

Cleaning procedure analysis through Hough Transform
for All Sky search for continuous waves

University of Florida International Research Experience
for Undergraduate

Rome summer 2015

Han Soul Lee (Florida Atlantic University)

Pia Astone (INFN, “la sapienza” University of Rome)

Abstracts

In order to remove unwanted artifacts injected in data from gravitational wave detectors, cleaning procedures were applied. The first filtering process, removal of time domain glitches was applied when fast fourier transform was applied to the data. Removal of spectral wandering lines and removal of spectral lines of constant frequency were then applied. The cleaning efficiency was compared between detectors through histograms of difference vetoed signal data and through Mock data challenge. Average 3 percent of peaks were detected as artifacts and removed. Higher concentration of noise and stronger signal to ratio was detected from Hanford interferometer.

Introduction

According to Einstein's theory of general relativity, gravitational wave is predicted to exist as a byproduct of curvature of space and time. Though the existence of gravitational wave is highly expected through renounced theory and indirect observational evidence, the gravitational wave have not been able to directly detect from earth. In attempt to detect the gravitational wave, laser interferometers such as VIRGO and LIGO have been constructed in various locations. Although these detectors are located in remote locations, unwanted signals such as environmental, mechanical and non-gravitational wave signals are contained in data. Therefore it is very important to apply the proper cleaning procedure to veto unwanted signals.

Source of Continuous waves

Neutron stars are expected to be the source of continuous waves. Neutron stars, which has a very high density and comparable size of a city has a very rapid rotational period. Not every neutron stars are predicted to be emitting gravitational wave. Continuous gravitational wave signals (CW) are emitted by asymmetrically rotating neutron stars. The asymmetry, which is defined by *ellipticity* of the star are caused by various mechanisms such as non-axisymmetric residual from the star's birth or a strong internal magnetic field not aligned to the rotational axis. The decrease in kinetic energy, which provides limit for the gravitational wave is called *spindown*, and *spinup* for vice versa. The spin-down, which is defined by star's parameters such as moment of inertia, rotational frequency, derivative of the rotational frequency, etc. sets the upper limit for the gravitational waves. The most probable gravitational wave frequency from the neutron stars is twice of its own rotational frequency.

Targeted and all sky search

Search for continuous wave can be roughly categorized into two sections: targeted and all sky search. Targeted search is performed for the source with parameters such as source positions, spindown, rotational frequency, etc. are exactly known. All sky search is performed without any known parameters and searched through various locations. About $O(10^6 \sim 10^7)$ neutron stars are expected, while 2400 of them have been detected. Although all sky search was conducted for this research, pulsars with simulated parameters were used to test the cleaning procedure.

Frequency Hough Transform

In order to reduce the computational loads, hierarchical procedure was created. The obtained data is first corrected for Doppler shift. After the correction, short fast fourier transform was applied to the database (SFDB) with short duration time called coherence time. Cleaning procedures are applied multiple times through the transform. Removal of time domain glitches is applied when SFDB is created. Once the procedure is applied, removal of spectral wandering

lines and removal of spectral lines of constant frequency are applied as well. The database after these procedures are applied is considered as ‘cleaned’ data, compared to ‘noclean’ data with only raw SFDB. From the database, time and frequency map called the peakmap is created. The peakmap is then served as an input of the Hough Map. Hough Map is the linear mapping between the Doppler shift corrected frequency plane into the source intrinsic frequency and spindown plane.

Cleaning Procedures

Removal of time domain glitches is applied when SFDB are created. The glitches, which appears randomly and enhances the noise level in a wide frequency band, are removed through applying highpass bilateral to the data. For the purpose of this research, this data was considered as NOCLEAN data.

Spectral Wandering Lines Removal (Time-Frequency filter)

Histogram of a low resolution is peakmap is constructed under the consideration of the possible presence of the continuous wave. Threshold is decided from the peakmap, and the noises presented as local maxima or minima above the thresholds are removed. Figure 1 and 2 demonstrates example of before and after the filtering.

Spectral Constant Frequency Removal

Persistency filtering is applied. The entire frequency is histogrammed. Statistical property of the graphs such as average and standard deviation is analyzed. The disturbed frequencies are determined and the frequency bin is removed from the data.

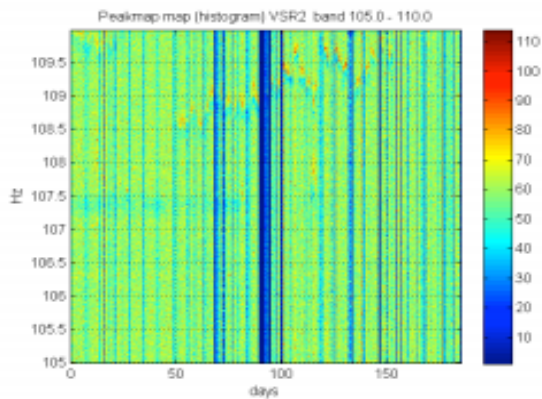


Figure 1. Peakmap before the wandering lines removal (time-frequency filtering) Wandering line is vividly on the top middle of the histogram. (left)

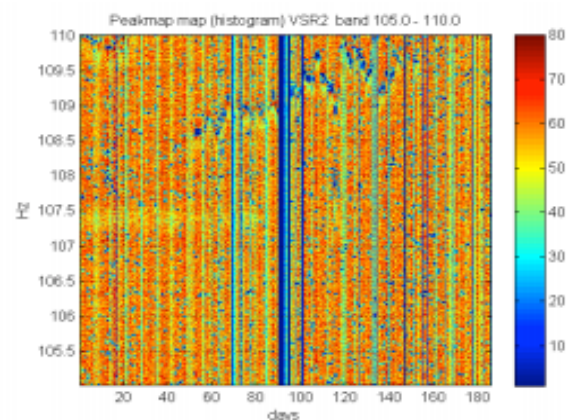


Figure 2. Peakmap after wandering lines removal(time-frequency filtering) (right)

The wandering line from figure 1 is now removed, and peaks throughout the graph is higher than the removed noises

Analysis

For consistency and accuracy purposes to detect the gravitational wave, data obtained from two detectors were analyzed. The cleaning procedure has been previously applied and analyzed for VIRGO. While VIRGO has better sensitivity for low frequency region, LIGO is expected to have better sensitivity for the higher frequency region. The first job I was asked to perform was to modify the existing code that has been applied to VIRGO and personalize it for LIGO data. The efficiency of detector was first determined by analyzing characteristics of vetoed signals from the filtering process. Two data set were obtained from LIGO Livingston and Hanford. The analysis is planned to be performed for 40~2000 hz. While data from both detectors initiated from 40 hz, frequencies up to 1505hz was prepared for analysis for Livingston and 1025 hz for Hanford.

