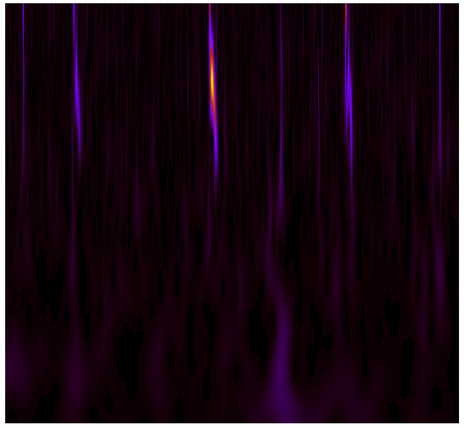
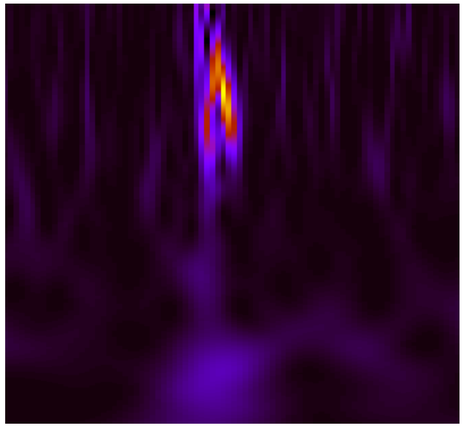
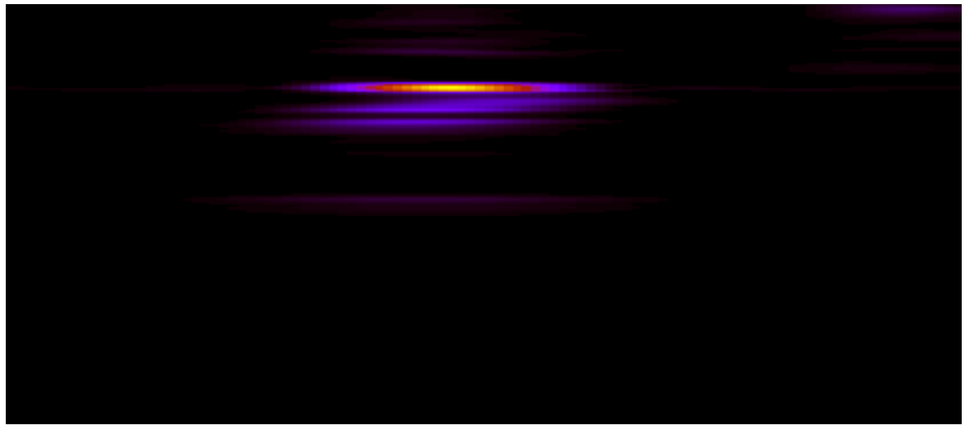
Machine Learning Meeting
April 26 2019



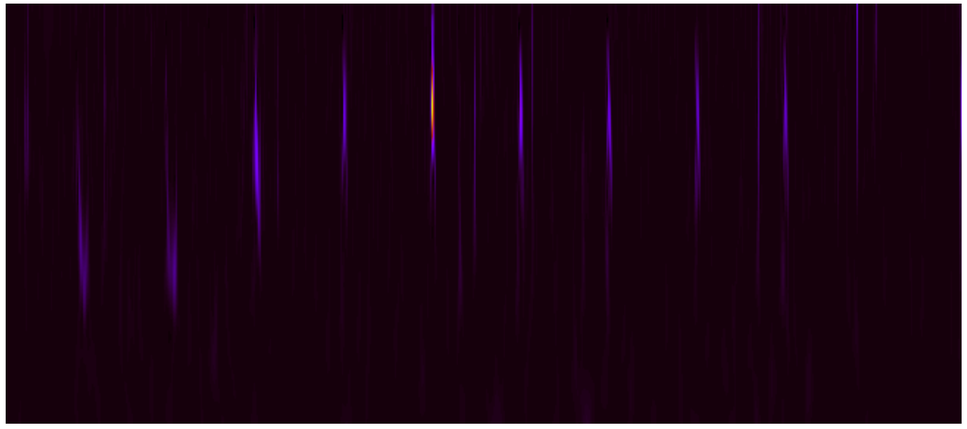
Preparation of Glitch dataset and class for MLAs

