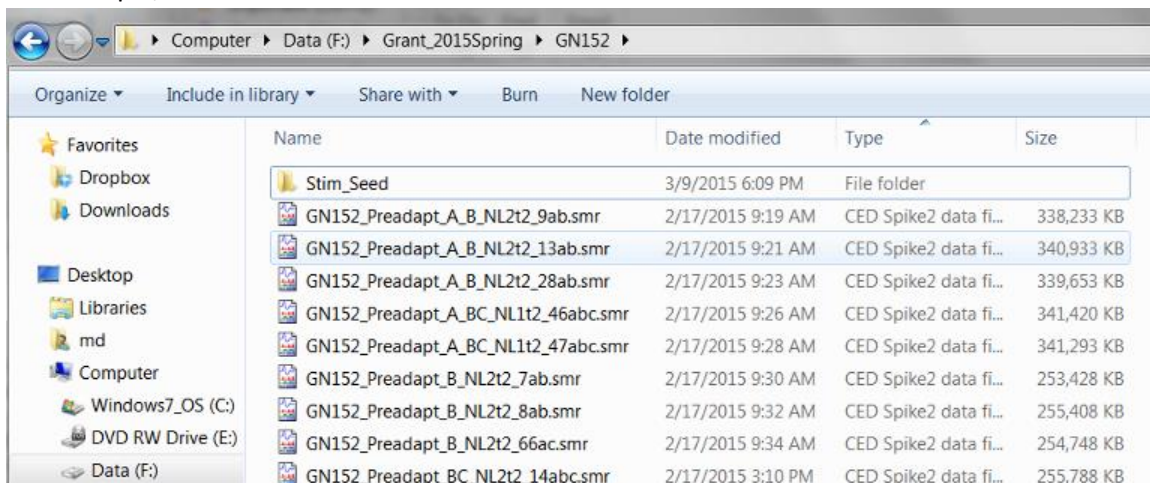
### 1<sup>st</sup> Step:

Export raw waveform data from Spike2 “.smr” file to Matlab “.mat” file.

**Important:** Please also read the notice at the end of this part for detailed requirements!

The current script used for this is called "Spike2Matlab\_CBC\_Batch\_v0.1.s2s", which you could find it inside the “Wave\_clus\_Ming\_v1.0”.

Before running the script, put Spike2 data you want to sort into a user-created folder. For example,



Then, open the script in Spike2, run it, and follow instructions from prompt.

First, select the folder where you put the raw Spike2 files; then select where you want to save the extracted ".mat" files. [Suggestion: create a folder in advance to store ".mat" file]

Finally, wait until Spike2 exported all ".mat" files.

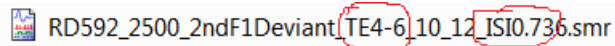
### IMPORTANT Notice1:

Your **spike2 data file name could only contain letters, numbers, and underscore** [ a-z, A-

2015/03/11

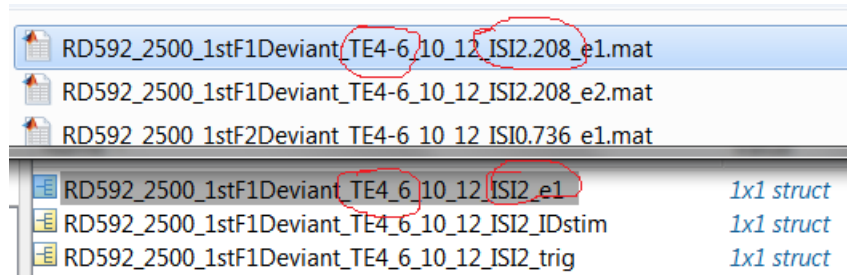
Z, 0-9, \_ ]. They cannot contain special characteristics like dash "-" or extra dot "." other than the one used before file extension. Otherwise, the Matlab clustering script won't work correctly.

One example of bad file names are shown below, which contains dash and extra dot:

 RD592\_2500\_2ndF1Deviant\_TE4-6\_10\_12\_ISI0.736.smr

The reason to not use special characteristics is to make the name of ".mat" file consistent with the name of variables it contains.

Notice the difference between the filename of .mat files and the variables it contains in the figure below. The scripts work on the assumption that the two names are the same. If they're not, it won't give you the output.



#### IMPORTANT Notice2:

If you know some electrodes are bad and don't have spikes or signals, delete them in advance or delete them after export. Bad channels may cause the clustering algorithm to crash or other unexpected error.

## 2nd Step:

Automatic spike detection, spike sorting, and data format conversion.

The current master wav\_clus script is called "autoss\_batch\_main", which could be found in "\*\Wave\_clus\_Ming\_v1.0\Wave\_clus\Batch\_files\ autoss\_batch\_main".