Peaks

Livingston

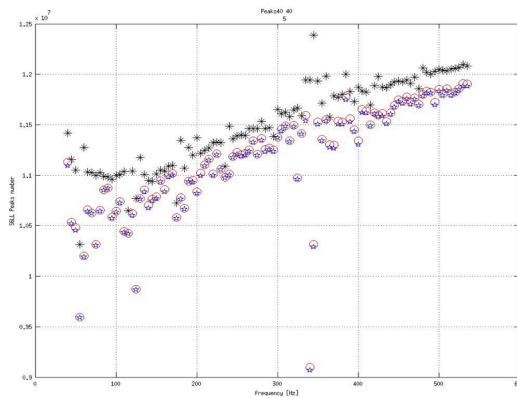


Figure 3.Livingston:Total count of peaks before and after the filtering(40~540 hz) (left)

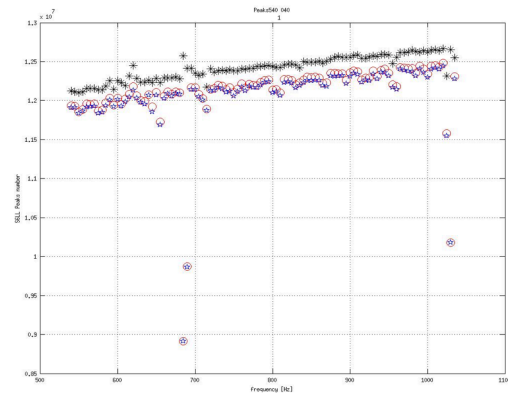


Figure 4. Livingston:Total count of peaks before and after the filtering(540~1505 hz) (right)

Figure 3

1150698640 total peaks (*)

1117779732 after t-f filter (o)(0.971)

1116098868 after persistence filter(☆)(0.970)

Total time vetoed 14.44 days

Figure 4

1241768478 total peaks in the peak map(*) (expected 9.086183e+09)

1210412255 after t-f filter (0.975)(o)

1207154768 after persistence filter(☆) (0.972)

Total time vetoed 18.3 days

Hanford

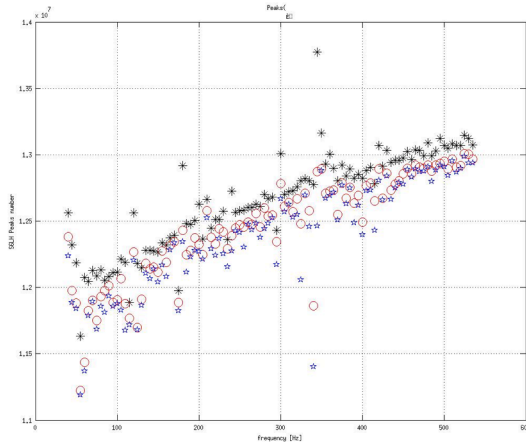


Figure 5. Hanford: Total count of peaks before and after the filtering (40~540 Hz) (left)

1265227333 total peaks in the peak map (*) (expected 1.245593×10^{10})

1246325456 after t-f filter (o) (0.985)

1238909194 after persistence filter (☆) (0.979)

Total time 11.95 days

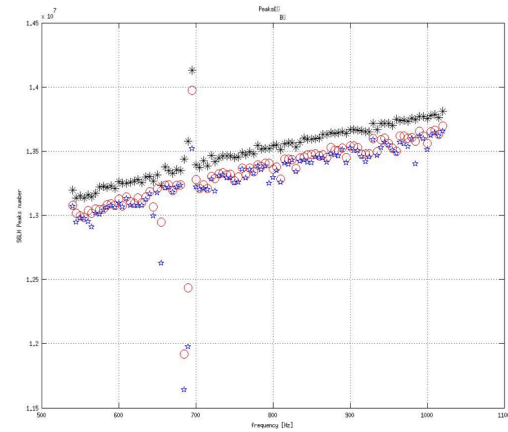


Figure 6. Hanford: Total count of peaks before and after the filtering (540~1025 Hz) (right)

1309492329 total peaks in the peak map (*) (expected 1.245593×10^{10})

1292977532 after t-f filter (o) (0.987)

1288436527 after persistence filter (☆) (0.984)

Total time 10.52 days

Both data obtained from Hanford and Livingston are demonstrating very similar patterns, where the slope of the graph is more gradual and stable for higher frequency region. More randomness and deviations are found from lower frequencies, where filtering had more dramatic effects. Though two detectors do demonstrate similar patterns, higher number of peaks are found from Hanford data. Average 3 percent of peaks were removed from the original peaks.

Histogram of time-frequency vetoed lines

The graph demonstrates the raw counts of vetoed signals per frequency.

Livingston

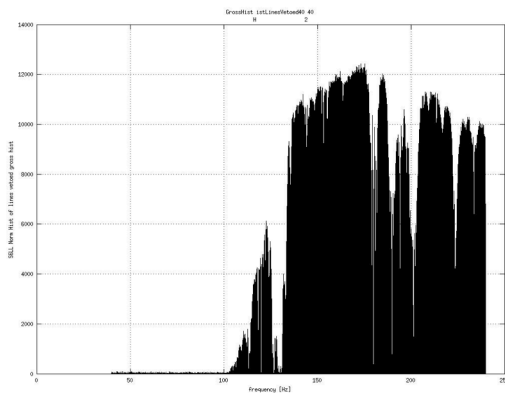


Figure 7 Liv. count of vetoed signals from 40~240 hz (left)

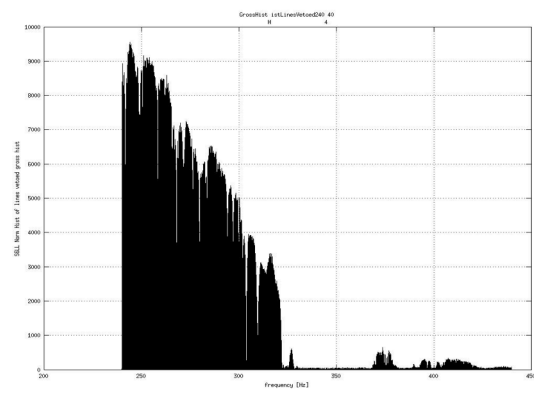


Figure 8 Liv. count of vetoed signals from 240~440 hz (right)

The most disturbed frequency region with high counts of vetoed signals were found.