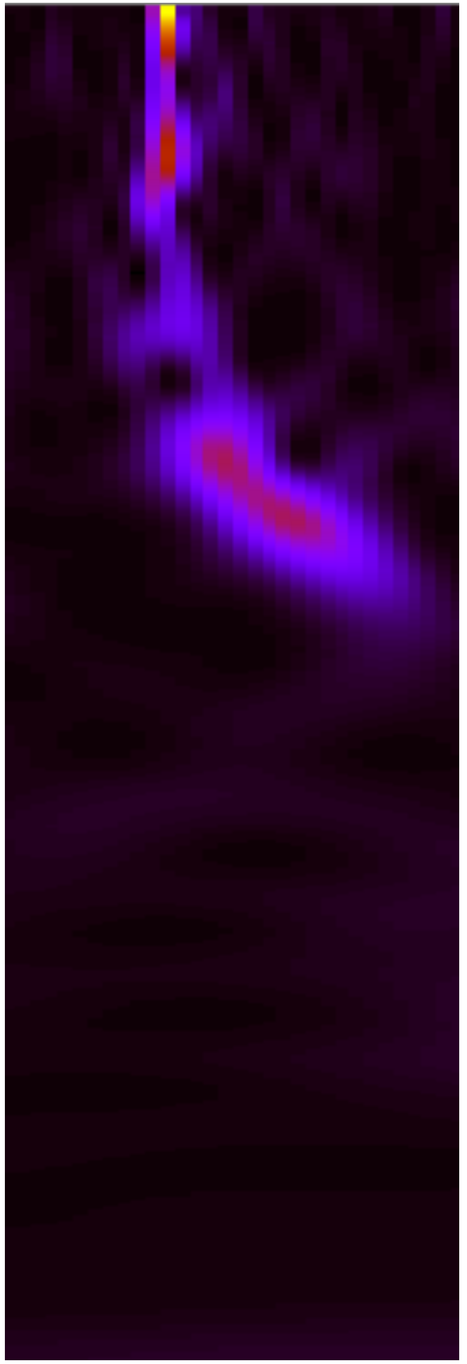
Jung, Pil-Jong



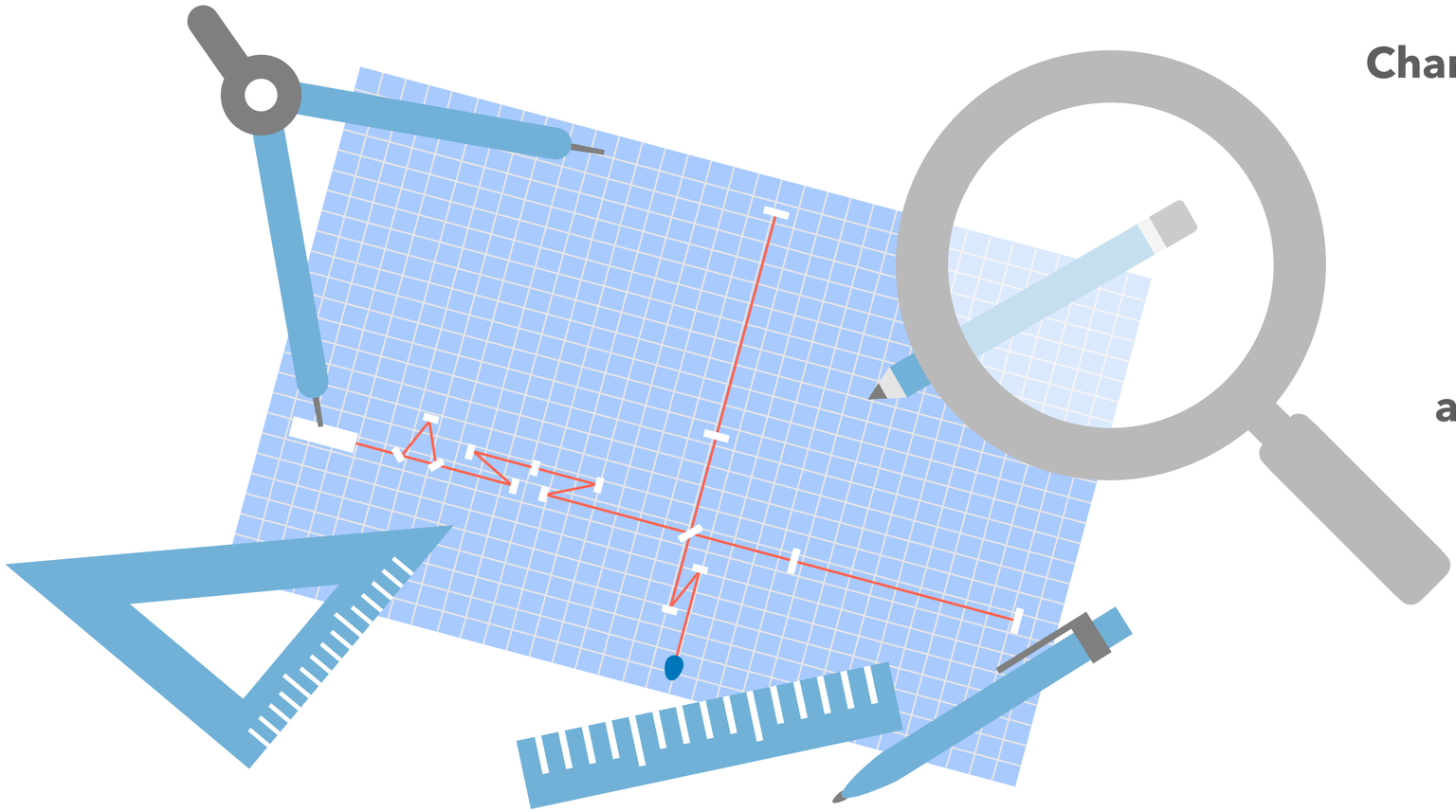


- Contants





Detector Characterization



Channel Safety Study

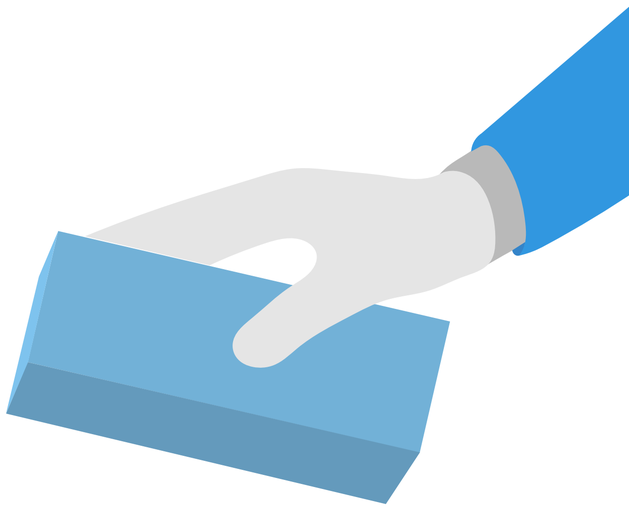
Data Quality Study

Glitch Study

and so on

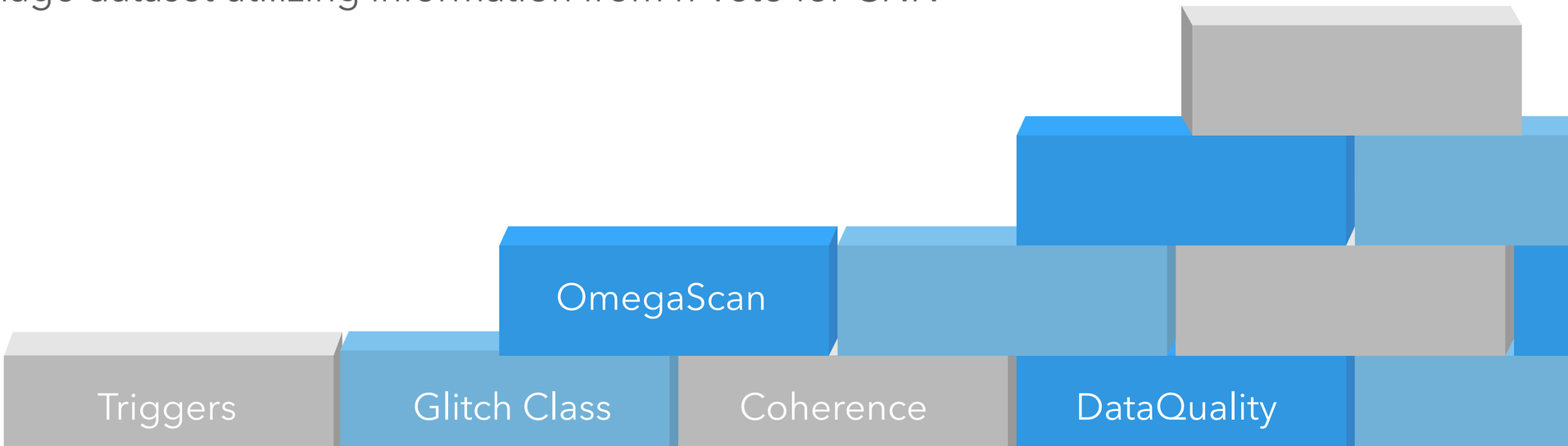
Needs

- The number of glitches that increase as the sensitivity
- KAGRA does not have much basic information
  such as Glitch type, DataQuality and so on
- Machine Running is required for efficiency of analysis



First Step

- Generate image dataset utilizing information from h-veto for CNN



Round details

Round 1, Winner = K1:PSL-PMC_SLOW_MON_OUT_DQ, window = 0.01, SNR thresh = 15.0 [top]

Winner: K1:PSL-PMC_SLOW_MON_OUT_DQ

SNR threshold: 15.0

Window: 0.01

Significance: 61.40

Veto segments: [txt] [xml]

Veto triggers: [txt]

Vetoed primary triggers: [txt]

Unvetoed primary triggers: [txt]

Omega scans:

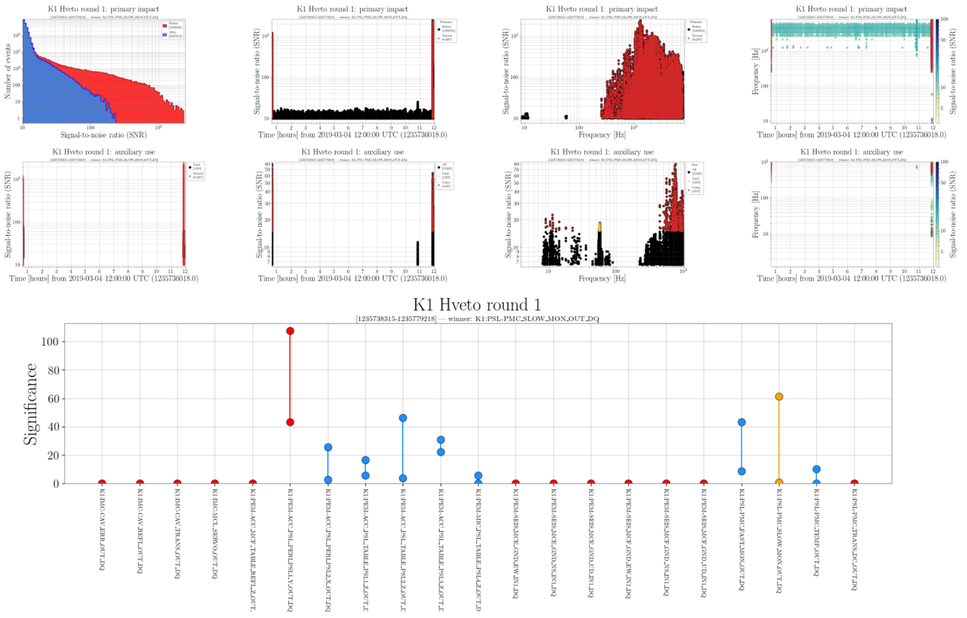
1235778802.418213 [SNR 31.2] [p] [a]

1235778802.502441 [SNR 17.8] [p] [a]

1235778802.494873 [SNR 34.8] [p] [a]

1235778802.442932 [SNR 141.8] [p] [a]

1235738349.607605 [SNR 158.2] [p] [a]



./triggers/K1-HVETO_RAW_TRIGS_ROUND_0-1235738315-40903.txt

./triggers/K1-HVETO_VETOED_TRIGS_ROUND_1-1235738315-40903.txt

./triggers/K1-HVETO_WINNER_TRIGS_ROUND_1-1235738315-40903.txt

Run h-veto

↓

Get vetoed triggers and coherence information

↓

Find loudest triggers without duplication

↓

Make table (.ecsv)