Before running "autoss\_batch\_main", copy ".mat" files exported from spike2 to a wav\_clus input folder [For example, create a folder on desktop called "SS\_Input"]. Don't use the original folder where you put the extracted ".mat" waveform files because clustering algorithm would modify and delete the files in the input folder.

Then, open "autoss\_batch\_main" and change the first coding line. Change the string path of "wav\_clus" to where you put "Wave\_clus\_Ming\_v1.0" folder (shown as purple in the figure below).

```
% Add the path of wave clus % EDIT HERE %Where wave_clus.m exists
addpath(genpath('C:\Users\md\Dropbox\Matlab_Spike2_Scripts\Matlab_Lab_Script\Wave_clus_Ming_v1.0'));
```

Then, run the script. As the prompt would ask, first select where you put ".mat" files [E.g., "SS\_Input" folder] and where you want to store the output [Suggestion: create an output folder like "SS\_Output" on desktop in advance].

Then, wait until the matlab finishes sorting all the channels.

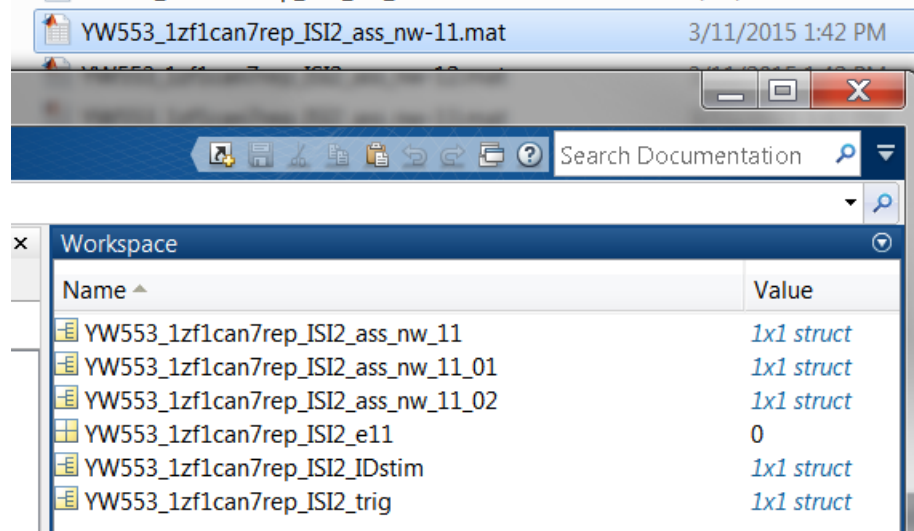
Upon successful completion of spike sorting, all the waveform “.mat” files in the input folder (your “SS\_Input”) should be deleted. If there’re files left, it means error occurred. The most possible reason is: 1, these are bad channels without spikes or signals; 2, the naming convention is not followed.

The important output generated by the script include:

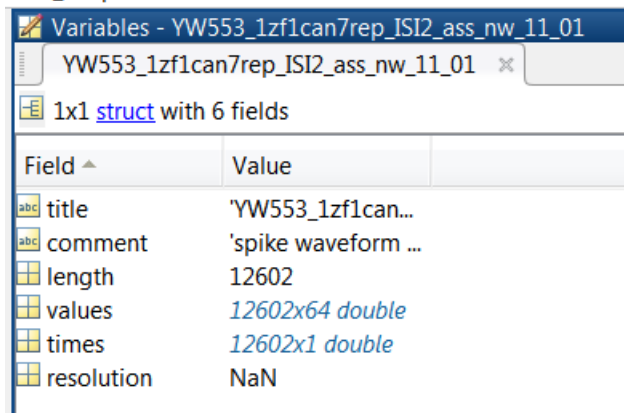
- a. **"\*\_ass-nw-\*.mat" file**: store the spikesorting results and it's compatible with ".mat" file that you would get when export wavemark channels from Spike2 to matlab. ("\*" indicate a number here: electrode number or cluster number).

Inside **"\*\_ass-nw-\*.mat" file**, important variables include:

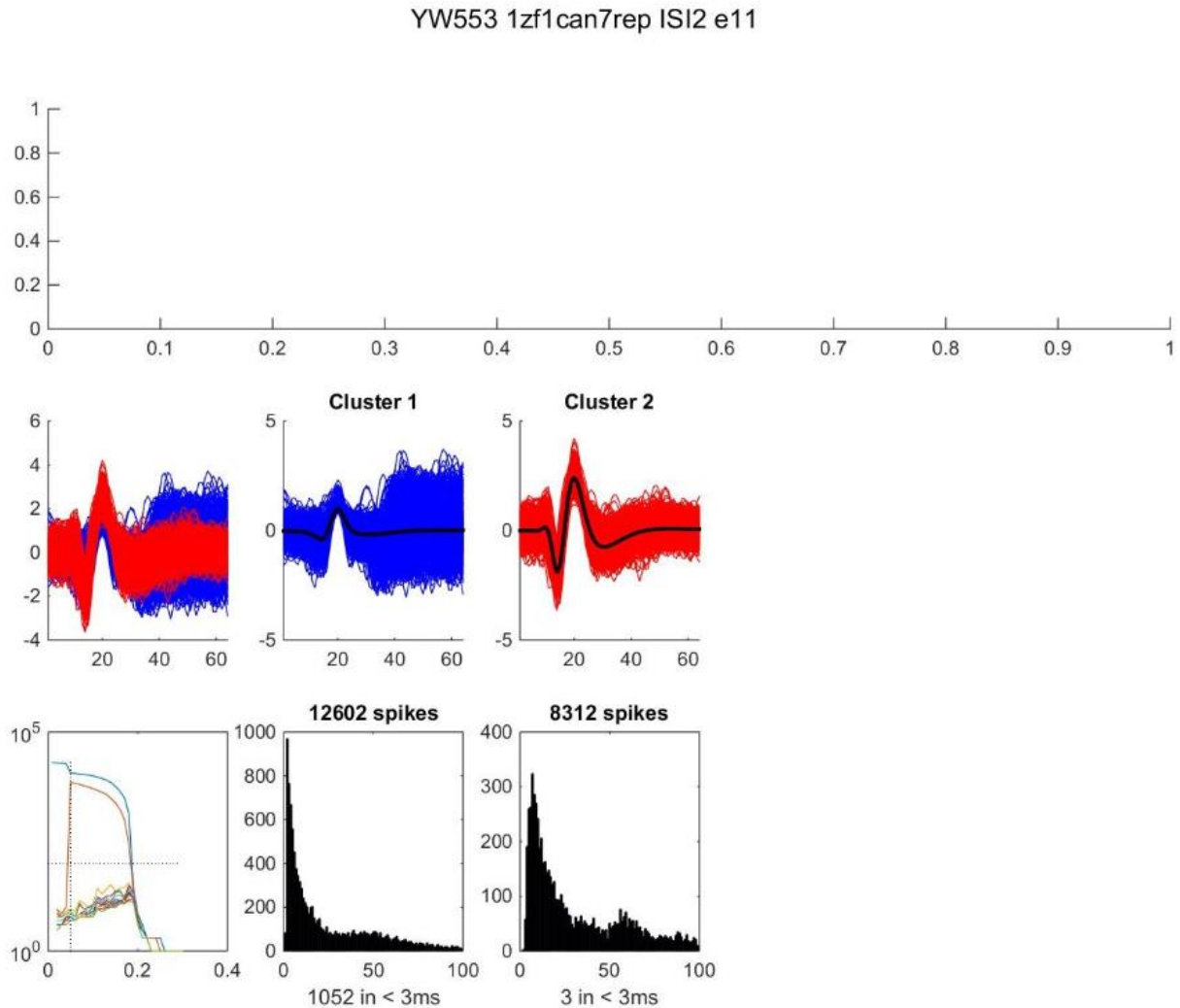
- 1) trig channel, IDstim channel, and data channel;
- 2) data channel **"\*\_ass\_nw\_"** contains all the spikes (Multiunit activity);
- 3) data channel **"\*\_ass\_nw\_\*\_0"** contains spikes from one particular cluster.



**"\*\_ass\_nw\_\*\_0" struct variable** contains title, times (spike timing), values (spike waveform),  $i^{\text{th}}$  row containing the waveform for  $i^{\text{th}}$  spike in spiketiming).



- b. **Figures for each processed channel:** used to get an intuitive feeling about the spikesorting results. E.g.,



- c. “cluster\_results.txt” log file: record all the files that spike sorting algorithm processed.  
Notice, files don’t follow naming convention would also be listed here. But you won’t get the most important **“\*\_ass-nw-\*.mat”** file.
- d. Other files:
- 1) “times\_filename.mat”: contains spike sorting results without format conversion, including sorting parameters.
  - 2) “filename\_spikes.mat”: contains spike sorting results without format conversion and no sorting parameters.
  - 3) other files: may be useful if you want to spike sort some channels using wav\_clus in a supervised way.

2015/03/11

**Possible future development:**

1. More user-friendly gui for supervised spike-sorting in wav\_clus.
2. Export well-sorted spikes back to Spike2 “.smr” file.
3. Other suggestions.
4. If you find bugs or problems, first check the Notices mentioned in this documentation. If unsolved, let me or Ankur know.