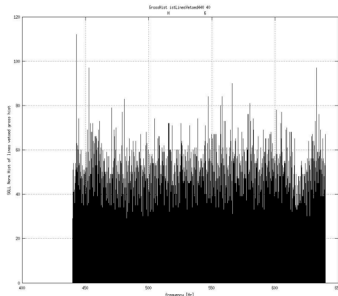


Figure 9 Liv. count of vetoed signals from 440~640 hz (left)

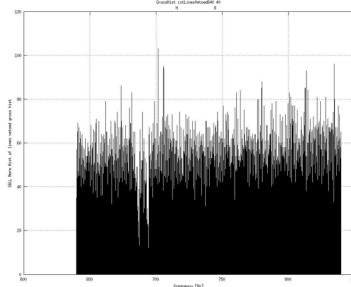


Figure 10 Liv. count of vetoed signals from 640~840 hz (middle)

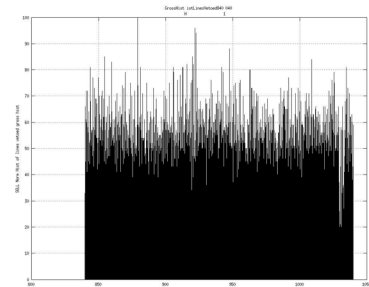
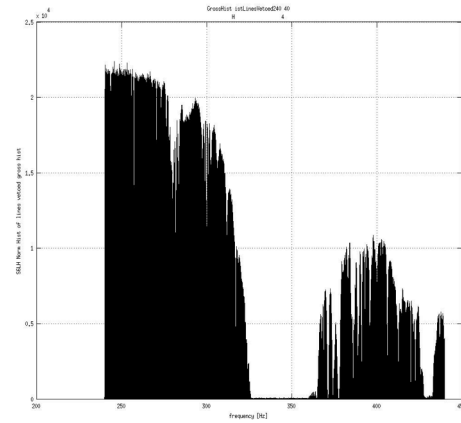
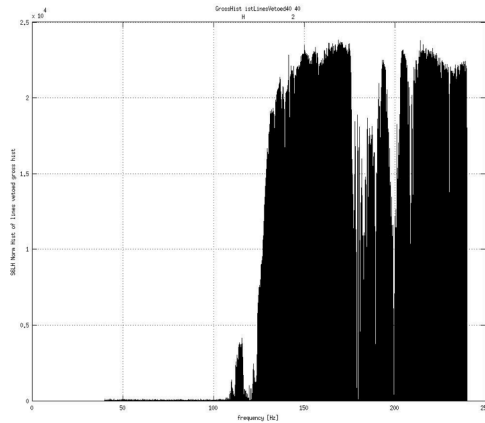


Figure 11 Liv. count of vetoed signals from 840~1040 hz (right)

Comparably minor counts of vetoes were found in high frequency region.

Hanford



The most disturbed frequency region with high counts of vetoed signals were found between 100hz to 340 hz. The second most vetoes came from 360~540 hz.

Figure 12 Han. count of vetoed signals from 40~240 hz (left)

Figure 13 Han. count of vetoed signals from 240~440 hz (right)

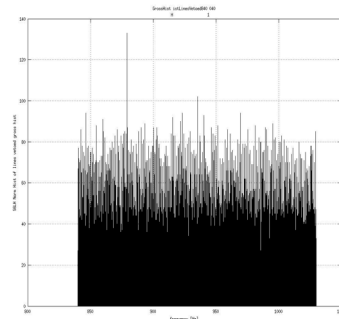
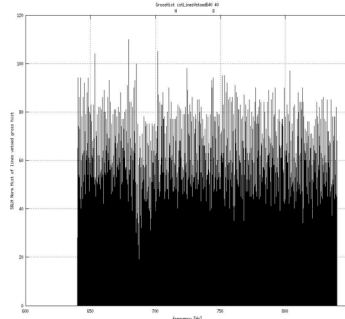
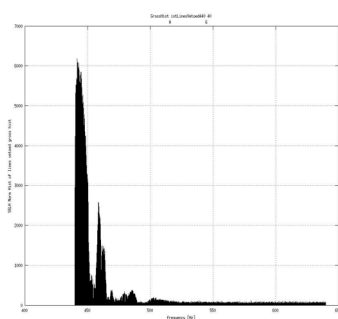


Figure 14 Han. count of vetoed signals from 440~640 hz (left)

Figure 15 Han. count of vetoed signals from 640~840 hz (middle)

Figure 16 Han. count of vetoed signals from 840~1025 hz (right)

Though two detectors do demonstrates similar pattern for their most disturbed frequency region (100~340 hz) approximately twice as much vetoes are detected from Hanford data. Additionally, the difference between counts from for the second most disturbed region (375hz~425) has more drastic

difference. While the vetoes from this region from Hanford is at most 800, the count is found near 10,000 for Hanford.

Time-Frequency vetoed Peaks

The graphs are as a peakmap for the vetoed frequencies. The amount of signals peak vetoed are indicated throughout the runtime. White band which occasionally appears on the graphs are the frequency bands that have been determined as noise and been removed from the filtering.

Livingston

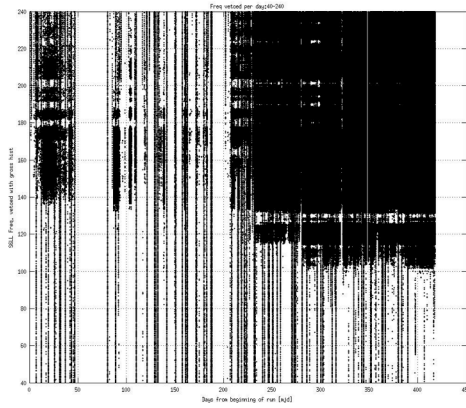


Figure 17 Liv.runtime vs vetoed signals (40~240hz)(left)

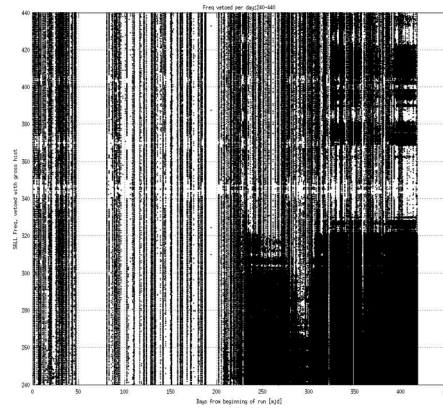


Figure 18 Liv.runtime vs vetoed signals (240~440hz)(right)

As indicated from the previous graphs, heavy vetoes are coming from 100hz to 340 hz region, with some less severe vetoes are also found from 370~420hz. It is interesting to note that these vetoes were mostly found after runtime day 200.