↓

Save OmegaScan plots

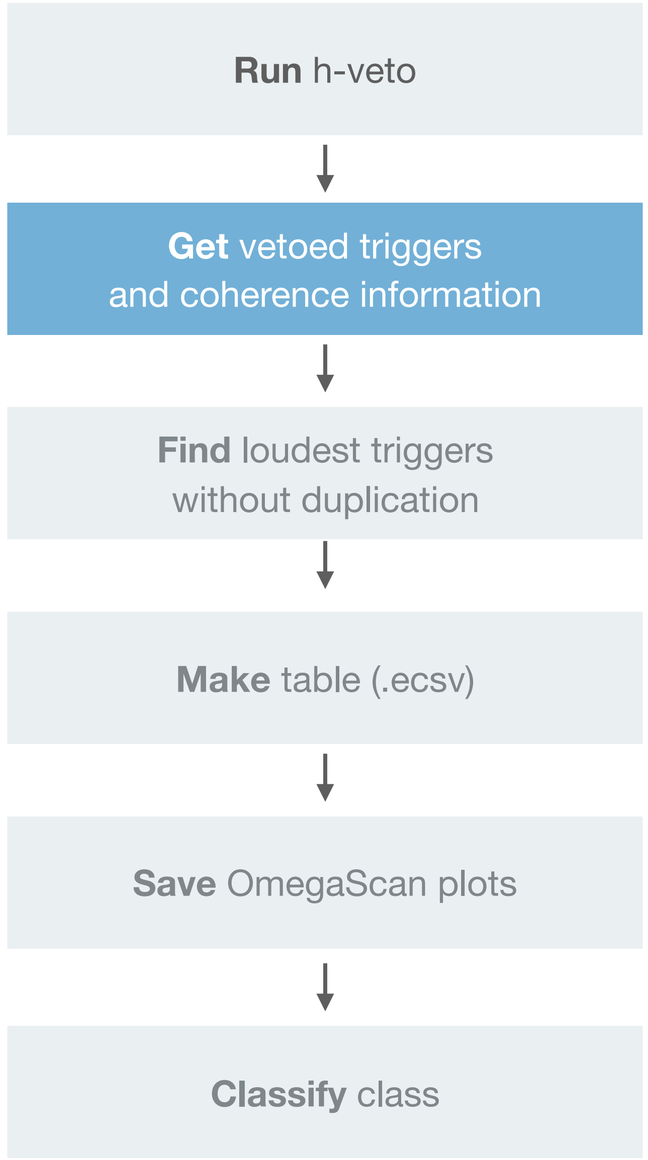
↓

Classify class

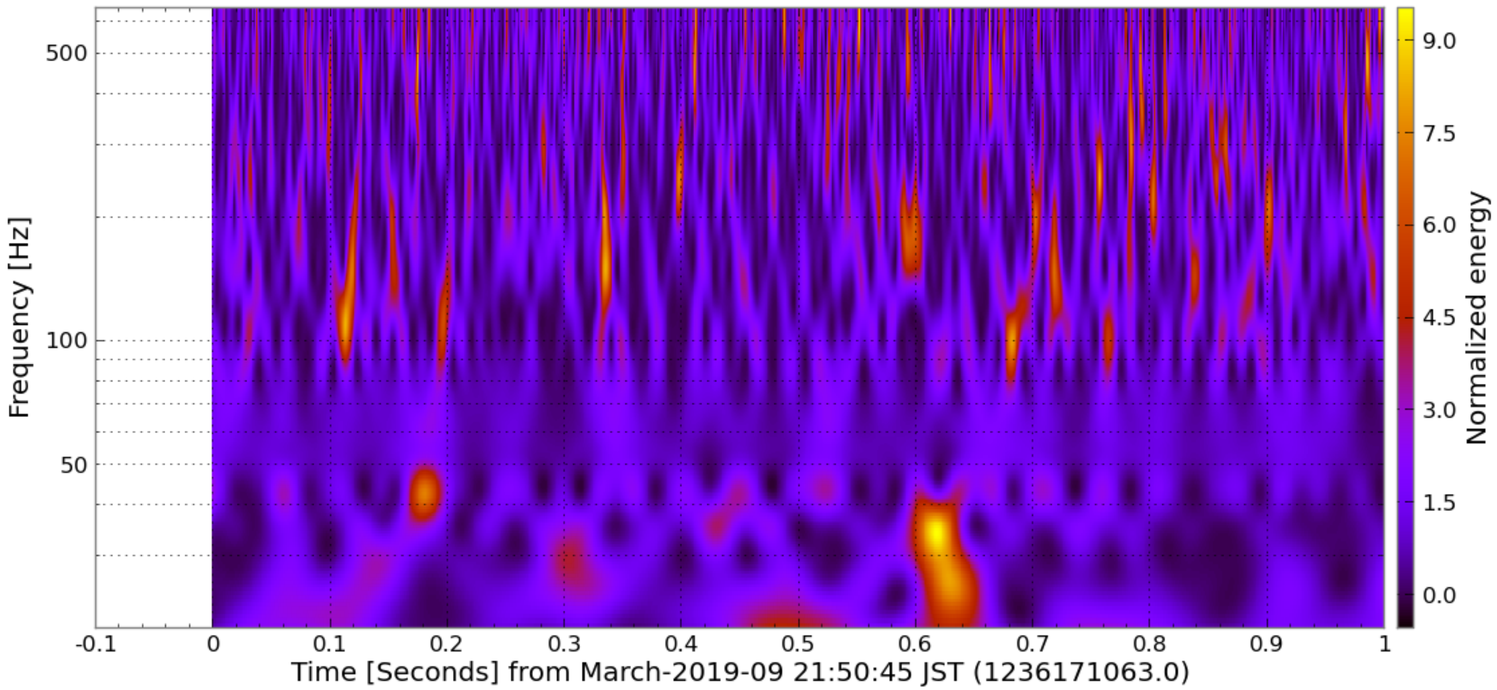
- Dataset with coherence information

$ cat ./triggers/K1-HVETO_VETOED_TRIGS_ROUND_1-1235738315-40903.txt

time peak_frequency snr channel
1235738318.291992 746.904296875 6.186349868774414 K1:PSL-PMC_MIXER_MON_OUT_DQ
1235738318.292968 751.12451171875 6.614719867706299 K1:PSL-PMC_MIXER_MON_OUT_DQ
1235738318.292968 736.7763061523438 6.464479923248291 K1:PSL-PMC_MIXER_MON_OUT_DQ
1235738318.293945 746.904296875 6.470779895782471 K1:PSL-PMC_MIXER_MON_OUT_DQ
1235738318.295898 746.904296875 6.432270050048828 K1:PSL-PMC_MIXER_MON_OUT_DQ
1235738318.297851 746.904296875 6.177060127258301 K1:PSL-PMC_MIXER_MON_OUT_DQ
1235738318.300781 736.7763061523438 6.754469871520996 K1:PSL-PMC_MIXER_MON_OUT_DQ
1235738318.300781 751.12451171875 6.807519912719727 K1:PSL-PMC_MIXER_MON_OUT_DQ
1235738349.482422 946.7334594726562 6.098869800567627 K1:PSL-PMC_MIXER_MON_OUT_DQ
1235738349.486328 946.7334594726562 6.722030162811279 K1:PSL-PMC_MIXER_MON_OUT_DQ
1235738349.490234 946.7334594726562 6.860730171203613 K1:PSL-PMC_MIXER_MON_OUT_DQ
1235738349.49414 946.7334594726562 6.501389980316162 K1:PSL-PMC_MIXER_MON_OUT_DQ
1235738349.521484 910.9091796875 7.0609002113342285 K1:PSL-PMC_MIXER_MON_OUT_DQ
1235738349.521484 893.5086669921875 7.740039825439453 K1:PSL-PMC_MIXER_MON_OUT_DQ
1235738349.521484 928.6485595703125 6.119629859924316 K1:PSL-PMC_MIXER_MON_OUT_DQ
1235738349.523437 360.9208984375 7.836510181427002 K1:PSL-PMC_MIXER_MON_OUT_DQ
1235738349.523437 340.630126953125 6.668270111083984 K1:PSL-PMC_MIXER_MON_OUT_DQ



- Dataset with coherence information



Run h-veto

↓

Get vetoed triggers
and coherence information

↓

Find loudest triggers
without duplication

↓

Make table (.ecsv)

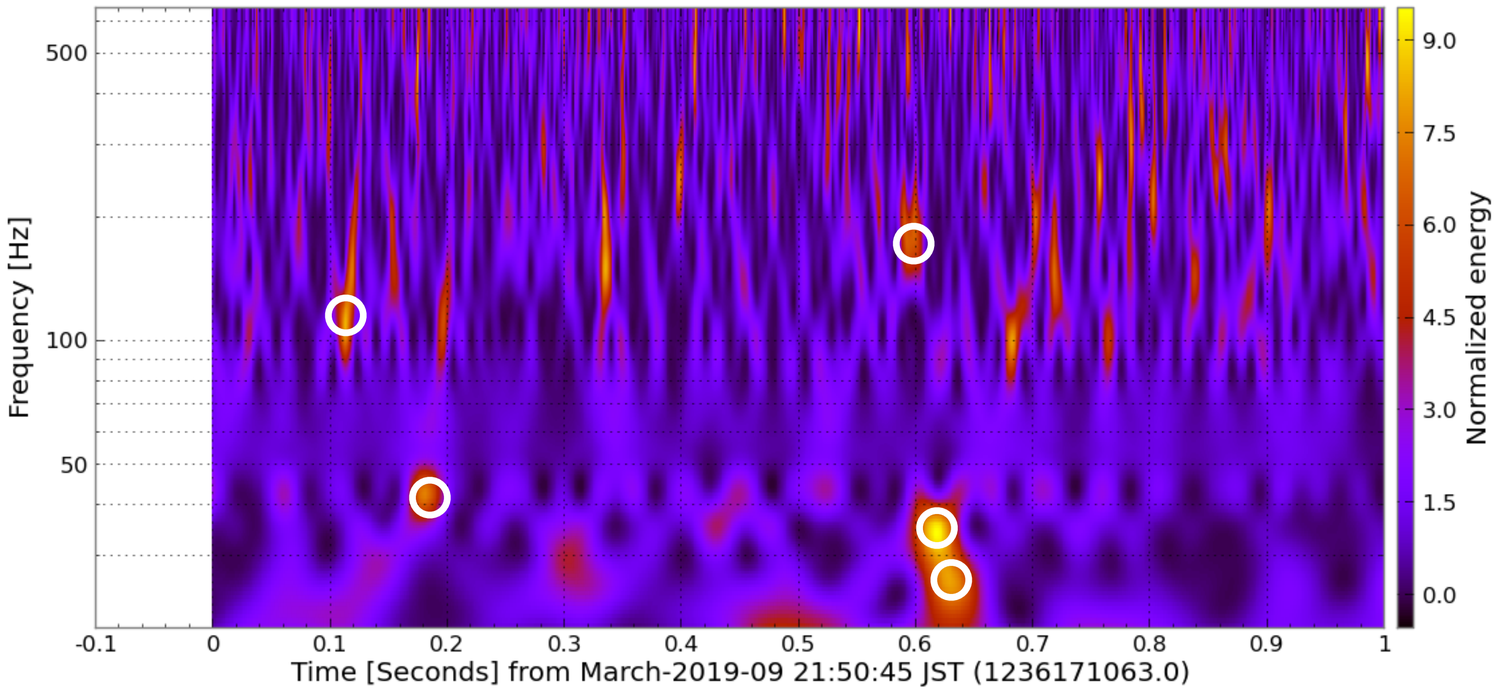
↓

Save OmegaScan plots

↓

Classify class

- Dataset with coherence information



Top 5 Loudest triggers

1236171063.61992 36.904296875 9.186349868774414 K1:PSL-PMC_MIXER_MON_OUT_DQ

1236171063.62968 41.12451171875 9.614719867706299 K1:PSL-PMC_MIXER_MON_OUT_DQ

1236171063.62968 36.7763061523438 9.464479923248291 K1:PSL-PMC_MIXER_MON_OUT_DQ

1236171063.623945 36.904296875 9.470779895782471 K1:PSL-PMC_MIXER_MON_OUT_DQ

1236171063.695898 36.904296875 9.432270050048828 K1:PSL-PMC_MIXER_MON_OUT_DQ

Run h-veto

↓

Get vetoed triggers
and coherence information

↓

Find loudest triggers
without duplication

↓

Make table (.ecsv)