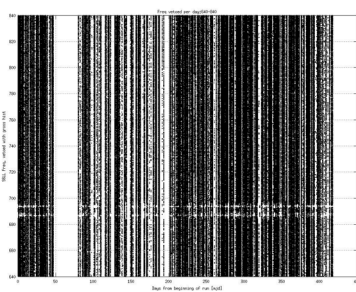


Figure 19 Liv.runtime vs vetoed signals (440~640hz)(left)

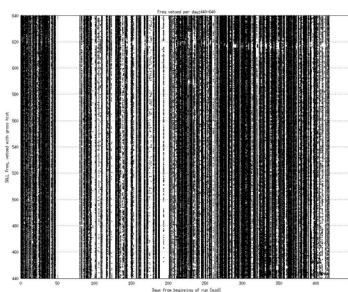


Figure 20 Liv.runtime vs vetoed signals (640~840hz)(middle)

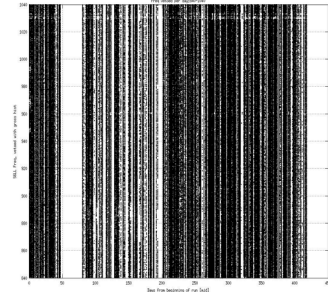


Figure 21 Liv.runtime vs vetoed signals (840~1040hz)(right)

The higher frequency regions have constant vetoed peaks throughout the runtime.

Hanford

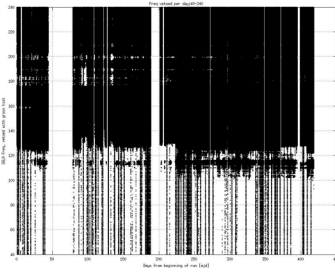


Figure 21 Han. runtime vs vetoed signals (40~240hz)(left)

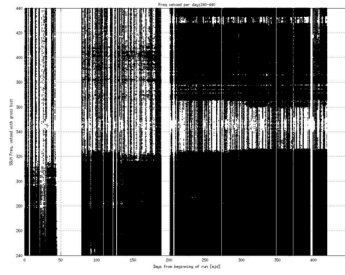


Figure 22 Han. runtime vs vetoed signals (240~440hz)(middle)

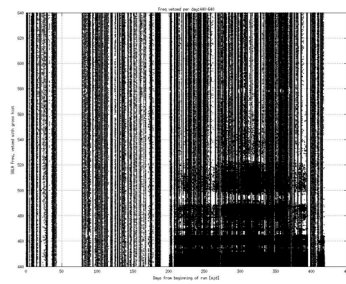


Figure 23 Han. runtime vs vetoed signals (440~640hz)(right)

Though the same frequency region is detected with artifacts, the intensity of veto is greater for Hanford compared to livingstone. The most disturbed region, 100~340 hz has constant vetos detected throughout the runtime. The second most disturbed region, 360~530 hz has comparably intense vetoes detected after the runtime day 200. These are the distinctive features for data from Hanford detector, where both detectors do share the same characteristics while more intense results are detected from Hanford. More noise seems to be contained in Hanford data.

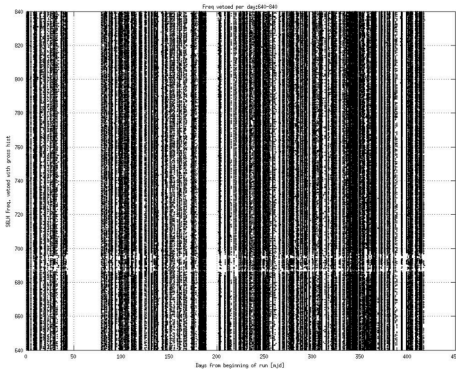


Figure 24 Han. runtime vs vetoed signals (640~840hz)(left)

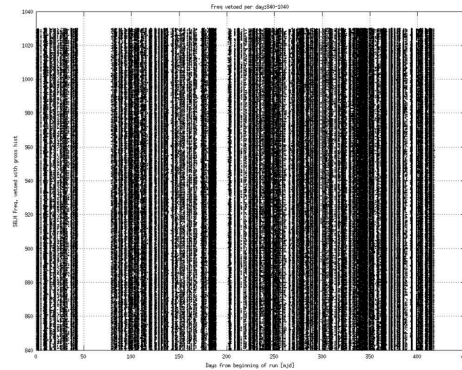


Figure 25 Han. runtime vs vetoed signals (840~1025hz)(right)

Again, higher frequency region has stable constant veto throughout the runtime.

Mock Data Challenge

To analyze and visualize the effect of the filtering around the significant signals, mock data challenge was performed for both of the detectors. Two types of injections were tested into the data.

Hardware injection

Hardware injection is a mechanical noise injected into the data through detector environment O(4) simulated pulsars were tested for hardware injection to compare the results before and after cleaning. Each of the pulsars had different parameters such as rotational frequency, expected gravitational wave frequency and frequency derivative.

Livingston

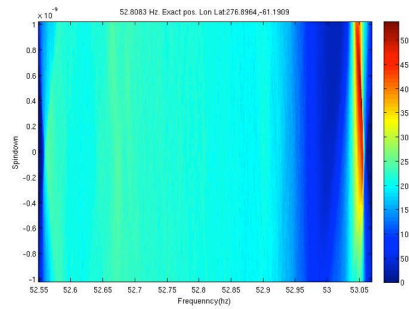


Figure 26 expt injection: 52.8083 Hz NOCLEAN(left)

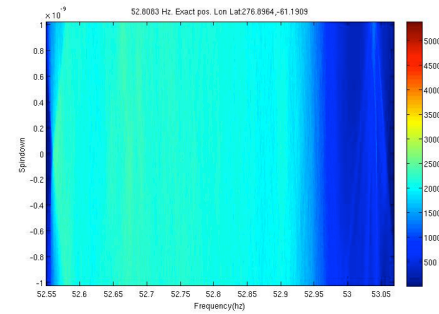


Figure 27 expt injection: 52.8083 Hz CLEAN(right)

Though injection isn't presented on either of the graphs, disturbances presented on the NOCLEAN graph have been removed after the cleaning procedure.

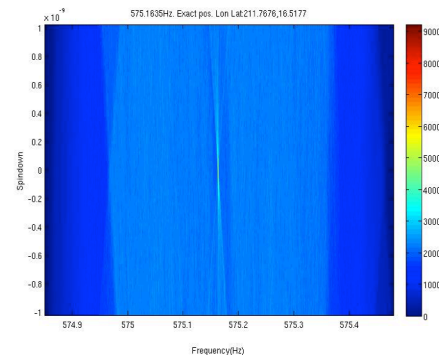
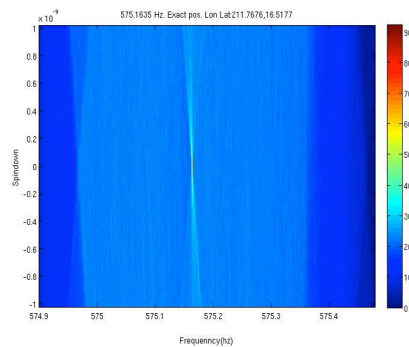


Figure 28 expt injection: 575.1635 Hz NOCLEAN(left)

Figure 29 expt injection: 575.1635 Hz CLEAN(right)

The injection is clear around the spin down 0 before cleaning. There is a slight region with a high concentration of peaks around the injection presented in NOCLEAN graph.

Figure 28 expt injection: 575.1635 Hz NOCLEAN(left)

Figure 29 expt injection: 575.1635 Hz CLEAN(right)

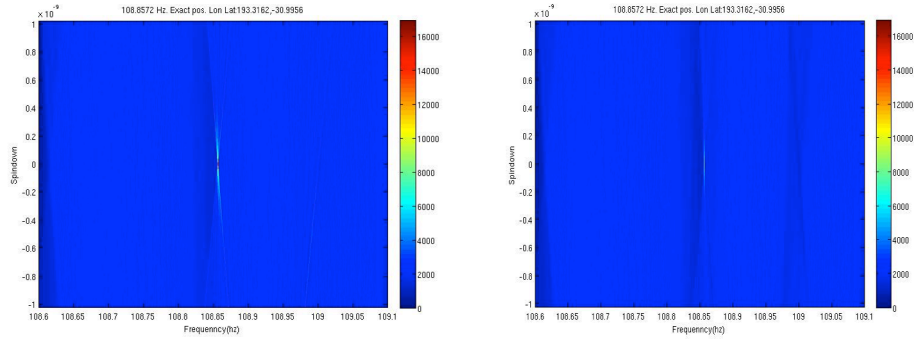


Figure 30 expt injection: 108.8572 Hz NOCLEAN(left)

Figure 31 expt injection: 108.8572 Hz CLEAN(right)

Injection presented in both graphs. Though the strength of signal is faded, noise has been obviously reduced. A slanted streak is removed after cleaning procedures.

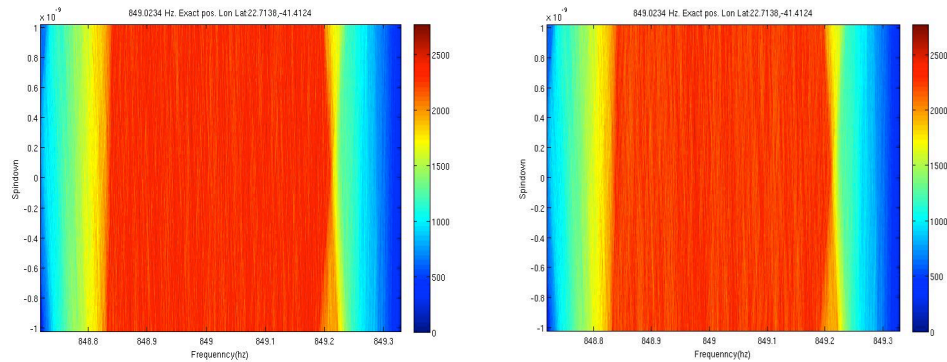


Figure 32 expt injection: 849.0234 Hz NOCLEAN(left)

Figure 33 expt injection: 849.0234 Hz CLEAN(right)

No injection presented from both graphs, color bar scales are lower compared to other pulsars. There seems to be a slight reduce in noise, though not dramatic.

Frequency vs. Critical Ratio

Critical ratio, which is the ratio between signal and noise was measured for each injections found on the Hough map. Critical ratio found on each maps were compared before and after cleaning.

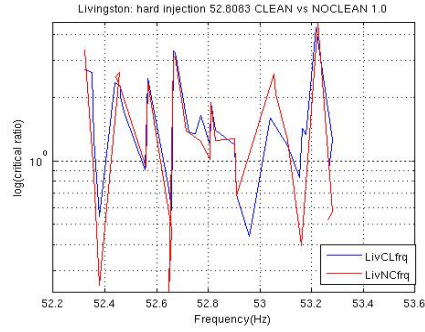


Figure 34 Frequency vs. Critical Ratio 52.8083hz (left)

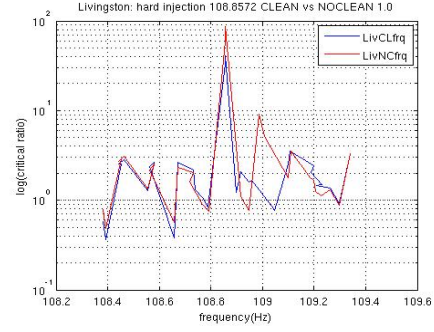


Figure 35 Frequency vs. Critical Ratio 108.8572hz (right)

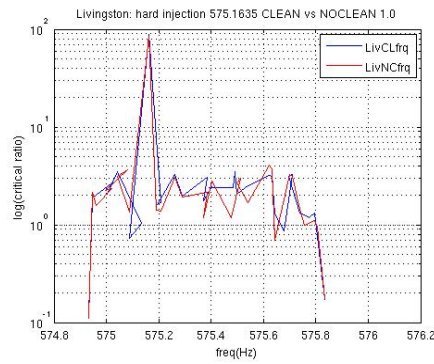


Figure 36 Frequency vs. Critical Ratio 575.1635hz (left)

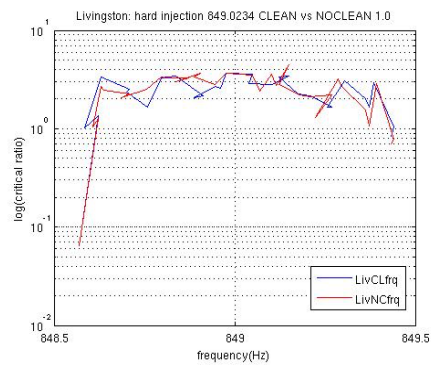


Figure 37 Frequency vs. Critical Ratio 849.0234hz (right)

Critical ratio was generally higher before cleaning. Though filtering procedures are reducing the noise, it is also weakening the strength of signals as well.