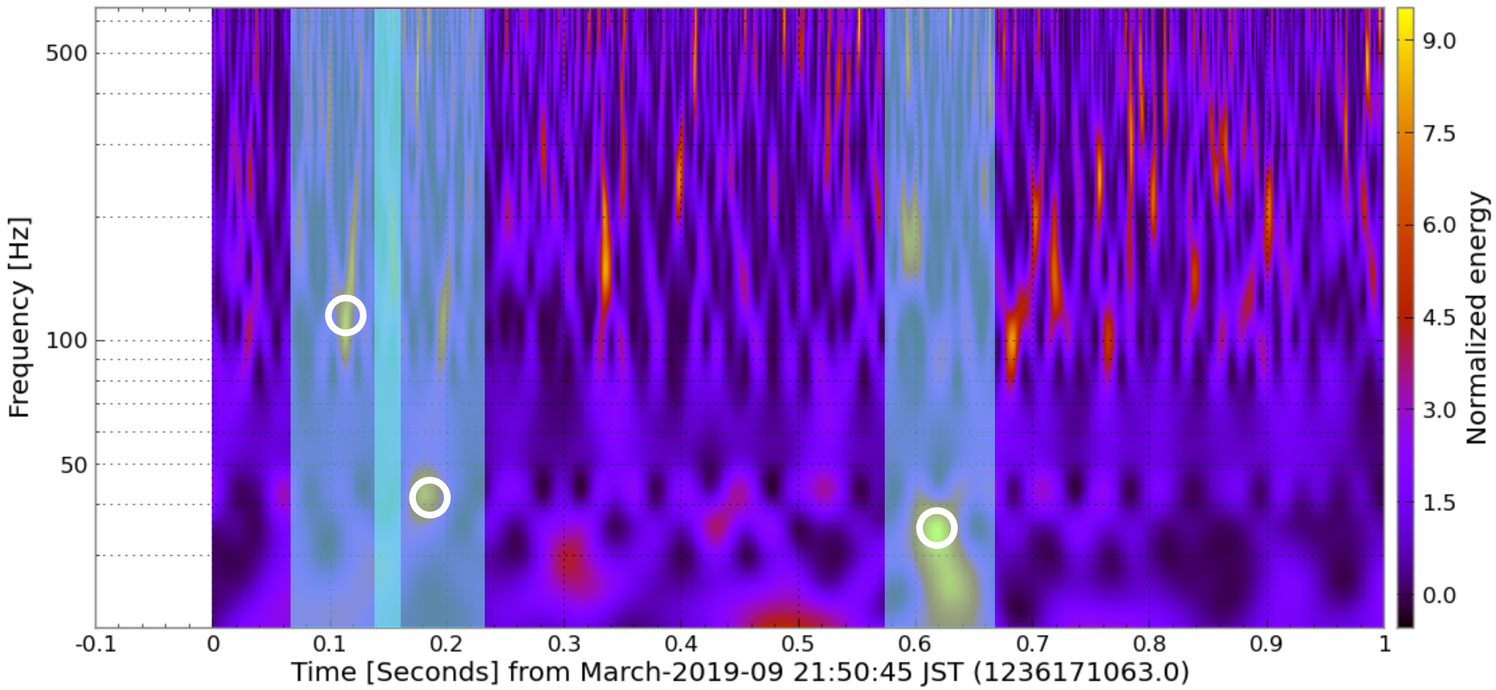
↓

Save OmegaScan plots

↓

Classify class

- Dataset with coherence information



Loudest trigger search conditions

time_window : 0.1 seconds

snr_threshold : 10

Loudest triggers

1236171063.61992 36.904296875 9.186349868774414 K1:PSL-PMC_MIXER_MON_OUT_DQ

1236171063.11992 42.904296875 9.186349868774414 K1:PSL-PMC_MIXER_MON_OUT_DQ

1236171063.01992 109.904296875 9.186349868774414 K1:PSL-PMC_MIXER_MON_OUT_DQ

Run h-veto

↓

Get vetoed triggers
and coherence information

↓

Find loudest triggers
without duplication

↓

Make table (.ecsv)

↓

Save OmegaScan plots

↓

Classify class

- Dataset with coherence information

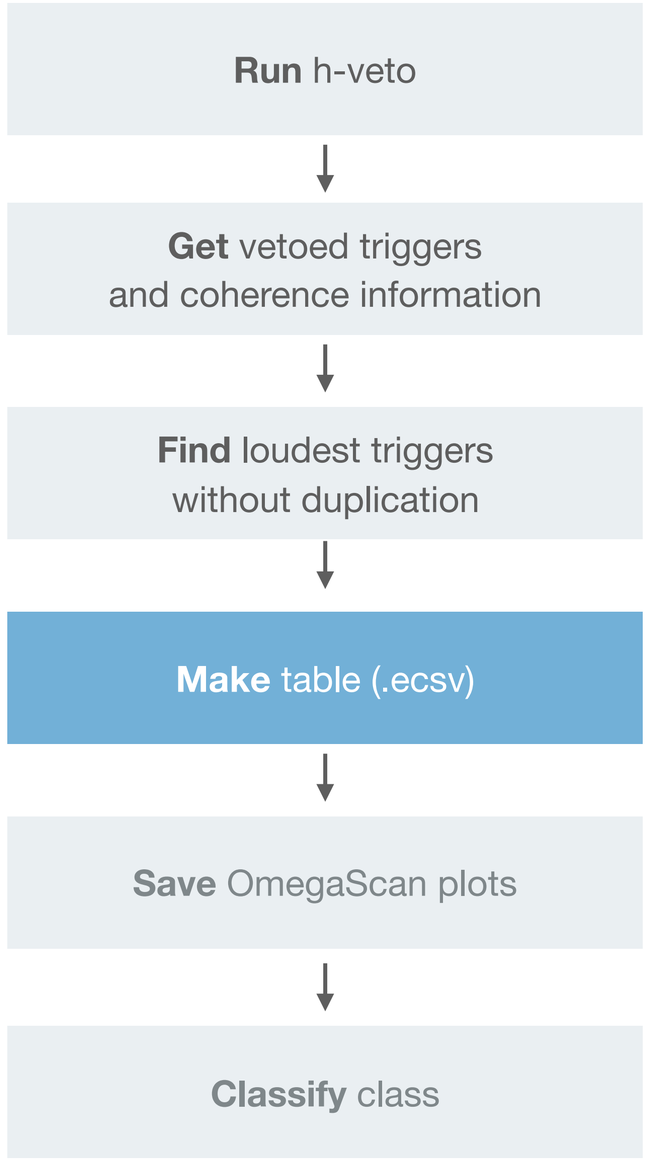
```
<Table length=1741>
round       primary_channel              aux_channel             peak_time           snr
 str2            str27                      str26                 float64          float64
----- ------------------------- ------------------------- ----------------- ------------------
   R1 K1:PSL-PMC_MIXER_MON_OUT_DQ K1:PSL-PMC_FAST_MON_OUT_DQ 1235778802.443847  479.7186279296875
   R1 K1:PSL-PMC_MIXER_MON_OUT_DQ K1:PSL-PMC_FAST_MON_OUT_DQ 1235778802.515625 114.83805084228516
   R1 K1:PSL-PMC_MIXER_MON_OUT_DQ K1:PSL-PMC_FAST_MON_OUT_DQ 1235778802.442138 113.29849243164062
   R1 K1:PSL-PMC_MIXER_MON_OUT_DQ K1:PSL-PMC_FAST_MON_OUT_DQ 1235778802.438232  113.1568832397461
   R1 K1:PSL-PMC_MIXER_MON_OUT_DQ K1:PSL-PMC_FAST_MON_OUT_DQ 1235778802.435547 112.51071166992188
   R1 K1:PSL-PMC_MIXER_MON_OUT_DQ K1:PSL-PMC_FAST_MON_OUT_DQ 1235778802.457519 109.77850341796875
   R1 K1:PSL-PMC_MIXER_MON_OUT_DQ K1:PSL-PMC_FAST_MON_OUT_DQ 1235778802.421875 109.09343719482422
  ...                       ...                       ...               ...                ...
   R2 K1:PSL-PMC_MIXER_MON_OUT_DQ K1:PSL-PMC_TRANS_DC_OUT_DQ 1235778803.03125 11.716090202331543
   R2 K1:PSL-PMC_MIXER_MON_OUT_DQ K1:PSL-PMC_TRANS_DC_OUT_DQ   1235778801.875 10.938090324401855
   R2 K1:PSL-PMC_MIXER_MON_OUT_DQ K1:PSL-PMC_TRANS_DC_OUT_DQ 1235778801.90625 10.931730270385742
   R4 K1:PSL-PMC_MIXER_MON_OUT_DQ K1:PSL-PMC_TRANS_DC_OUT_DQ  1235778801.6875 13.964690208435059
   R4 K1:PSL-PMC_MIXER_MON_OUT_DQ K1:PSL-PMC_TRANS_DC_OUT_DQ  1235778803.3125 12.295809745788574
   R4 K1:PSL-PMC_MIXER_MON_OUT_DQ K1:PSL-PMC_TRANS_DC_OUT_DQ 1235778803.15625 11.571940422058105
   R4 K1:PSL-PMC_MIXER_MON_OUT_DQ K1:PSL-PMC_TRANS_DC_OUT_DQ  1235778801.5625  10.33827018737793
   R4 K1:PSL-PMC_MIXER_MON_OUT_DQ K1:PSL-PMC_TRANS_DC_OUT_DQ 1235778848.453125 10.152999877929688

In [678]: t.meta
Out[678]: OrderedDict([('snr_threshold', 10), ('timewindow', 0.1)])
```