Hanford

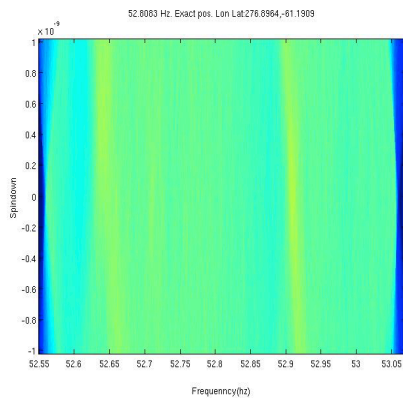


Figure 26 expt injection: 52.8083 Hz NOCLEAN(left)

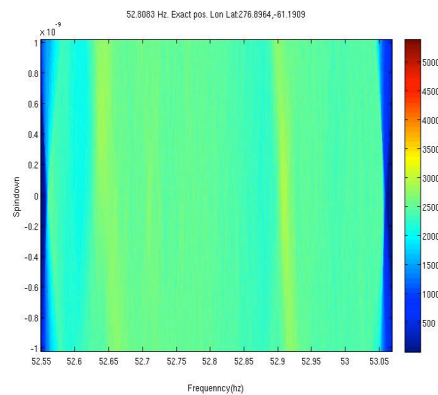


Figure 27 expt injection: 52.8083 Hz CLEAN(right)

No presence of signal. The artifact presented in Livingston is not included in Hanford data. Very minor difference between before and after cleaning.

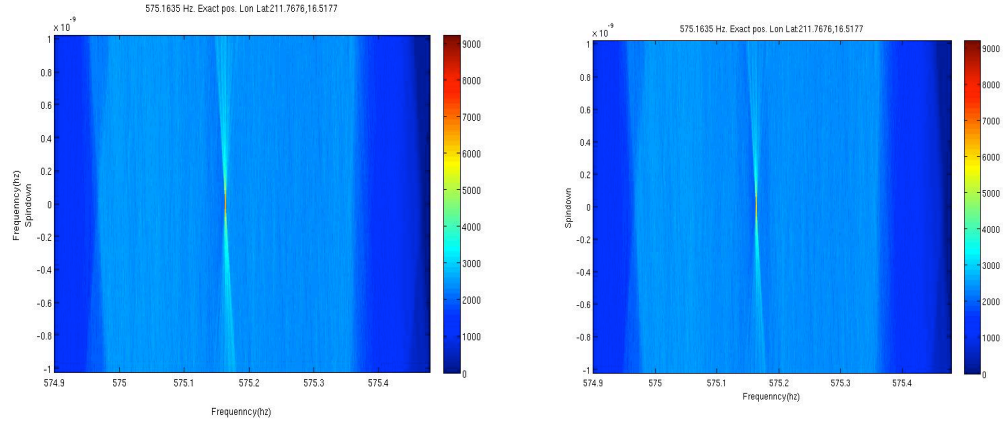


Figure 28 expt injection: 575.1635 Hz NOCLEAN(left)

Figure 29 expt injection: 575.1635 Hz CLEAN(right)

Presence of signals are strong before and after cleaning. Though the strength of signal is vivid, the streaks above and below the injection have been weakened after filtering.

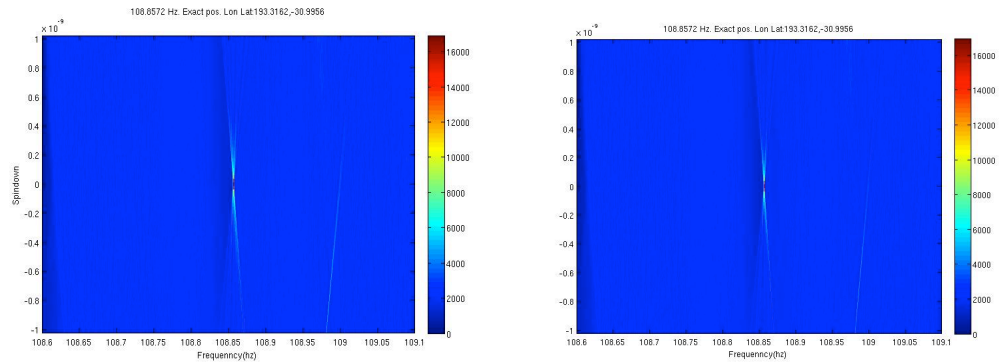


Figure 30 expt injection: 108.8572 Hz NOCLEAN(left)

Figure 31 expt injection: 108.8572 Hz CLEAN(right)

Though artifacts presented as a streak on the right side of the graphs is still presented after the filterings, noise steak around the injection has been reduced.

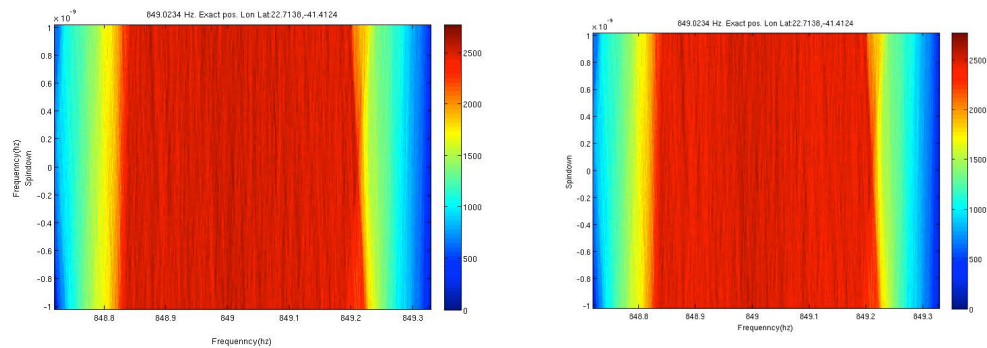


Figure 30 expt injection: 849.0234 Hz NOCLEAN(left)

Figure 31 expt injection: 849.0234 Hz CLEAN(right)

No presence of signal nor major effectivity. Very minor difference.

Injectons presented on the Hanford data is stronger, though noise have not been removed as successfully for Hanford, effectivity is still detected from the hough map.