Loudest trigger data table

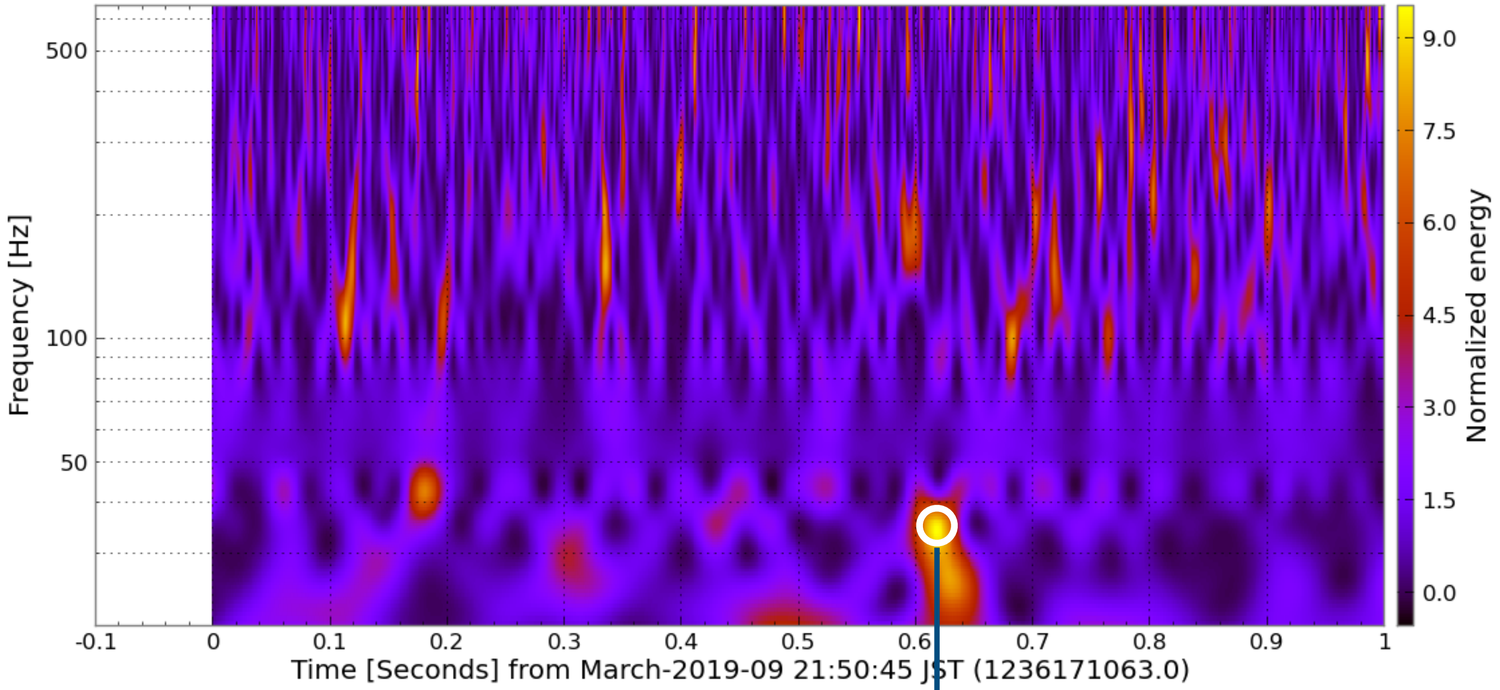
Save to “LoudestTriggers_Table_{date}.ecsv” file

But, Which format is suitable for saving this data?

XML, root or hdf5?



- Dataset with coherence information



CPU 1 duration 0.1s

CPU 2 duration 0.2s

CPU 3 duration 0.5s

CPU 4 duration 1.0s

CPU 5 duration 2.0s

CPU 6 duration 4.0s

Run h-veto

↓

Get vetoed triggers and coherence information

↓

Find loudest triggers without duplication

↓

Make table (.ecsv)