Frequency vs. Critical Ratio

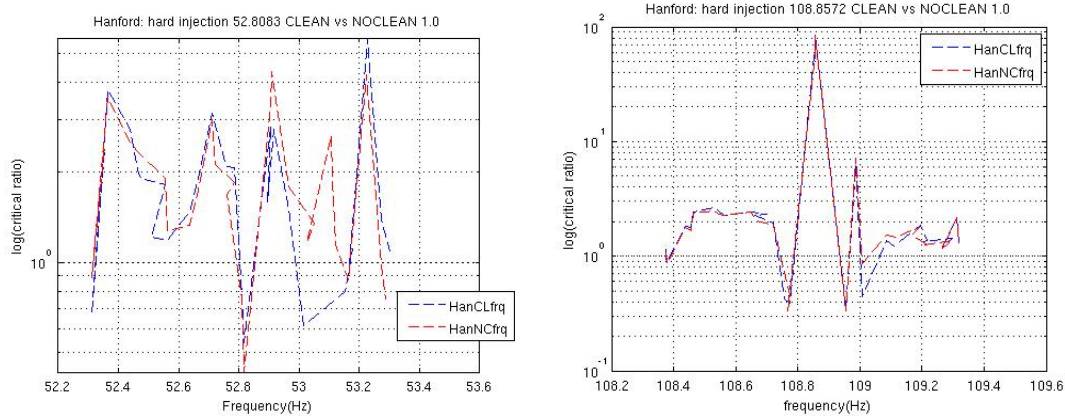


Figure 34 Frequency vs. Critical Ratio 52.8083hz (left)

Figure 35 Frequency vs. Critical Ratio 108.8572hz (right)

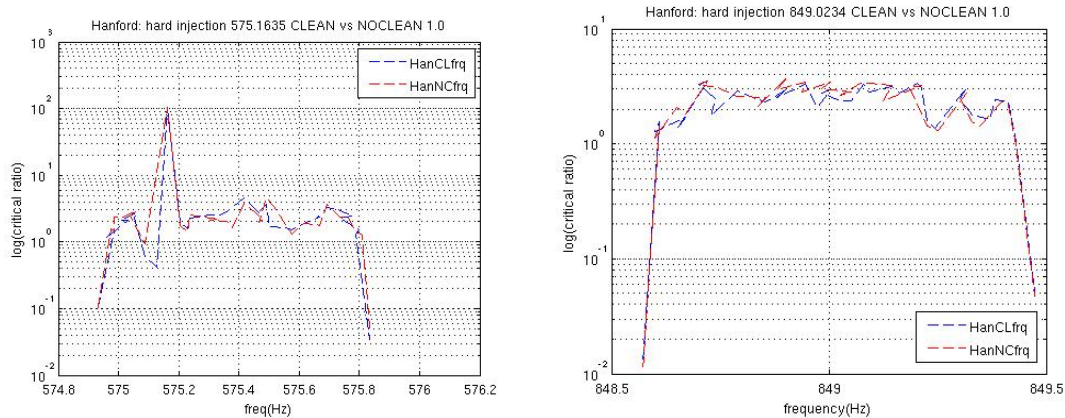


Figure 36 Frequency vs. Critical Ratio 575.1635hz (left)

Figure 37 Frequency vs. Critical Ratio 849.0234hz (right)

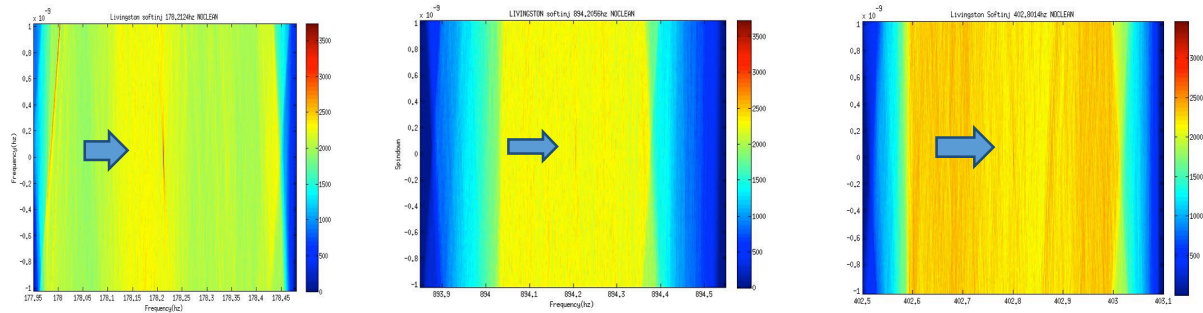
Again, the critical ratio is generally higher for NOCLEAN.

Software Injection

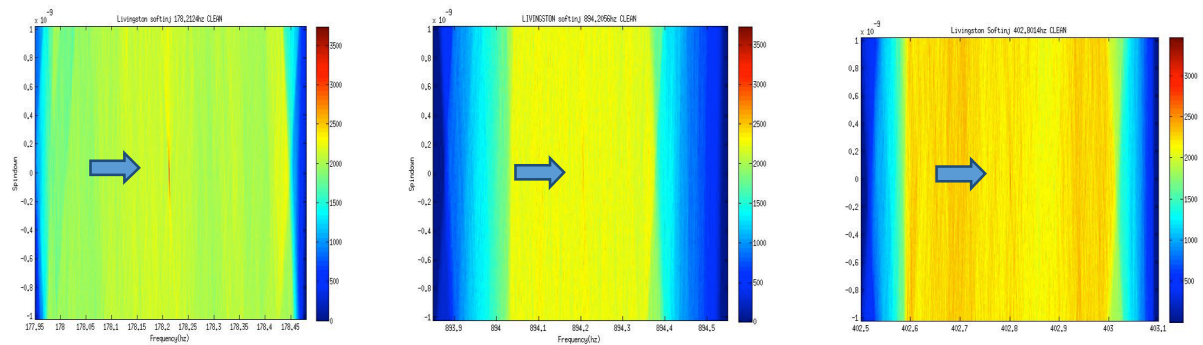
Software injection is an artificial injection created by software that has been injected into obtained data. Unlike hardware injections which are part of the data, software injection is injected after the data has been obtained. Artificial noises are injected throughout the frequency region. The cleaning results were compared before and after cleaning.

Livingston

Before cleaning:

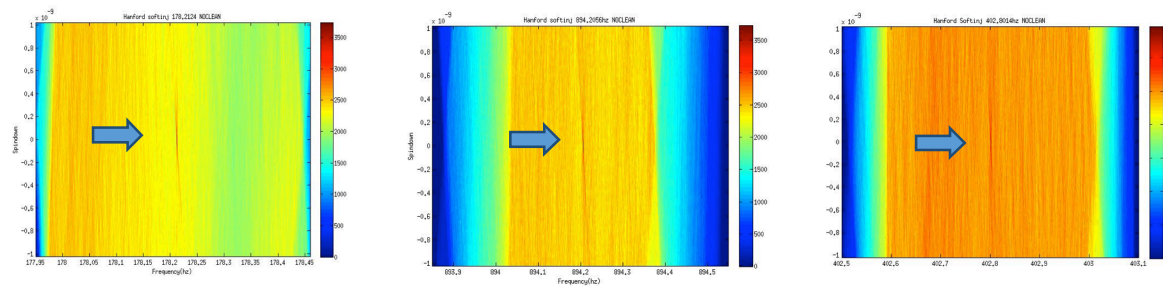


After cleaning:

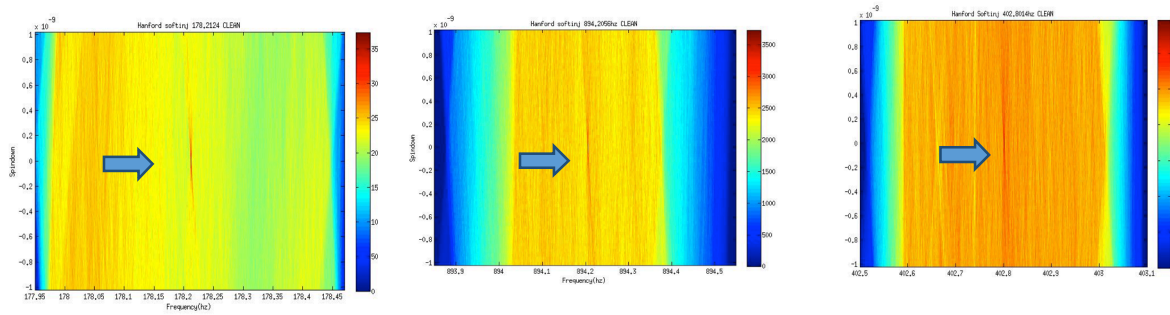


Hanford

Before cleaning:



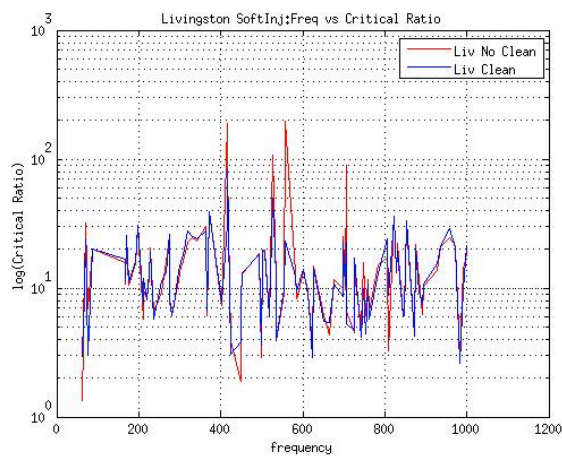
After cleaning:



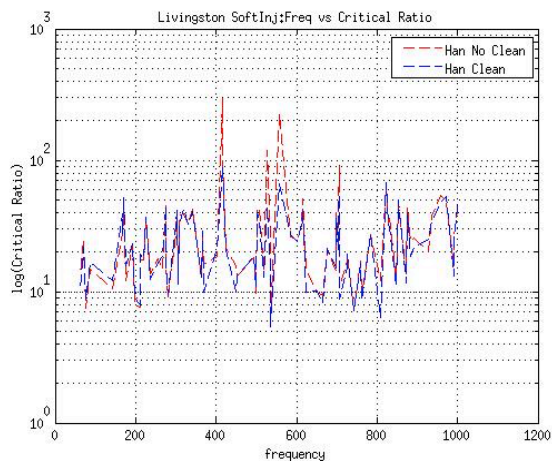
For both of the detectors, injections are well presented on the graphs. Filtering efficiencies are well demonstrated as there are visual difference in peak counts after the filtering procedures. Though less noises are detected from Livingston, Hanford, along with higher concentration of noises has stronger signals detected from the map.

Frequency vs Critical Ratio

Livingston



Hanford



As expected from previous analysis, the critical ratio is generally higher for Hanford. NOCLEAN data generally presents higher critical ratio, weakening the noises level as well as injections. Though two

graphs shows similar patterns from each other, they do demonstrate different characteristics. The difference is assumed to be caused by different noise ratio between two detectors.

Conclusions

Frequency region 135~325 hz seems to be the most disturbed regions where heavy concentration of signals are being vetoed through filtering. However, frequency region under 100hz and above 650 hz only returns minor count of veto signals. This result was confirmed from both Hanford and Livingston detectors. Though they demonstrated the similar patterns, more extreme results were detected from Hanford detector, where the most disturbed frequency region had over twice more vetoed counts compared to Livingston. Second most disturbed region, 375~480 hz had upto 15 times more veto counts compared to Livingston. The vetoes from second most disturbed lines were detected after the runtime day 200. This was also true for the most disturbed frequency region of Livingston data. Despite the high counts of noise detected from Hanford, stronger critical ratio was detected from Hanford. The injection signals analyzed through Hough map confirms this observation. Hanford injections are comparably stronger than Livingston despite the high concentration of noise. Though results from filtering did demonstrated some effects for both hardware and software injection, results did not demonstrate drastic difference. This result was expected from the analysis of counts of peaks after and before filtering, where about average 3 percent of peaks were removed from the process. The study was conducted to improve the efficiency for weak signals.

- [1] Astone, P., Colla, A., D'Antonio, S., Frasca, S., Palomba, C., & Serafinelli, R. (2014). Method for narrow-band search of continuous gravitational wave signals. *Physical Review D* *Phys. Rev. D* **89**, 062008, 13pp.
- [2] Astone, P., Colla, A., D'Antonio, S., Frasca, S. & Palomba, C. (2014). Method for all-sky search of continuous gravitational wave signals using the frequency-Hough transform. *Phys. Rev. D* **90**, 042002, 24pp.
- [3] LVC (2014) Narrow-band search of continuous gravitational-wave signals from Crab and Vela pulsars in Virgo VSR4 data, 17pp, [arXiv:1410.8310](https://arxiv.org/abs/1410.8310) .