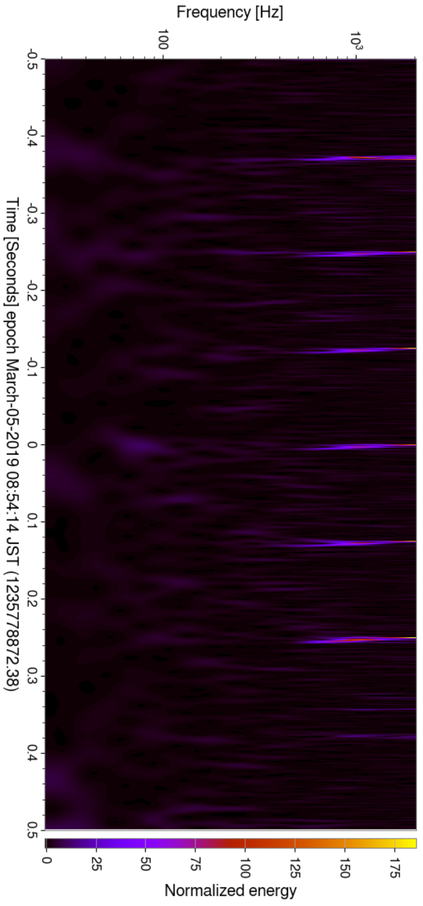
↓

Save OmegaScan plots

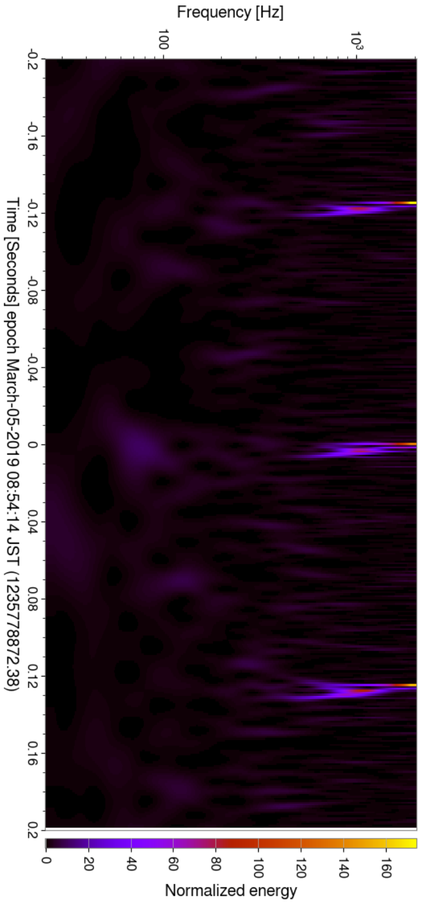
↓

Classify class

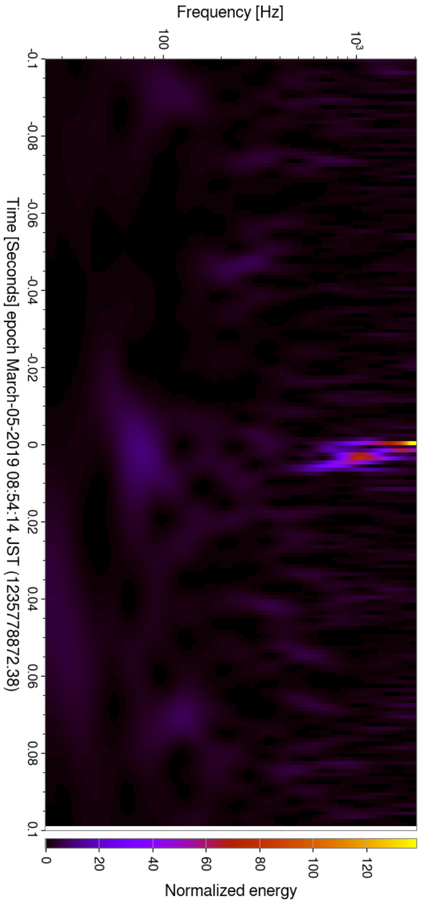
duration : 0.5s



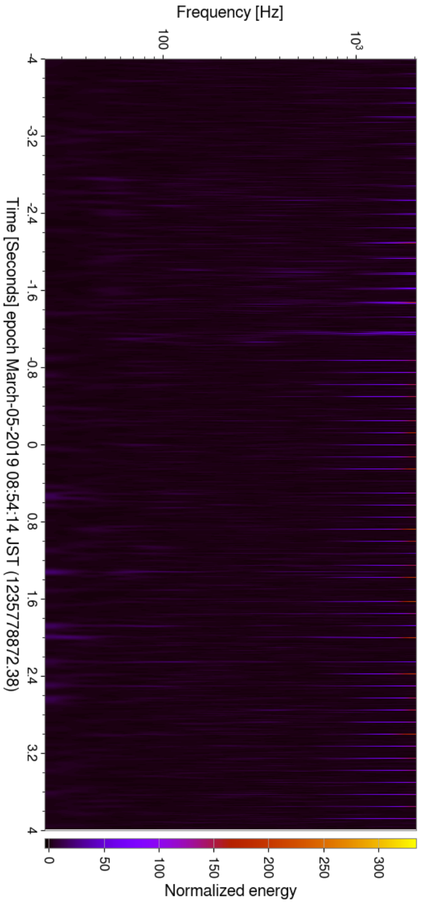
duration : 0.2s



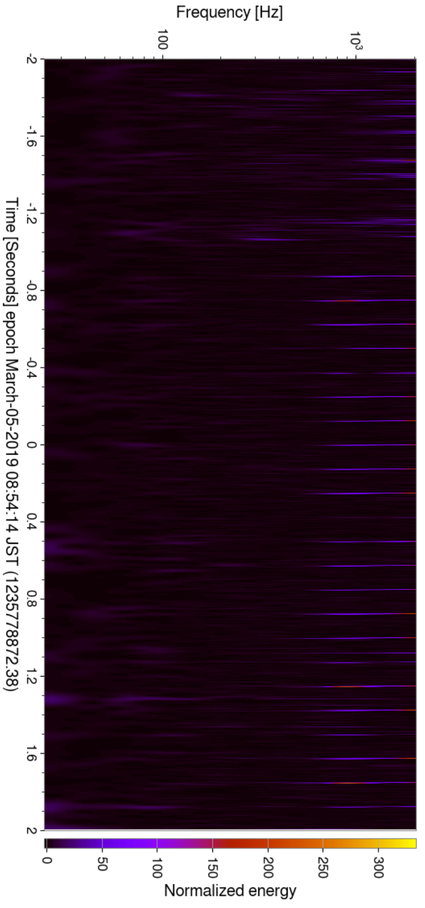
duration : 0.1s



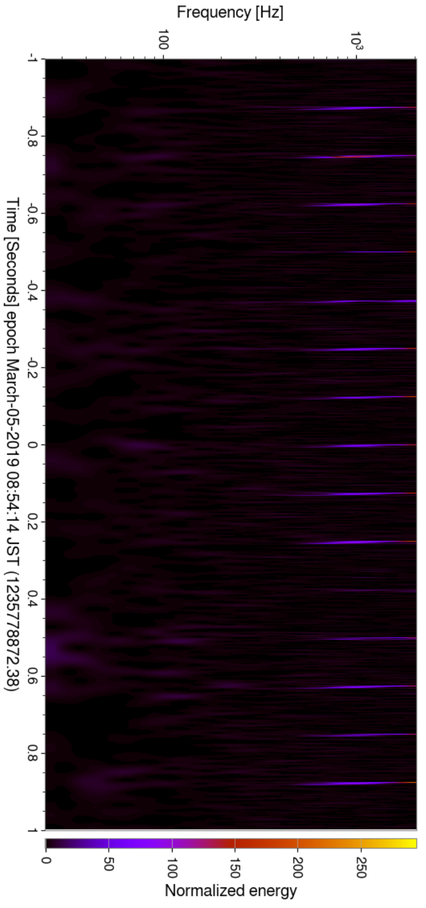
duration : 4.0s



duration : 2.0s



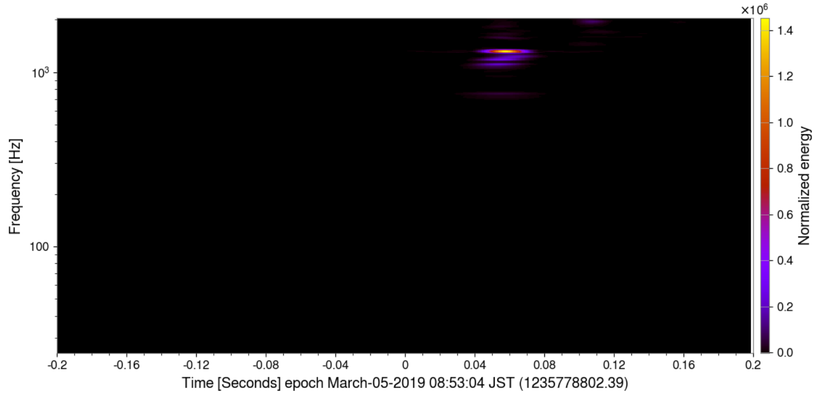
duration : 1.0s



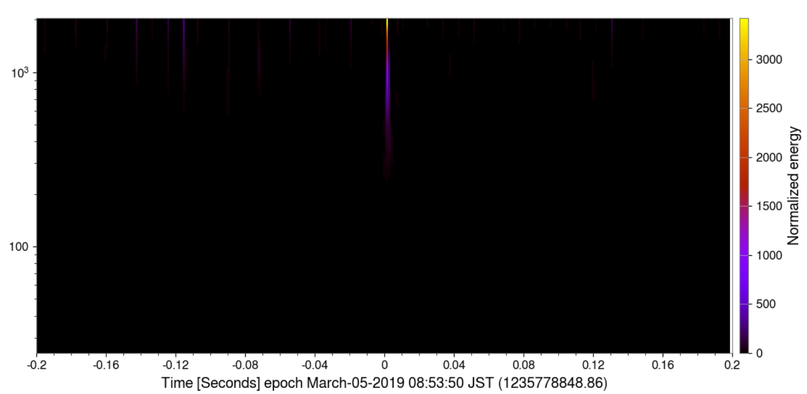
Channel : K1:PSL-PMC_FAST_MON_OUT_DQ

- Dataset with coherence information

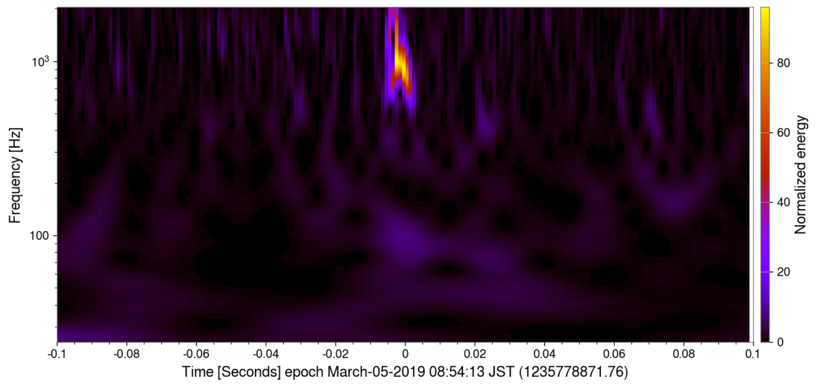
plate like



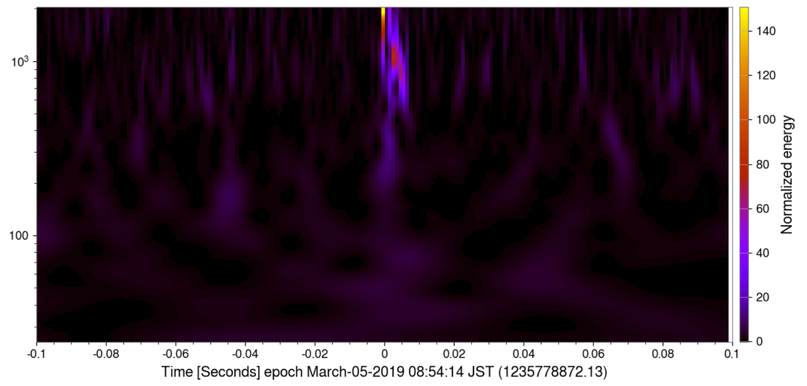
blip



“h” like



flag like



Run h-veto

↓

Get vetoed triggers
and coherence information

↓

Find loudest triggers
without duplication

↓

Make table (xml or h5)

↓

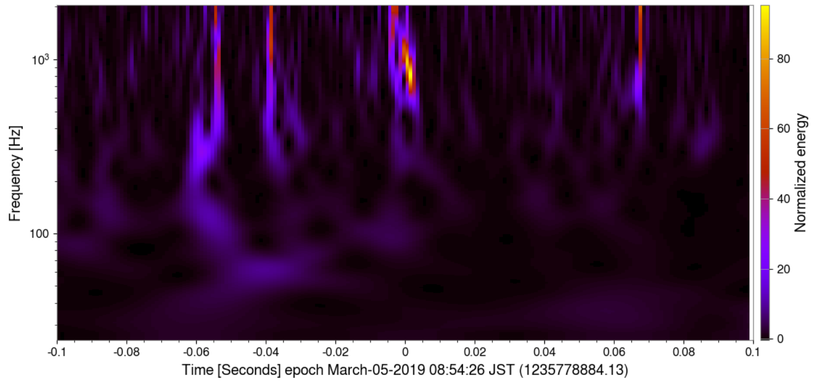
Save OmegaScan plots

↓

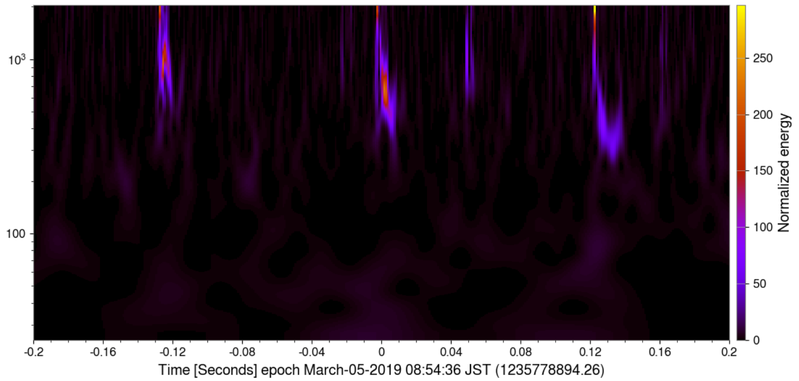
Classify class

- Dataset with coherence information

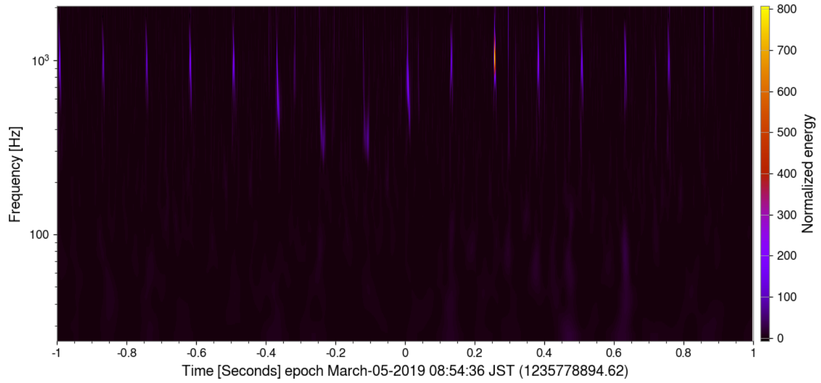
blip like



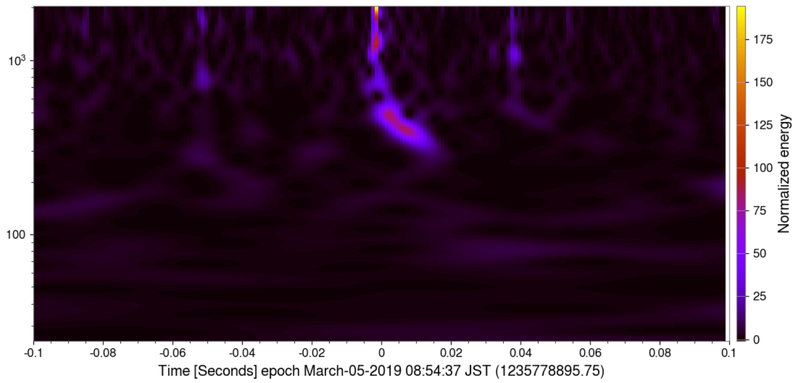
blip like



oscillating blips



fish hook like



Run h-veto

↓

Get vetoed triggers
and coherence information

↓

Find loudest triggers
without duplication

↓

Make table (xml or h5)

↓

Save OmegaScan plots

↓

Classify class

mini summary

- Algorithm Test to sort Loudest triggers in order
- Algorithm Test to plotting OmegaScan automatically
- Trigger and correlation information from h-veto can be saved as a table

beyond

- Bug Chenking
- Collect a sufficient number of plot data
- Separate class
- CNN running test
- Discuss other data storage formats other than png (h5, hdf5, roof and so on)

Project naming

機械学習を用いた

グリーチ

分類

Glitch Classification with Machine Learning

- Glitch-Ninja
- etc...